

# Analyses multivariées

## des indicateurs de santé des régions de France



# Sommaire

Avant-propos de M. Gilles Johanet	3
Avant-propos de M. Ludovic Lebart	4
<b>Introduction</b>	<b>5</b>
<b>Analyses thématiques</b>	<b>11</b>
Indicateurs socio-démographiques	11
Déroulement de l'analyse	11
Principaux résultats	11
Classification des régions	12
Indicateurs d'offre de soins	15
Déroulement de l'analyse	15
Principaux résultats	15
Classification des régions	16
Indicateurs d'état de santé	18
Déroulement de l'analyse	18
Principaux résultats	18
Classification des régions	19
<b>Analyse d'ensemble</b>	<b>21</b>
Déroulement de l'analyse	21
Principaux résultats	23
Classification des régions	24
<b>Conclusion</b>	<b>28</b>
<b>Annexes</b>	
Liste des indicateurs	31
Méthodes d'analyse	39
Bibliographie	43
<b>Cartes et graphiques</b>	<b>46</b>

**Avant-propos de M. Gilles Johanet,**  
directeur général de la Caisse nationale  
d'assurance maladie des travailleurs salariés

L'analyse de la Fédération nationale des ORS est éclairante par sa méthode pragmatique et renforce la démonstration déjà connue de l'apport de l'approche globale par rapport à l'approche sectorielle. En effet, pris séparément, les indicateurs socio-démographiques, d'offre de soins ou d'état de santé, vont conduire à des similitudes entre régions ou à des regroupements de régions qui vont être différents selon les indicateurs choisis (comme l'illustrent les cartes de ce document). Cependant, pris ensemble, les indicateurs risquent de rendre le constat affligeant par le cumul des handicaps (cas du Nord-Pas-de-Calais) ou par le cumul des avantages (cas de l'Ile-de-France).

L'analyse montre une nouvelle fois qu'une offre de soins équivalente en terme de densité médicale ou de lits par habitant ne conduit pas à un état de santé équivalent tant le poids des conduites à risque (consommation d'alcool et de tabac) ou des conditions de vie dégradées (chômage, forte densité d'enfants à l'aide sociale, forte densité de personnes n'ayant que des revenus de redistribution) est prépondérant.

Néanmoins, ces résultats, déjà partiellement connus grâce, en particulier, aux tableaux de bord «La santé observée» des ORS et de la Fnors, ont déjà contribué à l'amélioration des connaissances qui ont conduit à l'élaboration de la loi de lutte contre les exclusions et à l'introduction de la couverture maladie universelle (CMU). Un accès aux soins facilité depuis le 1<sup>er</sup> janvier 2000, ainsi que la réduction constante depuis plusieurs années du nombre de chômeurs devraient permettre l'amélioration des conditions sanitaires des régions les plus défavorisées.

Cette intéressante analyse, qui a bénéficié de l'appui de la Cnamts, pourra alors être reconduite prochainement avec des indicateurs actualisés, en particulier en intégrant les résultats du recensement de 1999, et donnera une première évaluation des mesures législatives prises, conjugués à l'effet de la croissance économique retrouvées à la fin des années quatre-vingt-dix.

**Avant-propos de M. Ludovic Lebart,**  
directeur de recherche au CNRS

Les professions de santé ont été pionnières dans l'utilisation d'instruments d'observations innovants, depuis le microscope et les appareils radioscopiques et radiographiques, jusqu'à la panoplie complexe et diversifiée de l'imagerie médicale moderne. Il s'agissait et il s'agit encore d'observer des réalités cachées parce qu'infiniment petites, ou cachées sous l'opacité des tissus et organes.

Le document que l'on va consulter tente de donner des représentations d'une autre réalité non-observable directement parce que « multivariée » ou encore « multidimensionnelle », dans un domaine qui paraît pourtant assez familier : la géographie de la santé.

De nombreux indicateurs sont en effet disponibles pour décrire chaque région de France à partir de sa situation socio-démographique, de son offre de soins médicaux, et de certains aspects de son état de santé. Chaque indicateur peut donner lieu à une carte géographique.

Mais comment synthétiser des centaines de cartes géographiques que l'on ne peut que consulter séquentiellement ? Comment rapprocher les indicateurs donnant lieu à des cartes similaires, comment regrouper les régions ayant des profils semblables ?

C'est là qu'interviennent les techniques d'analyse des données, outils statistiques dont les calculs fort complexes sont rendus accessibles par des logiciels de plus en plus performants et conviviaux.

Les auteurs de ce document ont ainsi dressé des typologies des régions françaises à partir de leurs caractéristiques socio-sanitaires. Un effort tout particulier a été porté sur la représentation graphique des résultats.

Une technique originale de représentation en trois dimensions, respectant les principales lois de la perspective, permet leur synthèse, tout en évitant au lecteur certaines erreurs classiques d'interprétation.

On ne peut pas attendre d'une représentation de phénomènes complexes qu'elle soit immédiatement accessible sans le moindre effort de la part du lecteur.

Mais ici la démarche a été en quelque sorte optimisée autant qu'il est possible de le faire sur un support papier, support évidemment plus passif et figé qu'un écran d'ordinateur.

Au prix d'une certaine attention portée aux règles de lecture, une personne non formée à ce type d'analyse peut maintenant avoir une idée de toute la richesse de l'information recueillie.

Les synthèses produites, les traits structuraux mis en évidence vont sans doute stimuler la réflexion, suggérer des hypothèses, peut-être conduire à des critiques des données de base.

Que les auteurs de cet ouvrage soient donc remerciés pour avoir fait tout ce qui était techniquement possible pour valoriser ainsi l'information disponible.

# Introduction

## **Une demande de la Caisse nationale d'assurance maladie des travailleurs salariés**

En 1997, la Fnors a publié «La santé observée dans les régions de France», synthèse nationale des 26 tableaux de bord régionaux sur la santé réalisés par les observatoires régionaux de la santé.

Ce document a permis de mettre en évidence d'importantes disparités régionales en matière de santé, grâce aux comparaisons pouvant être faites entre les nombreux indicateurs présentés (environ 400).

Ces indicateurs étaient analysés, commentés et illustrés par des cartes et graphiques. Ils portaient sur 9 grands thèmes : Constat d'ensemble, Offre et consommation de soins, Mère et enfant, Personnes âgées, Personnes handicapées, Pauvreté-précarité, Pathologies, Santé-environnement.

Les partenaires nationaux, régionaux et locaux des ORS et de leur Fédération ont apprécié cette vaste présentation de comparaisons entre les régions portant sur de nombreux aspects de la santé.

Est alors apparu l'intérêt de rapprocher les unes des autres les comparaisons faites thème par thème, pour mettre en évidence les corrélations pouvant exister entre des phénomènes ou des situations démographiques, sanitaires, sociales.

La mise en œuvre d'une telle approche transversale a ainsi été demandée par la Caisse nationale d'assurance maladie des travailleurs salariés (Cnamts) à la Fnors, et financée dans le cadre du Fonds national de prévention d'éducation et d'information sanitaires (FNPEIS) sur les deux années 1998 et 1999.

## **Des travaux antérieurs, à la demande du Haut comité de la santé publique**

Une étude avait également été réalisée par la Fnors, en 1997, dans le cadre d'un groupe de travail du Haut comité de la santé publique. Il s'agissait d'étudier la faisabilité de mise en évidence de corrélations entre les indicateurs socio-économiques et une série d'indicateurs du thème «Mère et enfant», tels qu'ils existaient dans la synthèse nationale des tableaux de bord.

Cette première analyse avait permis de dégager des corrélations nettes entre certains phénomènes : c'est dans les régions jeunes, ouvrières, où les difficultés sociales sont nombreuses (personnes isolées, beaucoup de jeunes adultes non diplômés...) que l'on retrouve une mortalité infantile plus élevée et un suivi médical plus lourd pour les praticiens qui doivent prendre en charge davantage de femmes et d'enfants.

Cependant, cette approche limitée à des analyses bivariées ne permettait pas de croiser les différents thèmes, ni de dresser des typologies de régions.

## **La démarche proposée par la Fnors**

La Fnors a sollicité les ORS pour réaliser une analyse statistique afin de répondre à un double objectif :

- mettre en évidence les corrélations pouvant exister entre des situations démographiques, sanitaires et sociales ;
- dresser une typologie des régions à partir des indicateurs du tableau de bord.

Trois ORS et la Fnors se sont proposés pour répondre à cette demande.

Ce document, qui fait appel à des méthodes d'analyse multivariée, est le fruit de ce travail (la description de ces méthodes figure en annexe du document).

## Indicateurs et régions retenus pour l'analyse

### • *Indicateurs*

En 1998, les ORS et leur Fédération ont créé une base de données qui regroupe tous les indicateurs des tableaux de bord régionaux sur la santé de toutes les régions de France. C'est dans cette base qu'ont été choisis les indicateurs utilisés dans l'étude.

La base est consultable sur le site internet de la Fnors ([www.fnors.org](http://www.fnors.org)) ; elle fait partie du site commun d'observation régionale en santé «SCORE-santé», mis en place par sept partenaires (Etat, Cnamts, Fnors, Inserm, InVS, CCMSA, Canam). La Fnors, avec les ORS, assure la maintenance et l'évolution du site.

Trois critères généraux ont guidé la sélection des indicateurs :

- le choix de l'année la plus récente quand plusieurs années étaient disponibles ;
- le choix systématique de taux plutôt que d'effectifs ;
- en matière de mortalité, le choix de taux comparatifs plutôt que d'indices comparatifs de façon à avoir une variable situant les régions les unes par rapport aux autres plutôt que par rapport à la moyenne nationale.

Outre ces critères généraux, la sélection des indicateurs a tenu compte d'autres contraintes ou impératifs liés à cette étude :

- éviter la redondance des indicateurs ;
- disposer de l'ensemble des valeurs régionales pour chaque indicateur ;
- choisir des indicateurs présentant un intérêt certain pour l'étude.

Compte tenu de ces critères de sélection, 97 indicateurs ont été retenus dans la base de données des ORS pour réaliser cette étude, qui se répartissent en trois groupes :

- un premier groupe d'indicateurs socio-démographiques, tels que la proportion des jeunes de moins de 20 ans dans la population, celle des plus de 75 ans, le pourcentage d'ouvriers dans la population active, le taux de chômage... ;
- un deuxième groupe d'indicateurs concernant l'offre de soins, tels que la densité médicale, le taux d'équipement en lits hospitaliers, en structures médico-sociales... ;
- un troisième groupe d'indicateurs relatifs à l'état de santé de la population, tels l'incidence de certaines maladies et les taux comparatifs de mortalité par cause.

Certains indicateurs qui auraient été particulièrement intéressants dans cette analyse n'étaient pas disponibles (consommation d'alcool, ou données de morbidité par exemple). D'autres enfin n'ont pas été retenus car d'interprétation délicate.

Les indicateurs et leur définition figurent en annexe de ce document.

## • *Régions*

Les régions retenues sont les 22 régions de France métropolitaine.

Il n'a pas été possible dans cette étude de prendre en compte les départements d'outre-mer, qui sont à la fois régions et départements (Guadeloupe, Guyane, Martinique, Réunion), en raison de nombreuses données manquantes parmi les indicateurs utilisés. Ces régions d'outre-mer pourraient être intégrées dans d'autres analyses menées ultérieurement dans ce domaine.

## **Méthodes d'analyse**

Pour traiter ce grand nombre d'indicateurs, les méthodes d'analyse multidimensionnelle retenues sont l'analyse en composantes principales (ACP) et la classification ascendante hiérarchique. Ces méthodes descriptives ont pour but de mieux apprécier les ressemblances entre les régions et les liaisons entre les indicateurs. Elles permettent d'établir une typologie, c'est-à-dire des groupes de régions dont les profils (valeurs prises par les indicateurs) se ressemblent.

Il convient toutefois de préciser que ces ressemblances et ces proximités ne permettent aucunement de tirer des conclusions en termes de cause à effet.

Une description de ces méthodes et des références bibliographiques figurent en annexe.

## **Déroulement de l'analyse**

Dans un premier temps, des analyses en composantes principales ont été effectuées sur chacun des trois groupes d'indicateurs sélectionnés, permettant de créer des typologies de régions pour chacun de ces trois thèmes : socio-économie, offre de soins et état de santé.

Enfin, a été réalisée une analyse globale prenant en compte tous les indicateurs sélectionnés. Celle-ci permet de révéler les relations pouvant exister entre les indicateurs de ces trois groupes et de créer une typologie socio-sanitaire globale des régions françaises.

## **Présentation des résultats**

Une première partie présente les résultats des analyses menées sur chacun des trois thèmes : socio-économie, offre de soins et état de santé.

Une seconde partie présente les résultats de l'analyse menée sur l'ensemble des

indicateurs (tous thèmes confondus). Cette analyse globale est davantage développée car c'est elle qui permet de mettre en évidence les corrélations existant entre tous les indicateurs et les similarités de comportements des régions vis-à-vis de ces indicateurs.

Dans tous les cas (analyses thématiques et analyse globale), de nombreuses analyses statistiques ont été réalisées sur les indicateurs sélectionnés et ont mis en évidence la présence d'indicateurs qui apportaient peu à l'analyse ou qui présentaient de fortes redondances avec d'autres indicateurs. Selon les cas, ils ont été supprimés ou regroupés. Pour ne pas alourdir l'exposé, seules les analyses « finales » sont présentées dans ce document, les travaux réalisés « en amont » étant simplement évoqués quand ils ont abouti à la suppression d'indicateurs importants en santé publique (taux de mortalité infantile ou taux d'équipement en lits de court séjour par exemple).

Pour la même raison, les résultats sont présentés simplement, en excluant tout vocabulaire spécialisé de façon à être compris par tous les lecteurs, y compris ceux qui ne sont pas formés à l'analyse des données.

Les résultats des analyses, c'est-à-dire l'ensemble des corrélations existant entre les variables et les similarités de comportement des régions vis-à-vis de ces variables, sont synthétisés par des représentations graphiques de points représentant des variables (points-variables) ou représentant des régions (points-régions).

Les règles de lecture des représentations obtenues sont les suivantes :

- la proximité entre deux points-variables indique une forte corrélation entre ces variables, donc une similitude des cartes géographiques correspondantes, et ce d'autant plus que ces points variables sont éloignés de l'origine ;
- la proximité de deux points-régions indique des situations analogues vis-à-vis de l'ensemble des variables, la proximité est toujours d'autant plus significative que l'on s'éloigne de l'origine ;
- la représentation simultanée des variables et des régions est à interpréter avec prudence car, sur les graphes, les échelles relatives sont différentes. Elle renseigne cependant sur la répartition géographique de chaque variable en comparant sa position à l'ensemble des points régions.

De façon plus générale, elle permet de confronter les répartitions géographiques de plusieurs variables en mettant en évidence les régions responsables des similitudes ou des différences observées.

Selon les cas, les points (variables et régions) sont représentés dans le plan formé par les deux premiers axes, ou, quand le troisième axe apporte de l'information supplémentaire, dans l'espace formé par les trois premiers axes (les autres axes n'apportent pratiquement aucune information complémentaire).

Une attention toute particulière a été apportée à la réalisation de ces représentations dans l'espace afin qu'elles délivrent, au prix de quelques règles de lecture, toute la richesse de ces analyses.

### **Logiciel utilisé**

Les analyses ont été réalisées à l'aide du logiciel SPAD® (système pour l'analyse des données), conçu et diffusé par le Centre international des statistiques et d'informatique appliquées (CISIA).

# Analyses thématiques

- **Indicateurs socio-démographiques**

## Déroulement de l'analyse

Les premières analyses ont été réalisées sur les 32 indicateurs socio-démographiques concernant les 22 régions de France métropolitaine sélectionnés dans la base de données.

Très rapidement est apparue la nécessité d'isoler l'Ile-de-France de l'analyse car cette région présente pour beaucoup d'indicateurs socio-démographiques des valeurs extrêmes qui tendent à masquer les disparités pouvant exister entre les autres régions.

L'Ile-de-France a donc été traitée en tant que région supplémentaire, c'est à dire qu'elle n'intervient pas dans la construction du modèle et qu'elle est simplement placée sur le graphique (ses coordonnées étant calculées pendant l'analyse).

Au terme des analyses, 26 indicateurs permettent à l'ACP finale de décrire le paysage socio-démographique des régions de France métropolitaine (voir liste en annexe).

## Principaux résultats

L'ensemble des corrélations existant entre les variables et les similarités de comportement des régions vis-à-vis de ces variables sont synthétisées par une représentation graphique montrant les résultats projetés dans l'espace formé par les trois premiers axes : voir graphique 1 (pages en fin de document).

La lecture de ce graphique doit être prudente (voir encadré sous le graphique) et ne permet en aucun cas d'établir des relations de cause à effet entre les indicateurs.

Elle permet néanmoins de souligner l'importance de certaines variables particulièrement discriminantes :

- le pourcentage de mères de 38 ans ou plus,
- le solde migratoire et le solde naturel (avec une forte corrélation négative entre ces variables),
- le pourcentage de ménages allocataires des CAF, dont les prestations légales représentent la totalité des ressources,
- le pourcentage de salariés des entreprises privées du secteur tertiaire,
- le pourcentage des personnes de 75 ans ou plus par rapport à la population totale,
- le pourcentage de cadres dans la population active,
- la proportion de femmes actives de 20-59 ans.

Elle permet également de souligner la situation caractéristique de deux régions dans le paysage socio-démographique français (Nord-Pas-de-Calais, et bien sûr Ile-de-France traitée en tant que région «supplémentaire» dans l'analyse).

La spécificité de ces deux régions est confirmée par la classification hiérarchique établie à partir des résultats de l'analyse (voir plus loin dans ce chapitre).

## **Classification des régions**

La classification établie à partir des résultats de l'analyse en composantes principales (réalisée sur 21 régions) met en évidence une typologie des régions en cinq classes. L'Ile-de-France, en tant que région «supplémentaire» constitue une sixième classe.

### ***Nord-Pas-de-Calais***

Le Nord-Pas-de-Calais est la région la plus peuplée de métropole après l'Ile-de-France, avec une densité de population trois fois supérieure à la moyenne nationale. Sa population essentiellement urbaine (forte représentation de la banlieue et faible représentation du secteur rural) est la plus «jeune» de France : 30 % de la population a moins de 20 ans contre 26 % au plan national. L'indice conjoncturel de fécondité (nombre d'enfants par femme) y est le plus élevé de France.

Le taux de chômage est le plus élevé de France pour les hommes comme pour les femmes, la région se distingue par ailleurs par le taux d'activité des femmes le plus bas de France. On note aussi que les 25-29 ans sont peu diplômés.

### ***Aquitaine, Auvergne, Bretagne, Limousin, Midi-Pyrénées, Poitou-Charentes***

Il faut préciser que la Bretagne se situe dans ce groupe en raison de sa forte ruralité, trait commun à ces six régions de l'Ouest et du Sud-Ouest de la France. Cette caractéristique ne lui permet pas d'appartenir à une autre classe même si, pour toutes les autres caractéristiques propres à ce groupe, la Bretagne se situe plutôt dans la moyenne française.

Ces régions du Sud-Ouest de la France sont plus rurales et plus âgées que le reste du pays : la proportion de jeunes de moins de 20 ans est la plus faible de France et celle des 75 ans ou plus est la plus élevée. La proportion de population vivant en zone rurale ainsi que la proportion d'agriculteurs exploitants sont parmi les plus fortes de métropole.

Elles ont une fécondité particulièrement basse : l'indicateur conjoncturel de fécondité est compris entre 1,43 et 1,63 contre 1,80 au plan national. D'ailleurs, l'accroissement de la population (des villes et des banlieues) dû au solde naturel est modéré.

Le pourcentage de ménages dont les prestations légales représentent la totalité des ressources est supérieur à la moyenne nationale. La densité de population est faible.

### ***Corse, Languedoc-Roussillon, Provence-Alpes-Côte-d'Azur***

Ces régions du Sud-Est se caractérisent par un secteur tertiaire développé (au détriment du secteur secondaire). Le chômage y est relativement important et la proportion de femmes actives relativement faible. C'est dans ces régions que les pourcentages de ménages dont les allocations représentent la totalité des ressources sont les plus élevés : entre 18,7 et 25,2 % contre 14,0 % au plan national.

Ces régions ont également en commun une structure de population âgée : peu de jeunes de moins de 20 ans et beaucoup de personnes âgées de 75 ans ou plus. De plus, la proportion de mères 38 ans ou plus y est très forte (juste derrière l'Ile-de-France).

D'autre part, elles se caractérisent par des soldes migratoires de population (dans les villes et banlieues) plus élevés qu'ailleurs.

### ***Alsace, Rhône-Alpes***

Ces régions de l'Est se rapprochent de la moyenne nationale pour la plupart des indicateurs étudiés. On peut souligner cependant une forte densité de population et une forte représentation des cadres. En revanche, il y a moins de personnes âgées que dans le reste du pays et la proportion d'agriculteurs-exploitants est plutôt faible.

Dans une moindre mesure, on observe également une faible représentation du secteur

rural, peu de personnes vivant exclusivement d'allocations de la CAF et un faible taux de jeunes de 25-29 ans non diplômés.

***Bourgogne, Centre, Basse-Normandie, Champagne-Ardenne, Franche-Comté, Haute-Normandie, Lorraine, Pays de la Loire, Picardie***

Ces régions ont en commun leur situation économique : les salariés des entreprises travaillent plus fréquemment dans le secteur secondaire et moins fréquemment dans le secteur tertiaire. Ceci se traduit par une forte représentation des ouvriers et une sous-représentation des cadres dans ces régions.

Ce sont les régions de France où l'on trouve le moins de ménages dont les prestations légales versées par les caisses d'allocations familiales représentent la totalité des ressources.

Les régions de ce groupe se distinguent par une proportion de moins de 20 ans dans la population, un indice conjoncturel de fécondité et par conséquent un taux d'accroissement naturel de leur population plus élevés que dans les autres régions.

***Ile-de-France***

L'Ile-de-France présente des caractéristiques socio-démographiques très marquées qui expliquent qu'elle ait été placée en région supplémentaire dans le cadre de cette analyse. Cette région se caractérise d'abord par sa forte activité économique basée principalement sur un secteur tertiaire très développé. La proportion de cadres dans la population active atteint 19,0 % contre 10,5 % en France alors que les ouvriers sont sous-représentés. La proportion des femmes actives est la plus élevée de France et le taux de chômage des femmes est très faible.

C'est également en Ile-de-France que le pourcentage de ménages allocataires des CAF dont les prestations légales représentent la totalité des ressources est le plus faible de France (9,6 % contre 14,0 % au niveau national).

L'Ile-de-France se caractérise également par sa forte urbanisation qui se traduit par une densité de population huit fois plus élevée qu'en moyenne en France.

La population francilienne se caractérise par une faible proportion de personnes âgées : 5,0 % de sa population est âgée de 75 ans ou plus contre 6,4 % au plan national.

Le taux d'accroissement dû au mouvement naturel de la population est le plus élevé de France (mais le déficit migratoire est important) et la fécondité est particulièrement importante chez les femmes âgées de 38 ans ou plus.

## • Indicateurs d'offre de soins

### Déroulement de l'analyse

Les premières analyses ont été réalisées sur les 33 indicateurs d'offre de soins concernant les 22 régions de France métropolitaine sélectionnés dans la base de données. Au terme des analyses, 17 indicateurs permettent à l'ACP finale de décrire le paysage sanitaire des régions de France métropolitaine (voir liste en annexe).

Cette importante suppression de variables entre le début de l'analyse et l'ACP finale s'explique par le fait que beaucoup d'entre elles ont pu être regroupées : les densités en médecins pour chaque spécialité, en particulier, ont pu être synthétisées par la densité toutes spécialités confondues.

De plus, des variables apportaient très peu à l'analyse et ont pu être retirées. C'est le cas notamment des indicateurs de médecine scolaire (infirmiers et médecins), des indicateurs de mode de prise en charge (psychiatrie, conventionnement des médecins libéraux) et des densités de professionnels de santé salariés.

### Principaux résultats

L'ensemble des corrélations existant entre les variables et les similarités de comportement des régions vis-à-vis de ces variables sont synthétisés par la représentation graphique proposée ci-contre (résultats projetés dans le plan formé par les deux premiers axes).

La lecture de ce graphique doit être prudente (voir l'encadré lecture du graphique) et ne permet en aucun cas d'établir des relations de cause à effet entre les indicateurs.

Elle permet néanmoins de souligner l'importance de certains indicateurs :

- les densités en médecins libéraux, tant généralistes que spécialistes, ainsi que les densités en chirurgiens-dentistes, kinésithérapeutes et infirmiers libéraux.  
Ce sont ces indicateurs, en effet, qui départagent le plus les régions françaises en matière d'offre de soins,
- dans une moindre mesure, apparaissent les indicateurs d'équipement hospitalier tels que le taux d'équipement en lits de soins de courte durée et de soins de suite.

Tous ces indicateurs sont corrélés positivement entre eux, et sont corrélés négativement avec un autre groupe d'indicateurs de prise en charge des personnes âgées et des personnes handicapées : taux d'équipement en hébergement collectif (médicalisé ou non) pour personnes âgées, pourcentage de personnes âgées de 65 ans ou plus bénéficiant de soins infirmiers et d'aide ménagère à domicile, proportion des adultes handicapés de 20-59 ans travaillant en structure de travail protégé.

- Les variables d'offre de soins en psychiatrie hospitalière sont également déterminantes dans cette analyse.

## **Classification des régions**

La classification établie à partir des résultats de l'analyse en composantes principales met en évidence une typologie des régions en cinq classes.

### ***Corse, Ile-de-France, Provence-Alpes-Côte-d'Azur***

Ces régions sont fortement dotées en médecins généralistes et spécialistes, en kinésithérapeutes et dentistes libéraux. De même, les taux d'équipement en lits de court séjour y sont très élevés.

En revanche, les taux d'équipement en hébergement collectif pour personnes âgées (médicalisé ou non) et la proportion de personnes âgées bénéficiant de soins infirmiers à domicile sont les plus faibles de France.

De même, les taux de personnes en structures spécialisées pour adultes handicapés sont les plus faibles de métropole.

### ***Aquitaine, Languedoc-Roussillon, Midi-Pyrénées***

Les régions de ce groupe ont en commun des taux d'équipement en lits d'hospitalisation pour les soins de suite, des densités de médecins généralistes, médecins spécialistes, dentistes, infirmiers et kinésithérapeutes libéraux très supérieurs à la moyenne nationale. La psychiatrie hospitalière adulte est fortement développée. En revanche, les taux d'équipement en hébergement collectif pour personnes âgées sont relativement faibles.

***Auvergne, Bretagne, Limousin***

Ces régions se caractérisent par des taux d'équipement en lits et places d'hospitalisation psychiatrique (enfants et adultes) et des taux d'hébergement collectif pour personnes âgées très supérieurs à la moyenne nationale. L'offre en structures de travail protégé (centres d'aide par le travail et ateliers protégés) est également importante.

***Alsace, Centre, Champagne-Ardenne, Lorraine***

Les régions de ce groupe ont en commun des densités en médecins généralistes, infirmiers et kinésithérapeutes libéraux très inférieures à la moyenne nationale. L'offre d'hébergement pour personnes âgées y est importante, mais les bénéficiaires d'aide ménagère à domicile peu nombreux.

***Bourgogne, Franche-Comté, Nord-Pas-de-Calais, Basse-Normandie, Haute-Normandie, Pays de la Loire, Picardie, Poitou-Charentes, Rhône-Alpes***

Les régions de ce dernier groupe présentent des indicateurs d'offre de soins globalement assez proches de la moyenne nationale.

Il faut souligner toutefois que l'offre en structures d'hébergement pour personnes âgées et l'offre en structures de travail protégé (pour personnes handicapées) y est relativement importante. A l'inverse, les densités hospitalières en médecins et para-médicaux libéraux sont en deçà de la moyenne nationale.

## • Indicateurs d'état de santé

### Déroulement de l'analyse

Les premières analyses ont été réalisées sur les 32 indicateurs d'état de santé concernant les 22 régions de France métropolitaine sélectionnés dans la base de données.

Au terme des analyses, 17 indicateurs permettent à l'ACP finale de décrire le paysage sanitaire des régions de France métropolitaine (voir liste en annexe).

Les variables supprimées pendant les analyses sont les indicateurs de mortalité des jeunes et de maladies professionnelles (peu discriminantes), les indicateurs de mortalité prématurée (hommes et femmes) fortement corrélés aux indicateurs de mortalité par cause qui leur ont été préférés car plus riches d'information.

Les taux comparatifs de mortalité par cardiopathies ischémiques par sexe ont été abandonnés au profit du taux comparatif deux sexes confondus.

### Les principaux résultats

L'ensemble des corrélations existant entre les variables et les similarités de comportement des régions vis-à-vis de ces variables sont synthétisées par la représentation graphique proposée ci-contre (résultats projetés dans le plan formé par les deux premiers axes).

La lecture de ce graphique doit être prudente (voir l'encadré lecture du graphique) et ne permet en aucun cas d'établir des relations de cause à effet entre les indicateurs.

Elle permet néanmoins de souligner l'importance de la mortalité liée (pour partie) à la consommation d'alcool et/ou de tabac : taux comparatifs de mortalité par cirrhose (hommes et femmes), par psychose alcoolique (hommes), par cancer du poumon (hommes et femmes), par cancer des voies aéro-digestives supérieures (hommes) et par cardiopathies ischémiques (deux sexes).

La mortalité par cancer de la prostate et par suicide, ainsi que la mortalité par accidents vasculaires-cérébraux (femmes) se révèlent également discriminants dans cette analyse.

Elle permet également de souligner la situation caractéristique de trois régions dans le paysage sanitaire français (Corse, Ile-de-France et Nord-Pas-de-Calais).

La spécificité de ces trois régions est confirmée par la classification hiérarchique établie à partir des résultats de l'analyse.

## **Classification des régions**

La classification établie à partir des résultats de l'analyse en composantes principales met en évidence une typologie des régions en sept classes.

### ***Nord-Pas-de-Calais***

Cette région présente à la fois une forte mortalité liée à l'alcoolisme et une forte mortalité liée au tabagisme. En effet, les taux comparatifs de mortalité par cirrhose (hommes et femmes), par cancer des voies aéro-digestives supérieures (hommes), par cardiopathies ischémiques (deux sexes), par cancer du poumon (hommes) et par bronchite chronique (hommes) sont les plus élevés de France. De même, les taux comparatifs de mortalité par psychose alcoolique (hommes) et par bronchite chronique chez les femmes figurent parmi les plus élevés de France.

Il faut également souligner une forte mortalité par cancer de la prostate et par maladies cérébro-vasculaires.

En revanche, l'incidence du sida est, après la Picardie, la plus faible de France.

### ***Corse***

Cette région se caractérise par une forte mortalité par maladies cérébro-vasculaires chez les femmes (la plus forte de France) et par une forte mortalité par cancer du poumon (hommes et femmes). En revanche, la Corse présente les plus faibles niveaux de mortalité pour les pathologies liées à l'alcool (cirrhose, psychose alcoolique et cancer des voies aéro-digestives supérieures).

La Corse se distingue également par la plus faible mortalité par cancer de la prostate et par une faible mortalité par suicide et par bronchite chronique (femmes).

### ***Ile-de-France***

La région Ile-de-France se caractérise par une faible mortalité cardio-vasculaire (maladies cérébro-vasculaires notamment), par une faible mortalité par bronchite

chronique et par une faible mortalité par suicide.

En revanche, la mortalité par cancer du poumon chez les femmes y est très élevée et la région enregistre la plus forte incidence du sida.

### ***Alsace, Lorraine***

Ces deux régions se caractérisent par une importante mortalité liée au tabagisme. En effet, les taux comparatifs de mortalité par cancer du poumon, par bronchite chronique (hommes) et par cardiopathies ischémiques y sont parmi les plus élevés de France. La mortalité par maladies cérébro-vasculaires y est également élevée.

### ***Bretagne, Champagne-Ardenne, Basse-Normandie, Haute-Normandie, Picardie***

Ces régions présentent comme le Nord-Pas-de-Calais une forte mortalité liée à l'alcool (cirrhose, psychose alcoolique, cardiopathies ischémiques, cancer des voies aéro-digestives supérieures), une forte mortalité par cancer de la prostate et une faible incidence du sida. Elles s'en différencient par une mortalité liée au tabac plus faible et une mortalité par suicide plus élevée.

Dans ce groupe, une région se distingue : la Bretagne, qui présente les taux comparatifs de mortalité par alcoolisme (hommes), par suicide, par bronchite chronique (femmes) et par cancer de la prostate les plus élevés du groupe.

### ***Aquitaine, Languedoc-Roussillon, Midi-Pyrénées, Provence-Alpes-Côte d'Azur, Rhône-Alpes***

Ces régions du Sud de la France se caractérisent par des taux comparatifs de mortalité inférieurs à la moyenne nationale. En particulier, la mortalité par cirrhose, par psychoses alcooliques, par cardiopathies ischémiques, par cancer des voies aéro-digestives supérieures, par cancer de la prostate et par suicide est faible.

Mais ces régions présentent, après l'Ile-de-France, les taux d'incidence du sida les plus élevés.

### ***Auvergne, Bourgogne, Centre, Franche-Comté, Limousin, Poitou-Charentes, Pays de la Loire***

Ces régions présentent une mortalité liée à l'alcool et au tabac plus faible que dans le Nord du pays, mais une mortalité par suicide supérieure à la moyenne nationale.

# Analyse d'ensemble

## Déroulement de l'analyse

Les premières analyses ont été réalisées sur l'ensemble des indicateurs sélectionnés : 32 indicateurs socio-démographiques, 33 indicateurs d'offre de soins et 32 indicateurs d'état de santé, soit au total 97 indicateurs socio-sanitaires concernant les 22 régions de France métropolitaine.

Très vite est apparu le poids important des indicateurs socio-démographiques, plus discriminants que les indicateurs d'état de santé et d'offre de soins, en raison notamment des valeurs extrêmes prises par ces indicateurs en Ile-de-France (voir l'analyse portant sur les variables socio-démographiques dans la première partie du document).

Deux solutions permettaient alors de poursuivre l'analyse sans que les indicateurs socio-démographiques ne masquent les disparités régionales d'offre de soins et d'état de santé :

- traiter l'Ile-de-France en tant que région supplémentaire ; dans ce cas, les données concernant cette région n'intervenaient pas dans l'analyse et la région était simplement positionnée sur le graphique,
- traiter les variables socio-démographiques en tant que variables supplémentaires : dans ce cas, les variables actives étaient celles concernant l'offre de soins et l'état de santé (variables dites actives car ce sont elles qui participent à la construction du modèle), les variables socio-démographiques étant simplement projetées dans ce modèle.

Les deux solutions ont été considérées et la première a été écartée, car, après le retrait de l'Ile-de-France, d'autres régions présentaient également des situations atypiques pour les indicateurs socio-démographiques (Corse, Limousin, Nord-Pas-de-Calais) ; il aurait également fallu les placer en tant que régions supplémentaires pour ne pas perturber le modèle.

C'est donc la seconde solution qui a été retenue, permettant de mettre en évidence les disparités régionales d'offre de soins et de santé, les variables socio-démographiques apportant un éclairage supplémentaire.

Au terme des analyses, 9 variables d'offre de soins et 18 variables d'état de santé permettent à l'ACP finale de décrire le paysage sanitaire des régions de France métropolitaine.

Dans ce paysage, ont pu être projetés les 32 indicateurs socio-démographiques, mais aussi les indicateurs d'offre de soins et de santé ne participant pas à la construction du modèle (leurs coordonnées étant calculées pendant l'analyse).

L'importante suppression de variables au cours de l'analyse s'explique par la redondance ou le peu d'apport à l'analyse de certaines variables, qui ont pu être supprimées, ou par certains regroupements de variables qui ont pu être opérés (des exemples sont indiqués dans les analyses thématiques).

De plus, cette diminution du nombre de variables était souhaitable d'un point de vue purement technique. En effet, l'analyse en composantes principales peut être perturbée si elle est réalisée sur un nombre de variables explicatives très supérieur au nombre d'individus (ici 22 régions).

Les **9 indicateurs d'offre de soins** retenus sont les densités de médecins libéraux généralistes et spécialistes, les densités d'infirmiers libéraux, de masseurs-kinésithérapeutes libéraux et de chirurgiens-dentistes libéraux, le nombre de lits d'hospitalisation complète en soins de suite, l'offre d'hébergement pour personnes âgées (toutes structures confondues), la proportion d'enfants bénéficiaires de l'allocation éducation spéciale (AES) et l'offre de prise en charge des toxicomanes.

Les **18 indicateurs d'état de santé** recouvrent surtout la mortalité liée à la consommation d'alcool et/ou de tabac : taux comparatifs de mortalité par cirrhose, par psychose alcoolique (hommes), par cancer des voies aéro-digestives supérieures (hommes), par cancer du poumon (hommes) et par bronchite chronique.

Les autres indicateurs d'état de santé concernent la mortalité cardio-vasculaire (maladies cérébro-vasculaires et cardiopathies ischémiques), la mortalité par cancer de la prostate, par cancer du côlon-rectum, par cancer de l'utérus, par suicide, l'incidence de la tuberculose et du sida et le taux de mortalité post-néonatale.

## Principaux résultats

L'ensemble des corrélations existant entre les variables et les similarités de comportement des régions vis-à-vis de ces variables sont synthétisées par la représentation graphique montrant les résultats projetés dans l'espace formé par les trois axes : voir graphique 2 (pages en fin de document).

La lecture de ce graphique doit être prudente (voir encadré lecture du graphique) et ne permet en aucun cas d'établir des relations de cause à effet entre les indicateurs.

Elle permet néanmoins de souligner certaines **associations fortes** :

- mortalité élevée par alcoolisme et par cancer de la prostate dans des régions ouvrières à faible densité médicale et para-médicale,
- forte mortalité par suicide, forte mortalité post-néonatale et taux d'équipement élevé en structures médico-sociales (personnes âgées et handicapés) dans les régions à forte population péri-urbaine,
- des mères plus âgées, une forte incidence du sida et de la tuberculose et de fortes densités en spécialistes dans les régions à forte proportion de cadres.

Elle permet également de souligner la situation caractéristique de trois régions (Corse, Ile-de-France, Nord-Pas-de-Calais) dans le paysage socio-sanitaire français.

La spécificité de ces trois régions est confirmée par la classification hiérarchique qui propose à partir de l'ACP, une classification en sept classes :

- Corse,
- Ile-de-France,
- Nord-Pas-de-Calais,
- Languedoc-Roussillon, Midi-Pyrénées, Provence-Alpes-Côte-d'Azur
- Alsace, Lorraine
- Bretagne, Champagne-Ardenne, Basse-Normandie, Haute-Normandie, Picardie,
- Aquitaine, Auvergne, Bourgogne, Centre, Franche-Comté, Limousin, Pays de la Loire, Poitou-Charentes, Rhône-Alpes.

## **Classification des régions**

La typologie établie à partir des résultats de l'analyse en composantes principales permet de mettre en évidence sept groupes de régions.

Trois régions se singularisent : Ile-de-France, Corse et Nord-Pas-de-Calais.

### ***Ile-de-France***

L'Ile-de-France se distingue par la position extrême qu'elle occupe pour de nombreux indicateurs.

C'est particulièrement vrai pour les indicateurs socio-démographiques : le PIB par habitant, la proportion de cadres dans la population active et la proportion de femmes de 20-59 ans actives (parmi l'ensemble des femmes de cette tranche d'âge) y sont les plus élevés de France.

A l'inverse, c'est dans cette région que la population de 60 ans ou plus et le pourcentage de personnes vivant uniquement grâce aux prestations légales sont les plus faibles.

Sa forte urbanisation se traduit par une densité de population élevée et par une importante proportion de population vivant en banlieue. La proportion d'étrangers, le taux d'accroissement naturel de la population et la proportion de mères âgées de 38 ans ou plus y sont également les plus élevés de France.

Ces spécificités ne sont pas sans conséquence sur la situation sanitaire et de fait, l'Ile-de-France se distingue par des points forts : très faible mortalité par maladies cérébro-vasculaires et par bronchite chronique chez les hommes et faible mortalité par suicide, mais aussi par des points faibles avec en particulier les plus fortes incidences de la tuberculose et du sida de France métropolitaine.

L'offre de soins dans cette région se caractérise par de faibles densités en lits de moyen séjour, en infirmiers libéraux et en structures médico-sociales. En revanche, la région est fortement dotée en médecins spécialistes libéraux.

### ***Corse***

Cette région se caractérise par une faible mortalité liée à l'alcool (cirrhose hommes et femmes, psychose alcoolique et cancer des voies aéro-digestives supérieures chez les hommes).

La mortalité par suicide et par cancer de la prostate y est également très faible.

En revanche, la mortalité liée au tabagisme y est relativement importante (maladies

cérébro-vasculaires chez les hommes et chez les femmes, cancer du poumon chez les hommes).

C'est dans cette région que l'on enregistre également le plus faible niveau de mortalité post-néonatale, mais il faut préciser que le taux est calculé sur de faibles effectifs.

Les densités en infirmiers et kinésithérapeutes libéraux y sont relativement élevées, mais les structures d'hébergement pour personnes âgées peu nombreuses.

La proportion d'enfants bénéficiaires de l'allocation d'éducation spéciale (AES) parmi les moins de 20 ans est élevée. La proportion de toxicomanes pris en charge dans la population est également élevée.

Cette région peu peuplée (une grande partie de la population est concentrée dans les deux principales villes) se caractérise par une situation économique peu favorisée avec en particulier le plus faible produit intérieur brut (PIB) par habitant de France, un très faible taux d'activité féminine et la plus forte proportion de ménages vivant uniquement d'allocations. Le secteur d'activité dominant est le secteur tertiaire (administration, tourisme) alors que le secteur secondaire (industrie) est très peu développé.

La proportion de moins de 20 ans est faible et les mères âgées de 38 ans ou plus sont relativement nombreuses.

### *Nord-Pas-de-Calais*

La région Nord-Pas-de-Calais se caractérise par une situation sanitaire défavorable avec notamment une forte mortalité liée à l'alcoolisme et/ou au tabagisme (cirrhose, psychose alcoolique, cancer des voies aéro-digestives supérieures, cancer du poumon chez les hommes, bronchite chronique, cardiopathies ischémiques et maladies cérébro-vasculaires).

La mortalité par cancer de la prostate y est également élevée.

En revanche, l'incidence du sida est particulièrement faible dans cette région.

Le nombre de toxicomanes pris en charge pour 100 000 habitants est le plus élevé de France ainsi que le nombre d'enfants bénéficiaires de l'allocation d'éducation spéciale (AES).

Les densités en dentistes libéraux et en lits de moyen séjour y sont particulièrement faibles.

Après l'Ile-de-France, la région Nord-Pas-de-Calais est la région la plus urbanisée de France ce qui se traduit par une forte proportion de personnes vivant en banlieue et une forte densité de population. La structure par âge de la population est particulièrement jeune et c'est dans cette région que la part des naissances issues de mères âgées de moins de 18 ans est la plus élevée.

### ***Languedoc-Roussillon, Midi-Pyrénées, Provence-Alpes-Côte d'Azur***

Ces régions du Sud de la France se caractérisent par une offre de soins importante (fortes densités en médecins et paramédicaux libéraux, forte densité en lits de moyen séjour). A noter, toutefois, le peu de structures d'hébergement pour personnes âgées eu égard à l'importante population âgée dans ces régions.

La situation sanitaire y est globalement favorable avec en particulier une faible mortalité liée à la consommation d'alcool (cirrhose alcoolique, psychose alcoolique, cancers des voies aéro-digestives supérieures).

La mortalité par cardiopathies ischémiques, par suicide, par bronchite chronique, par cancer de la prostate et par cancer du côlon chez la femme y est également faible. L'incidence du sida, en revanche, y est assez importante.

Ces régions relativement dynamiques sur le plan économique se caractérisent par un secteur tertiaire important et une forte proportion de cadres dans leur population. Elles exercent une attractivité qui contribue à leur forte croissance démographique. Ce contexte favorable doit cependant être tempéré, notamment par la forte proportion de personnes dont les seules ressources financières sont les allocations qu'elles perçoivent.

### ***Alsace, Lorraine***

Ces régions de l'Est de la France présentent une mortalité cardio-vasculaire particulièrement élevée (maladies cérébro-vasculaires et cardiopathies ischémiques) ainsi qu'une surmortalité par cancer du poumon chez les hommes.

L'incidence du sida y est en revanche peu élevée.

Les densités en médecins et paramédicaux libéraux y sont relativement faibles, particulièrement en Lorraine.

### ***Bretagne, Champagne-Ardenne, Basse-Normandie, Haute-Normandie, Picardie***

Ces régions se caractérisent par une forte mortalité liée à l'alcool (cirrhose alcoolique, psychose alcoolique, cancer des voies aéro-digestives supérieures), ainsi qu'une forte mortalité par suicide et par cancer de la prostate.

L'incidence du sida est en revanche peu élevée dans ces régions.

L'offre de soins y est globalement peu importante mais les structures d'hébergement pour personnes âgées y sont relativement nombreuses.

Au niveau socio-démographique, ces régions se distinguent peu de la moyenne nationale sinon par des proportions de jeunes et d'ouvriers élevées.

*Aquitaine, Auvergne, Bourgogne, Centre, Franche-Comté, Limousin, Pays de la Loire, Poitou-Charentes, Rhône-Alpes*

Ce groupe de neuf régions se distingue peu de la moyenne nationale.

Il faut toutefois souligner la situation de l'Aquitaine qui, par certaines caractéristiques (offre de soins, incidence du sida), se rapproche des régions du Sud, ainsi que la situation du Limousin caractérisée par une population très agricole et assez âgée.

# Conclusion

Pour répondre à la demande formulée par la Caisse nationale d'assurance maladie, le groupe de travail a cherché à mettre en évidence les principales corrélations pouvant exister entre les situations démographiques, sanitaires et sociales des 22 régions de France métropolitaine et à dresser des typologies de ces régions à partir des indicateurs des tableaux de bord régionaux sur la santé.

Différentes analyses en composantes principales, réalisées sur une partie ou sur tous les indicateurs, ont permis de mettre en évidence trois types associations fortes :

- mortalité élevée par alcoolisme, mortalité élevée par cancer de la prostate, régions ouvrières, faibles densités médicales et paramédicales,
- forte mortalité post-néonatale, forte mortalité par suicide, taux d'équipement élevé en structures médico-sociales (personnes âgées et handicapées), régions à forte population péri-urbaine,
- mères plus âgées que la moyenne, forte incidence du sida, forte incidence de la tuberculose, densités élevées en spécialistes libéraux, régions où les cadres sont nombreux.

Une typologie en sept groupes de régions synthétisant l'ensemble de l'information a été établie sur une sélection de 97 indicateurs socio-démographiques, d'offre de soins et d'état de santé.

Dans cette typologie, trois régions se singularisent : le Nord-Pas-de-Calais, l'Ile-de-France et la Corse. Les autres régions se répartissent en quatre groupes : l'Alsace-Lorraine, un arc Nord-Nord Ouest (Champagne-Ardenne, Picardie, Haute-Normandie, Basse-Normandie, Bretagne), un arc Sud-Sud Est (Midi-Pyrénées, Languedoc-Roussillon, Provence-Alpes-Côte d'Azur) et un axe médian regroupant les autres régions.

- Le groupe Alsace et Lorraine se caractérise par une forte mortalité par maladies vasculaires cérébrales et par cancer du poumon. L'incidence du sida est particulièrement faible. L'offre de soins tant médicale que paramédicale est peu élevée..
- Dans les régions jeunes et ouvrières de l'arc Nord / Nord-Ouest, l'offre de soins est peu importante et l'incidence du sida est faible. Mais ce groupe se distingue du précédent par une forte mortalité par alcoolisme et par suicide.
- En Nord-Pas-de-Calais, la mortalité présente à la fois les particularités négatives de l'Alsace (liées au tabagisme) et des régions de l'arc Nord / Nord-Ouest (liées à l'alcoolisme). Ces deux caractéristiques font de cette région jeune et urbaine une région très défavorisée du point de vue sanitaire. Cependant, comme dans le groupe Alsace-Lorraine et dans l'arc Nord / Nord-Ouest, l'incidence du sida est peu élevée.
- L'Ile-de-France se singularise à bien des égards, avec des indicateurs prenant souvent des valeurs extrêmes. Elle présente une forte incidence du sida et de la tuberculose, et une faible mortalité par maladies vasculaires cérébrales et par suicide. L'offre médicale, notamment en spécialistes est importante dans cette région économiquement favorisée.
- Les régions du Sud–Sud Est présentent une forte incidence du sida et une faible mortalité par alcoolisme. L'offre de soins est également importante dans ces régions ayant à la fois des proportions élevées de cadres et de ménages vivant uniquement d'allocations.
- La Corse, comme les régions du Sud-Est, se caractérise par une faible mortalité alcoolique. On y observe en revanche une forte mortalité par tabagisme. Les équipements pour personnes âgées sont peu nombreux, mais l'offre de soins paramédicale est plutôt bonne dans cette région économiquement défavorisée.
- L'axe médian (régions au Sud d'une ligne Pays de la Loire – Franche-Comté et au Nord d'un axe Aquitaine – Rhône-Alpes) regroupe toutes les régions françaises pour lesquelles les indicateurs discriminant les autres classes se situent dans la moyenne nationale.

Rappelons encore une fois que la méthode statistique appliquée ici ne permet en aucun cas d'établir des liens de cause à effet entre les variables qui participent à l'analyse.

Toutefois, les découpages du territoire qui résultent de cette analyse, notamment certaines oppositions nord/sud, ne sont pas inconnus des professionnels de la santé publique.

De même, les variables qui se révèlent les plus discriminantes, celles ayant trait à des conditions de vie dégradées (chômage, proportion de personnes n'ayant que des revenus de redistribution) et à des conduites à risque (alcoolisme et, dans une moindre mesure, tabagisme, suicide et sida) sont régulièrement dénoncées comme génératrices d'inégalité de santé en France.

Le niveau géographique retenu ici, la région, bien que fortement agrégé et recouvrant des disparités internes, confirme être un échelon pertinent d'observation de la santé. Toutefois, une analyse comparable menée à un niveau géographique plus fin, le département ou le canton, permettrait certainement de préciser certains résultats.

Il conviendrait enfin de reproduire régulièrement ces analyses avec les données les plus récentes possibles pour disposer des évolutions en cours. Depuis le milieu des années 1990, par exemple, la mortalité par sida, la répartition du chômage, les niveaux de mortalité infantile et périnatale ont considérablement évolué dans les régions françaises. De même, le recensement général de la population réalisé en 1999 et dont les résultats sont progressivement diffusés par l'Insee révèle quelques modifications du paysage socio-démographique de 1990.

# Annexes

## 1 - Liste des indicateurs

Dans les pages suivantes sont présentés les 97 indicateurs retenus dans le cadre de cette analyse.

Ils sont classés en trois groupes :

- indicateurs socio-démographiques (32 indicateurs),
- indicateurs d'offre de soins (32 indicateurs),
- indicateurs d'état de santé (33 indicateurs).

Pour chacun d'eux sont indiqués une définition, la date de validité de l'indicateur et sa source, sa valeur en France métropolitaine et ses valeurs minimum et maximum dans les régions de France métropolitaine.

Les croix figurant dans les dernières colonnes indiquent dans quelles analyses (socio-démographique, offre de soins, état de santé, globale) l'indicateur a été retenu pour l'analyse en composantes principales présentée dans ce document.

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Indicateurs socio-démographiques									
Ville-centre	Pourcentage de population vivant dans les villes-centres (définition Insee).	1990 Insee	41,8	36	62,3	x			x
Banlieue	Pourcentage de population vivant en banlieue (définition Insee).	1990 Insee	32,2	5,8	49,5	x			x
Rural périurbain	Pourcentage de population vivant dans le rural périurbain (définition Insee).	1990 Insee	22,3	9,7	40,3				x
Rural profond	Pourcentage de population vivant dans le rural profond (définition Insee).	1990 Insee	3,7	0,1	21,8	x			x
Taux de variation dû au mouvement naturel	Part de variation de la population des villes et des banlieues en 1982-90 dû au solde naturel. Le solde naturel est la différence entre le nombre de naissances et le nombre de décès.	1982-90 Insee	0,6	0,1	0,7	x			x
Taux de variation dû au mouvement migratoire	Part de variation de la population des villes et des banlieues en 1982-90 dû au solde migratoire. Le solde migratoire est la différence entre l'effectif des immigrants et celui des émigrants.	1982-90 Insee	-0,1	-0,9	0,9	x			x
Produit intérieur brut par habitant	Produit intérieur brut (PIB) par habitant, en francs constants (valeur 1980). Le PIB mesure la richesse produite par l'ensemble de l'économie.	1991 Insee	63042	43651	63301	x			x
Effectifs des salariés du secteur primaire	Pourcentage des salariés des entreprises privées du secteur primaire.	1989 Insee	0,9	0,4	2,3	x			x
Effectifs des salariés du secteur secondaire	Pourcentage des salariés des entreprises privées du secteur secondaire.	1989 Insee	30	6,7	47,2	x			x
Effectifs des salariés du secteur BGCA	Pourcentage des salariés des entreprises privées du bâtiment, génie civil et agricole (BGCA).	1989 Insee	8,3	6,5	10,8				x
Effectifs des salariés du secteur tertiaire	Pourcentage des salariés des entreprises privées du secteur tertiaire.	1989 Insee	60,8	45,5	83,9	x			x
Densité de population	Nombre d'habitants par km2.	1990 Insee	102	29	319	x			x
Population < 20 ans	Pourcentage des moins de 20 ans dans la population totale.	1996 Insee	26	20,6	29,9	x			x
Population de 60-74 ans	Pourcentage des 60-74 ans dans la population totale.	1996 Insee	13,8	12,5	18,8				x
Population ≥ 75 ans	Pourcentage des 75 ans et plus dans la population totale.	1996 Insee	6,4	5	9,9	x			x
Ensemble des étrangers	Pourcentage de la population étrangère (CEE et hors CEE) dans la population totale.	1990 Insee	6,3	0,9	9,9	x			x
Agriculteurs	Pourcentage d'agriculteurs-exploitants dans la population active. La population active comprend les actifs ayant un emploi, les chômeurs et les appelés du contingent.	1990 Insee	4	1,9	9,8	x			x
Ouvriers	Pourcentage d'ouvriers dans la population active.	1990 Insee	30,1	25,8	38,2	x			x
Cadres	Pourcentage de cadres dans la population active.	1990 Insee	10,7	6,2	19	x			x

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

25-29 ans non diplômés	Pourcentage des 25-29 ans n'ayant pas de diplôme dans la population de cette tranche d'âge.	1990 Insee	18,9	13,3	29,6	x			x
Activité des femmes de 20-59 ans	Pourcentage de femmes actives de 20-59 ans dans la population de cette classe d'âge.	1990 Insee	69,2	51,3	73,7	x			x
Chômage des hommes	Taux de chômage des hommes (nombre de chômeurs rapportés à la population active). La définition du chômage est celle du Bureau international du travail (BIT).	1994 Insee	10,6	6,6	14,4	x			x
Chômage des femmes	Taux de chômage des femmes (nombre de chômeurs rapportés à la population active). La définition du chômage est celle du Bureau international du travail (BIT).	1994 Insee	14	9,3	19,2	x			x
Chômage de longue durée des hommes	Pourcentage d'hommes âgés de 25 à 49 ans au chômage depuis plus d'un an dans la population totale de la tranche d'âge. La définition du chômage est celle de l'Agence nationale pour l'emploi (ANPE) = demandes d'emploi en fin de mois.	1990 ANPE, Insee	2,5	1,2	4,9				x
Chômage de longue durée des femmes	Pourcentage de femmes âgées de 25 à 49 ans au chômage depuis plus d'un an dans la population totale de la tranche d'âge. La définition du chômage est celle de l'Agence nationale pour l'emploi (ANPE) = demandes d'emploi en fin de mois.	1990 ANPE, Insee	4,6	2,8	6,6				x
Chômage de longue durée	Pourcentage de personnes âgées de 25 à 49 ans au chômage depuis plus d'un an dans la population totale de la tranche d'âge. La définition du chômage est celle de l'Agence nationale pour l'emploi (ANPE) = demandes d'emploi en fin de mois.	1990 ANPE, Insee	3,5	1,5	5,1	x			x
Indice conjoncturel de fécondité	L'indice conjoncturel de fécondité représente le nombre moyen de naissances par femme	1989-91 Insee	1,8	1,4	2	x			x
Mères < 18 ans	Pourcentage de mères de moins de 18 ans	1992 Insee	0,6	0,3	1,4	x			x
Mères ≥ 38 ans	Pourcentage de mères de 38 ans et plus	1992 Insee	4,3	3,3	5,6	x			x
75 ans ou plus en zone rurale	Pourcentage de personnes âgées de 75 ans et plus vivant dans des communes rurales.	1990 Insee	29,9	10,2	60,5	x			x
Prestations légales dans les ressources des ménages	Pourcentage de ménages dont les prestations légales représentent la totalité des ressources. Les prestations légales sont versées par les caisses d'allocations familiales au titre du Fonds national des prestations familiales, du Fonds national de l'habitat, du Fonds national d'aide au logement et du Compte de tiers spécifique pour le RMI.	1996 Cnaf	14	11,8	25,2	x			x

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Indicateurs d'offre de soins									
Lits de soins de courte durée	Taux d'équipement en lits publics et privés d'hospitalisation complète pour les soins de courte durée (médecine, chirurgie, gynéco-obstétrique). (nb lits pour 1000 hab)	1993 Drees, Insee	4,8	4	6			x	
Lits de soins de suite	Taux d'équipement en lits publics et privés d'hospitalisation complète pour les soins de suite (réadaptation fonctionnelle et autres lits de suite tels que convalescence). (nb lits pour 1000 hab)	1993 Drees, Insee	1,6	0,9	2,8			x	x
Médecins généralistes libéraux	Densité de médecins libéraux généralistes (pour 1000 habitants)	1995 Drees	114,3	93,6	148,5			x	x
Médecins spécialistes libéraux	Densité de médecins libéraux spécialistes (pour 1000 habitants)	1995 Drees	86,5	45,9	134,7			x	x
Mode conventionnel des médecins libéraux	Pourcentage d'omnipraticiens et de spécialistes de secteur 2 parmi les médecins libéraux.	1993 Cnamts	28,4	9,1	52,3				
Cardiologie	Densité de cardiologues libéraux (pour 100 000 habitants). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1993 Drees, Insee	6,4	3,6	12,6				
Pédiatrie	Densité de pédiatres libéraux (pour 100 000 habitants de moins de 16 ans). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1993 Drees, Insee	25,8	11,4	40,2				
Radiologie	Densité de radiologues libéraux (pour 100 000 habitants). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1993 Drees, Insee	8,2	5,5	12,9				
Gynéco-obstétrique	Densité de gynécologues-obstétriciens libéraux (pour 100000 femmes de 15-49 ans). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1993 Drees, Insee	40	27,2	55,1				
Psychiatrie et neuro-psychiatrie	Densité de psychiatres et de neuro-psychiatres libéraux (pour 100 000 habitants). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1993 Drees, Insee	11	3,5	19,4				
Infirmiers salariés	Densité d'infirmiers salariés (y compris psychiatriques) pour 100 000 habitants. Les salariés sont des salariés exclusifs.	1994 Drees, Insee	493,9	354,9	636,4				
Infirmiers libéraux	Densité d'infirmiers libéraux pour 100 000 habitants. Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1994 Drees, Insee	79,6	47,7	198,1			x	x

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Masseurs-kinésithérapeutes salariés	Densité de masseurs-kinésithérapeutes salariés( pour 100 000 habitants). Les salariés sont des salariés exclusifs.	1994 Drees, Insee	17,8	9	27,2				
Masseurs-kinésithérapeutes libéraux	Densité de masseurs-kinésithérapeutes libéraux pour 100 000 habitants. Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1994 Drees, Insee	58,3	34,3	95,2		x		x
Chirurgiens-dentistes	Densité de chirurgiens-dentistes libéraux (pour 100000 habitants). Les libéraux regroupent les libéraux exclusifs et les libéraux partiels.	1994 Drees, Insee	63,2	36,9	90		x		x
Capacité d'hospitalisation en psychiatrie générale	Taux d'équipement en hospitalisation complète ou partielle en psychiatrie générale (adultes) pour 1000 habitants. L'hospitalisation partielle inclut le placement familial thérapeutique.	1993 Drees, Insee	1,7	1	3,2		x		
Capacité d'hospitalisation en psychiatrie infanto-juvénile	Taux de lits et places en hospitalisation complète ou partielle infanto-juvénile pour 1000 enfants de 0 -16 ans inclus.	1993 Drees, Insee	0,9	0,5	1,4		x		
Modes de prise en charge en psychiatrie	Pourcentage de patients de la file active pris en charge en ambulatoire ou à temps partiel. La file active des secteurs de psychiatrie est l'ensemble des patients vus au moins une fois par an par l'un des membres de l'équipe du secteur.	1993 DGS , Insee	67,8	51,9	74,6				
Lits de gynécologie-obstétrique	Taux d'équipement en lits de gynécologie-obstétrique publics et privés installés (pour 1000 femmes de 15 à 44 ans).	1993 Drees, Insee	2,2	1,9	3,3				
Lits de pédiatrie (médecine)	Taux d'équipement en lits de pédiatrie installés (pour 1000 enfants de moins de 16 ans). En pédiatrie, les enfants sont pris en charge jusqu'à l'âge de 15 ans et trois mois.	1993 Drees, Insee	1,1	0,8	1,5				
Hébergement collectif pour personnes âgées	Taux d'équipement en places d'hébergement pour 1000 personnes de 75 ans ou plus : logements-foyers (LF), maisons de retraite (MR), services de soins de longue durée (LS), autres structures.	1992 Drees, Insee	166,8	67	239,2		x		x
Hébergement collectif médicalisé pour personnes âgées	Taux d'équipement en places de sections en cure médicale (SCM) et en lits de soins de longue durée (LS) pour 1000 personnes de 75 ans ou plus. Les sections de cure médicale sont destinées à prendre en charge les personnes nécessitant un suivi médical et des soins dans les logements-foyers et dans les maisons de retraite.	1992 Drees, Insee	53,2	26,2	76,8		x		
Soins infirmiers à domicile pour personnes âgées	Pourcentage de bénéficiaires de soins infirmiers à domicile parmi la population de 65 ans ou plus.	1992 Cnamts Insee	0,5	0,2	0,9		x		

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Aide ménagère à domicile	Pourcentage de bénéficiaires de l'aide médicale à domicile parmi la population des 65 ans ou plus.	1992 Cnav, Drees, Insee	5,6	4,4	7,2		x		
Toxicomanes pris en charge	Taux de toxicomanes pris en charge dans les structures sanitaires et sociales (pour 100 000 habitants).	1994 Drees, Insee	34,7	10	54				x
Bénéficiaires de l'AES	Taux de bénéficiaires de l'allocation d'éducation spéciale (AES) pour 1000 jeunes de moins de 20 ans. L'AES est une prestation familiale versée aux personnes qui ont la charge d'un enfant de moins de 20 ans handicapé.	1993 Cnaf, MSA, Insee	6	4,8	8,3				x
Jeunes en éducation spéciale	Taux de jeunes en éducation spéciale (pour 1000 jeunes de moins de 20 ans).	1993 Drees, Insee	7,2	4	12,5				
Centres d'aide par le travail	Taux de personnes en ateliers protégés (AP) et en centres d'aide par le travail (CAT) (pour 100 000 personnes de 20-59 ans).	1993 Drees, Insee	287,1	156,8	463,5		x		
Structures pour handicapés lourds	Taux de personnes dans les structures pour adultes lourdement handicapés : foyers de vie (ou foyers occupationnels), maisons d'accueil spécialisées (MAS), foyers à double tarification (FDT). (pour 100 000 pers. De 20-59 ans)	1993 Drees, Insee	113,4	29,1	359,2		x		
Médecins de PMI par femme enceinte	Nombre de médecins (temps complet ou équivalent temps plein) de Protection maternelle et infantile (PMI) par femme enceinte.	1993 Drees, Insee	2,34.10 <sup>3</sup>	1,24.10 <sup>3</sup>	3,26.10 <sup>3</sup>		x		
Puéricultrices de PMI par naissances	Nombre de puéricultrices de Protection maternelle et infantile (PMI) par naissance.	1993 Drees, Insee	4,98.10 <sup>3</sup>	3,12.10 <sup>3</sup>	6,37.10 <sup>3</sup>		x		
Encadrement par la santé scolaire	Taux d'encadrement en nombre de médecins (équivalent temps plein) rapporté au nombre d'élèves. Le service de santé scolaire, dénommé service de promotion de la santé en faveur des élèves, comprend des médecins titulaires à temps plein ou partiel, et des vacataires, ainsi que des infirmières de santé scolaire.	1995 Minist. Educ. Nat.	1,39.10 <sup>4</sup>	1,24.10 <sup>4</sup>	1,62.10 <sup>4</sup>				
Encadrement par les infirmières scolaires	Taux d'encadrement en nombre d'infirmières rapporté au nombre d'élèves. Le service de santé scolaire, dénommé service de promotion de la santé en faveur des élèves, comprend des médecins titulaires à temps plein ou partiel, et des vacataires, ainsi que des infirmières de santé scolaire	1995 Minist. Educ. Nat.	4,05.10 <sup>4</sup>	3,26.10 <sup>4</sup>	7,46.10 <sup>4</sup>				

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Indicateurs d'état de santé									
Mortalité néonatale	Taux de mortalité néonatale (décès avant 28 jours rapportés au nombre de naissances vivantes).	1993-95 Insee	3,1	2,2	4,1				
Mortalité post-néonatale	Taux de mortalité post-néonatale (décès entre 28 jours et 1 an rapportés au nombre de naissances vivantes).	1993-95 Insee	2,6	1,17	3,3				x
Mortalité infantile	Taux de mortalité infantile (décès des enfants de moins de 1 an rapportés au nombre de naissances vivantes).	1993-95 Insee	5,7	4,69	6,7			x	
Mortalité des enfants de 1 à 4 ans	Taux de mortalité des enfants de 1 à 4 ans (nb de décès rapportés au nombre d'enfants du même âge). Les valeurs sont des moyennes sur trois ans.	1993-95 Inserm, Insee	11,5	8,39	17,13				
Mortalité des enfants de 5 à 14 ans	Taux de mortalité des enfants de 5 à 14 ans (nb de décès rapportés au nombre d'enfants du même âge). Les valeurs sont des moyennes sur trois ans.	1993-95 Inserm, Insee	7,2	5,38	15,66				
Mortalité prématurée chez les hommes	Taux comparatif de mortalité prématurée (avant 65 ans) chez les hommes, pour l'ensemble des causes de mortalité.	1993-95 Inserm, Insee	340,7	272,3	453,6				
Mortalité prématurée chez les femmes	Taux comparatif de mortalité prématurée (avant 65 ans) chez les femmes, pour l'ensemble des causes de mortalité.	1993-95 Inserm, Insee	138,6	121,6	178,5				
Mortalité par cardiopathies ischémiques hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	119,5	95,2	155,9				
Mortalité par cardiopathies ischémiques femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	57,2	45,9	78,9				
Mortalité par cardiopathies ischémiques (2 sexes)	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	82,4	67,7	105,4			x	x
Mortalité par maladies vasculaires cérébrales hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	91,3	70,5	122			x	x
Mortalité par maladies vasculaires cérébrales femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	68,2	54,3	97,5			x	x
Mortalité par maladies vasculaires cérébrales (2 sexes)	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	77,4	60,9	106,7				
Mortalité par cancer du côlon-rectum hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	37,5	26	49,7				x
Mortalité par cancer du côlon-rectum femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	21,8	19,1	28,1				x

Indicateur	Définition	date et sources	moyenne (France métro)	mini	maxi	Indicateurs retenus dans les ACP			
						socio-démo	offre soins	état santé	globale

Mortalité par cancer de l'utérus	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	9	7,4	12,1				x
Mortalité par cancer du sein	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	32,1	26,9	40,1			x	
Mortalité par cancer de la prostate	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	47	28,2	57,1			x	x
Mortalité par cirrhose hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	22,7	13,1	38,8			x	x
Mortalité par cirrhose femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	8,4	3,2	20,4			x	x
Mortalité par psychose alcoolique et alcoolisme hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	7,3	3,4	12,6			x	x
Mortalité par cancers des VADS hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	42,2	26	71,4			x	x
Mortalité par cancer du poumon hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	81,1	58,3	116,8			x	x
Mortalité par cancer du poumon femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	10,1	7,2	14,2			x	
Mortalité par bronchite chronique hommes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	41,8	33,7	67,7			x	x
Mortalité par bronchite chronique femmes	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	14,4	11,9	19,5			x	x
Mortalité par suicide	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	20,5	11,8	32			x	x
Mortalité par accidents de la circulation	Taux comparatif de mortalité tous âges pour cette cause	1993-95 Inserm, Insee	14,2	9,1	22,8			x	
Maladies profession.	Taux de premières indemnités pour 100 000 salariés du régime général d'assurance maladie.	1993-95 Cnamts, Insee	57,2	24,34	148,8				
Interpellations pour usage de stupéfiants	Taux de toxicomanes interpellés pour usage ou usage-revente de stupéfiants pour 100 000 habitants.	1994 Ocrtis, Drees, Insee	89	38	156				
Incidence du sida	Nombre de cas de sida diagnostiqués dans l'année rapporté à la population (par million d'habitants).	1996 InVS, Insee	44,3	19,4	150,9			x	x
Incidence de la tuberculose	Nombre de nouveaux cas de tuberculose déclarés pendant l'année (nombre pour 100 000 habitants).	1995 InVS, Insee	14,7	5,87	35,55				x

## 2 - Méthodes d'analyse

### Analyse en composantes principales (ACP)

L'ACP s'applique à des tableaux croisant des individus et des variables quantitatives. Dans cette étude, les individus sont les régions françaises et les variables quantitatives sont les variables d'état de santé, d'offre de soins et socio-démographiques retenues.

L'ACP cherche à évaluer :

- les ressemblances entre les individus, c'est-à-dire s'ils possèdent des valeurs proches pour l'ensemble des variables,
- les liaisons entre les variables, c'est-à-dire leurs corrélations.

Un individu peut être représenté comme un point de l'espace vectoriel à  $K$  dimensions  $R_K$  dont chaque dimension représente une variable. L'ensemble des individus constitue le nuage  $N_I$  dont le centre de gravité est confondu avec l'origine (car en ACP le tableau des données est centré).

Dès que  $K$  est supérieur à 3, l'étude directe du nuage est impossible. Mais, l'ACP fournit des images planes d'un nuage situé dans un espace de grande dimension. Le nuage des individus peut donc être étudié par cette méthode.

Une variable peut être représentée comme un vecteur de l'espace vectoriel à  $I$  dimensions noté  $R_I$  dont chaque dimension représente un individu. L'ensemble des extrémités des vecteurs représentant les variables constitue le nuage  $N_K$ . Les variables en ACP étant centrées réduites, chaque vecteur représentant une variable a pour norme 1. Le nuage  $N_K$  est donc situé sur une sphère de rayon 1.

Le cosinus de l'angle formé par les vecteurs représentant deux variables est égal au coefficient de corrélation entre ces deux variables (car les variables sont centrées et réduites). L'étude directe des coefficients de corrélation entre les variables est impossible du fait de la dimension  $R_I$  du nuage. Mais l'ACP fournit des variables synthétiques (appelées composantes principales) qui constituent un résumé de l'ensemble des corrélations entre les variables et qui sont la base d'une représentation plane approchée des variables et de leurs angles.

### *Les individus*

Le but est de déterminer une suite  $(u_s)$  de directions de  $R_K$  appelées axes factoriels qui prises deux à deux définissent des plans factoriels sur lesquels on projette le nuage  $N_I$  des individus afin d'obtenir une image plane des distances inter-individuelles.

L'individu  $i$  se projette en  $H_i$  sur  $u_1$ . On cherche  $u_1$  qui rend maximum  $S_i OH_i^2$ , c'est-à-dire qui rend maximum l'inertie par rapport à l'origine  $O$ . L'inertie est la moyenne des carrés des distances des points au centre de gravité. Elle traduit la dispersion du nuage autour de son centre de gravité.

Puis on cherche  $u_2$  orthogonal à  $u_1$  et rendant maximum  $S_j OH_j^2$ . Rendre maximum  $S_j OH_j^2$  revient à rendre minimum  $S_j (jH_j)^2$  (critère des moindres carrés). Ainsi, les axes factoriels rendent minimum l'écart entre le nuage des individus et sa projection.

### *Les variables*

Le but est d'obtenir une suite de variables synthétiques  $(v_s)$  sur laquelle projeter le nuage  $N_K$  des variables afin d'obtenir une image plane des corrélations entre les variables.

Les variables étant centrées réduites, leur projection sur  $v_1$  est égale à leur coefficient de corrélation avec  $v_1$ . Rechercher  $v_1$  qui rend maximum  $S_k OH_k^2$  (c'est-à-dire qui rend maximum la somme des cosinus carrés des angles entre les vecteurs et leur projection) équivaut à rechercher la combinaison linéaire la plus liée à l'ensemble des variables. Les axes factoriels étant orthogonaux deux à deux, on obtient une suite de variables synthétiques  $(v_s)$  non corrélées entre elles qui résument l'ensemble des variables initiales.

L'importance d'un axe se mesure par le pourcentage d'inertie qui lui est associé, c'est-à-dire le rapport entre l'inertie associée au facteur (ou valeur propre) et l'inertie totale du nuage étudié. Le pourcentage d'inertie associé à un axe indique la qualité de la représentation du nuage par l'axe. C'est pourquoi, on s'intéresse généralement aux axes dont le pourcentage d'inertie est élevé.

Les angles entre les variables étant déformés par la projection, on limite l'étude des plans factoriels aux variables bien représentées (c'est-à-dire dont l'image est proche du cercle de corrélation). De même, on s'intéresse plus spécialement aux régions éloignées de l'origine des axes. C'est pourquoi, seules les variables fortement corrélées aux axes et

seules les régions ayant des contributions aux axes importantes sont représentées sur les graphiques.

*Une première approche séparée des graphiques permet :*

- 1 - de regrouper visuellement les variables liées entre elles (car des variables très corrélées positivement sont situées du même côté d'un axe);
- 2 - d'identifier visuellement les liaisons interrégionales pour chaque axe (car les régions situées du même côté d'un axe sont proches et les régions situées de part et d'autre de l'origine s'opposent pour cet axe).

Les graphiques des variables et des individus peuvent être étudiés simultanément. En effet, une région se situe du côté des variables pour lesquelles elle a de fortes valeurs et à l'opposé des variables pour lesquelles elle a de faibles valeurs. Cette étude conjointe des deux types de graphiques permet de trouver pour chaque axe un dénominateur qui, à la fois, relie les variables situées du même côté d'un axe et oppose les variables situées de part et d'autre de l'origine. On obtient ainsi la signification des axes factoriels.

Quand une région est trop dissemblable des autres, elle perturbe le modèle qui n'est alors plus que le reflet de cette seule région. C'est le cas de l'Ile-de-France par exemple dans l'analyse socio-démographique. Il est cependant possible de construire un modèle sans faire intervenir cette région tout en gardant la possibilité de la comparer aux profils obtenus. Cette région porte alors le nom dans l'analyse de «région supplémentaire».

## **Classification**

Les méthodes de classification ou de typologie ont pour but de regrouper les individus en un nombre restreint de classes homogènes. Il s'agit de décrire les données en procédant à une réduction du nombre des individus.

On distingue 2 grands types de méthodes de classification :

- les méthodes non hiérarchiques qui produisent directement une partition en un nombre fixé de classes ;
- les méthodes hiérarchiques qui produisent des suites de partitions en classes de plus en plus vastes.

La méthode utilisée ici est la classification ascendante hiérarchique : chaque classe d'une partition est incluse dans une classe de la partition suivante. La suite des partitions obtenues est usuellement représentée sous la forme d'un arbre de classification.

A chaque partition correspond une valeur numérique appelée niveau d'agrégation représentant le niveau auquel ont lieu les regroupements. Plus cet indice est élevé, plus les parties regroupées sont hétérogènes.

Le principal problème des méthodes de classification hiérarchique consiste à définir le critère de regroupement de 2 classes, ce qui revient à définir une distance entre classe. Tous les algorithmes de classification hiérarchique se déroulent de la même manière : on recherche à chaque étape les 2 classes les plus proches, on les fusionne, et on continue jusqu'à ce qu'il n'y ait plus qu'une seule classe.

La procédure RECIP de SPAD effectue la classification hiérarchique d'un ensemble d'individus caractérisés par leurs premières coordonnées factorielles. Cette procédure utilise le critère d'agrégation de Ward qui minimise l'inertie intra-classe et maximise l'inertie inter-classe. Ce critère de regroupement suit le principe suivant : fusionner les 2 classes pour lesquelles la perte d'inertie est la plus faible. Ce qui revient à réunir les 2 classes les plus proches en prenant comme distance entre 2 classes la perte d'inertie que l'on encourt en les regroupant.

Pour améliorer la partition obtenue, une procédure d'agrégation permet d'augmenter l'inertie entre les classes à chaque itération. Cette procédure (appelée consolidation) s'arrête dès que l'inertie entre les classes cesse de croître de façon sensible. A la fin de la classification, on dispose de la meilleure classification possible.

La typologie, une fois constituée, peut-être décrite en fonction de la situation des régions à l'égard des indicateurs : pour chaque classe, on calcule la moyenne des valeurs prises par chacun des individus qui la composent pour chacune des variables et l'on obtient ainsi un profil d'ensemble de la classe.

### 3 - Bibliographie

- Aïach P., Carr-Hill R. (1998) : Les inégalités sociales d'état de santé en France et en Grande-Bretagne, ed. INSERM.
- Aïach P., Cebe D. (1994) : Les inégalités sociales de santé, Problèmes économiques, pp. 20-29.
- Anderson RT., Sorlie P., Backlund E., Johnson N., Kaplan GA. (1997) : Mortality effects of community socioeconomic status, *Epidemiology*, 8(1), pp. 42-47.
- Babazono A, Hillman AL. (1994) : A comparison of international health outcomes and health care spending, *International journal of technology assessment in health care*, 10(3), pp. 376-381.
- Badeyan G., Pinteaux A., Roussille B. (1996) : La santé en France : un bilan contrasté, *Données sociales 1996*, INSEE, pp. 265-275.
- Blank N., Diderichsen F. (1996) : Inequalities in health : the interaction between socio-economic and personal circumstances, *Public Health*, 110(3), pp. 157-162.
- Bobak M., Koupilova I., Prikazsky V., Kriz B. (1996) : Relation between socioeconomic factors and mortality in districts in the Czech Republic, *Cas Lek Cesk*, 135(23), pp. 753-757.
- Brooks-Gunn J., Duncan GJ. (1997) : The effects of poverty on children, *Future Child*, 7(2), pp. 55-71.
- Charlton BG. (1994) : Is inequality bad for the national health ?, *The Lancet*, 343, pp. 221-222.
- Charlton JRH, Hartley M., Silver R., Holland W. (1983) : Geographical variation in mortality from conditions amenable to medical intervention in England and Wales, *The Lancet*, 26, pp. 691-696.
- Chetter IC., Dolan P., Spark JI., Scott DJ., Kester RC. (1997) : Correlating clinical indicators of lower-limb ischaemia with quality of life, *Cardiovasc Surg*, 5(4), pp. 361-366.
- Commissariat Général Du Plan (1993) : Santé 2010 - Equité et efficacité du système de santé : les enjeux, *La documentation française*.
- Desai S., Alva S. (1998) : Maternal education and child health : is there a strong causal relationship ?, *Demography*, 35(1), pp. 71-81.
- Desplanques G. (1993) : L'inégalité sociale devant la mort, *Données sociales 1993*, INSEE, pp. 251-258.
- Dinh QC. (1990) : Le relief social de la mortalité infantile, *Données sociales 1990*, INSEE, pp. 245-250.
- Doblhammer G., Kytir J. (1998) : Social inequalities in disability-free and healthy life expectancy in Austria, *Wien Klin Wochenschr*, 110(11), pp. 393-396.
- Elola J., Daponte A., Navarro V. (1995) : Health indicators and the organisation of health care systems in Western Europe, *American journal of public health*, vol. 85, n°10.
- Escofier B., Pagès J. (1990) : Analyses factorielles simples et multiples - objectifs, méthodes et interprétation, Dunod.
- Evans RG., Barer ML., Marmor TR (1996) : Etre ou ne pas être en bonne santé - biologie et déterminants sociaux de la maladie, *Les presses de l'université de Montréal*.

Fédération nationale des observatoires régionaux de la santé (1997) : La santé observée dans les régions de France.

Feinstein JS. (1993) : The relationship between socioeconomic status and health : a review of the literature, *The milbank quarterly*, 71(2), pp. 279-322.

Foucart T. (1997) : l'analyse des données, DIDACT Statistiques Presses Universitaires de Rennes

Foulon D. (1993) : La fréquentation hospitalière régionale, *Solidarité santé - Etudes statistiques*, n°2.

Frenkiel J., Hecketsweiler P. (1992) : Les inégalités inter-régionales de l'offre de soins en France, *Solidarité santé - Etudes statistiques*, n°4.

Fuchs VR. (1986) : Some economic aspects of mortality in developed countries, *The health economy*, Cambridge Mass., Harvard university press, pp. 181-199.

Gibert F., Letourmy A. (1974) : Santé, environnement, consommations médicales. Un modèle et son estimation à partir des données de mortalité, *Cerebe*.

Guignon N. (1990) : Alcool et tabac : des plaisirs qui empoisonnent, *Données sociales 1990*, 6, pp. 254-257.

Guignon N. (1994) : Les consommations d'alcool, de tabac et de psychotropes en France en 1991-1992, *Echanges santé-social*, 74, pp. 40-43.

Hammarström A. (1994) : Health consequences of youth unemployment, *Public health*, 108, pp. 403-412.

Haut Comité de Santé Publique (1994) : La santé en France - annexe : travaux des groupes thématiques, *La documentation française*.

Helmert U., Shea S. (1994) : Social inequalities and health status in Western Germany, *Public health*, 108, pp. 341-356.

Hitiris T., Posnett J. (1992) : The determinants and effect of health expenditure in developed countries, *Journal of health economics*, 11, pp. 173-181.

Holcik J., Koupilova I., Gerylovova A. (1997) : Assessment of health needs, their geographic distribution and association of health care indicators in Districts within the Czech Republic, *Cas Lek Cesk*, 136(21), pp. 662-665.

INSEE (1990) : La France et ses régions.

Jougla E, Ducimetière P., Bouvier-Colle MH, Hatton F. (1987) : Relation entre le niveau de développement du système de soins et le niveau de la mortalité évitable selon les départements français, *Revue d'épidémiologie et santé publique*, 35, pp. 365-377.

Karim E., Mascie-Taylor CG. (1997) : The association between birthweight, sociodemographic variables and maternal anthropometry in urban sample from Dakha, Bangladesh, *Ann Hum Biol*, 24(5), pp. 387-401.

Leal M. do C., Szwarcwald CL. (1997) : Characteristics of neonatal mortality in the State of Rio de Janeiro, Brazil, in the 1980's : a spatio-temporal analysis, *Rev Saude Publica*, 31(5), pp. 457-465.

Lebart L., Sandier S., Tonnelier F. : Aspects géographiques du système des soins médicaux - *Annales du CREDOC N°4* (1974).

Lebart L., Moineau A., Piron M., : Statistique exploratoire multidimensionnelle – 2ème cycle, Editions DUNOD .

Le Roux S. (1984) : Les inégalités devant la santé - rapport au ministre des affaires sociales et de la solidarité nationale, La documentation française.

Lucas-Gabrielle V., Tonnelier F., Vigneron E., : Une typologie des paysages socio-sanitaires en France – CREDES (1998).

Luoto R., Keskimaki I., Reunanen A. (1997) : Socioeconomic variations in hysterectomy : evidence from a linkage study of the finnish hospital discharge register and population census, *J Epidemiol Community Health*, 51(1), pp. 67-73.

Marmot MG., Mc Dowell ME. (1986) : Mortality decline and widening social inequalities, *The Lancet*, 2, pp. 274-275.

Meraud C. (1990) : Inégalités régionales dans la répartition de l'enveloppe sanitaire, *Solidarité santé - Etudes statistiques*, n°6.

Michel E. (1993) : Disparités régionales de la mortalité en France, *Solidarité santé - Etudes statistiques*, n°1.

Or Z. (1987) : Les déterminants de l'état de santé dans les pays industrialisés : une analyse économétrique, thèse d'économie de la santé, université de Paris I.

Perez DJ., McGee R., Campbell AV., Christensen EA., Williams S. (1997) : A comparison of time trade-off and quality of life measures in patients with advanced cancer, *Qual Life Res*, 6(2), pp. 133-138.

Rognerud MA., Kruger O., Gjertsen F., Thelle DS. (1998) : Strong regional links between socio-economic background factors and disability and mortality in Oslo, Norway, *Eur J Epidemiol*, 14(5), pp. 457-463.

Sharland M., Atkinson P., Maguire H., Begg N. (1997) : Lone parent families are independent risk factor for lower rates of childhood immunisation in London, *Commun Dis Rep CDR Rev*, 7(11), pp. 169-172.

Sorlie PD., Backlund E., Keller JB. (1995) : US mortality by economic, demographic, and social characteristics : the national longitudinal mortality study, *American journal of public health*, vol.85, n°7.

Tonnelier F. (1991) : Les inégalités géographiques de densité médicales sont stables depuis près d'un siècle, *Solidarité santé - Etudes statistiques*, n°3.

Tonnelier F. (1997) : Géographie de la santé, *Actualité et dossier en santé publique*, n°19.

Velez-Pastrana MC., Rodriguez Gomez JR., Martinez Lugo M. (1997) : Social support, activity, and health among the elderly, *Bol Asoc Med P R*, 89(10-12), pp. 174-183.

Vogel RL., Ackerman RJ., Is primary care physician supply correlated with health outcomes?, *Int J Health Serv*, 28(1), pp. 183-196.

# Cartes et graphiques

- CARTE 1 Classification réalisée sur les indicateurs socio-démographiques
- GRAPHIQUE 1 Les 22 régions de France métropolitaine et leurs caractéristiques socio-démographiques
- CARTE 2 Classification réalisée sur les indicateurs d'offre de soins
- CARTE 3 Classification réalisée sur les indicateurs d'état de santé
- CARTE 4 Classification réalisée sur l'ensemble des indicateurs
- GRAPHIQUE 2 Les 22 régions de France métropolitaine et leurs caractéristiques socio-sanitaires

## CARTE 1

### Représentation cartographique de la classification réalisée sur les indicateurs socio-démographiques

Cinq groupes de régions (et Ile-de-France, région "supplémentaire")

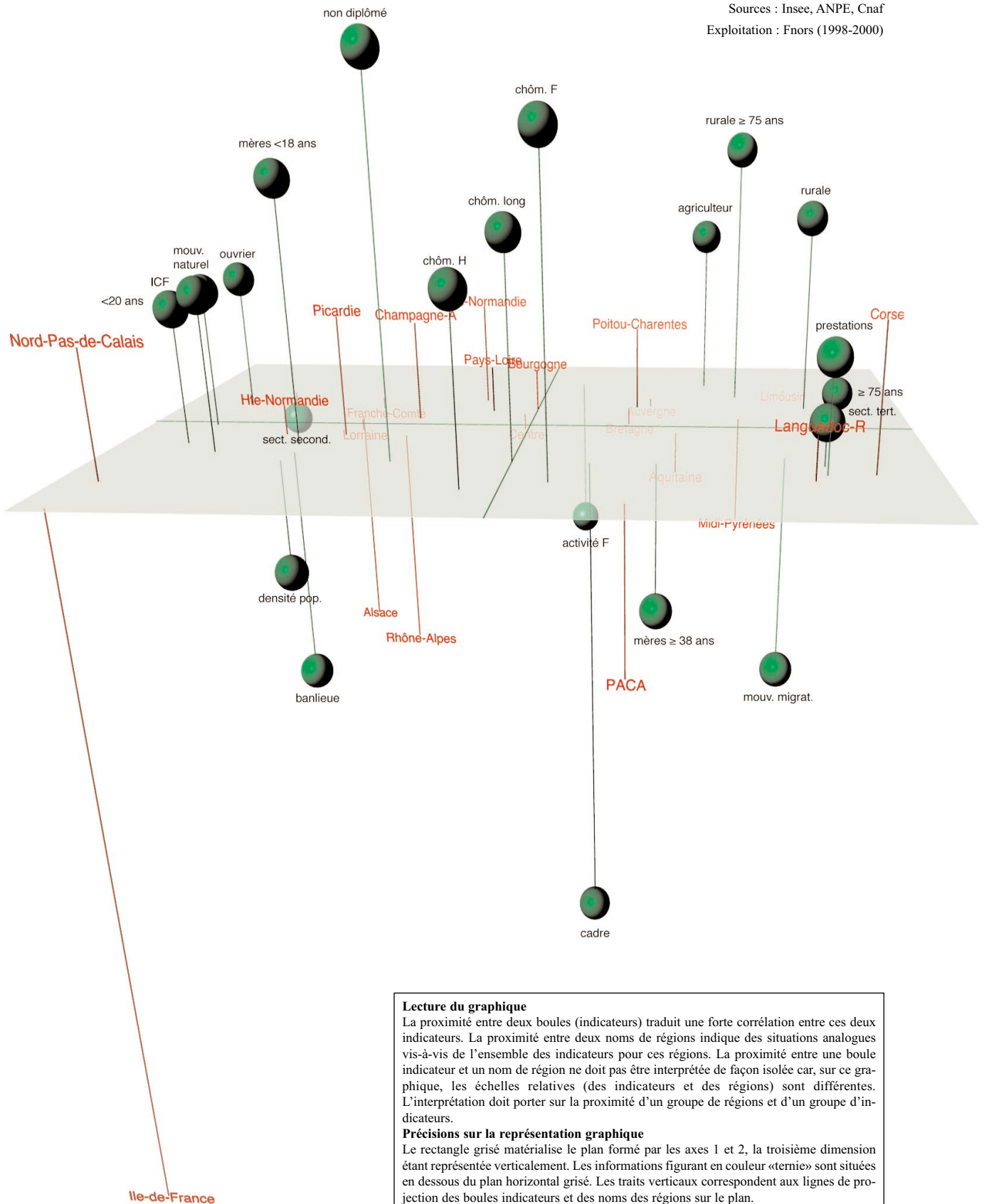


Typologie des régions de France métropolitaine établie à partir de l'analyse en composantes principales (indicateurs socio-démographiques)

Sources : Insee, ANPE, Cnaf  
Exploitation : Fnors (1998-2000)

# GRAHIQUE 1 : Les 22 régions de France métropolitaine et leurs caractéristiques socio-démographiques

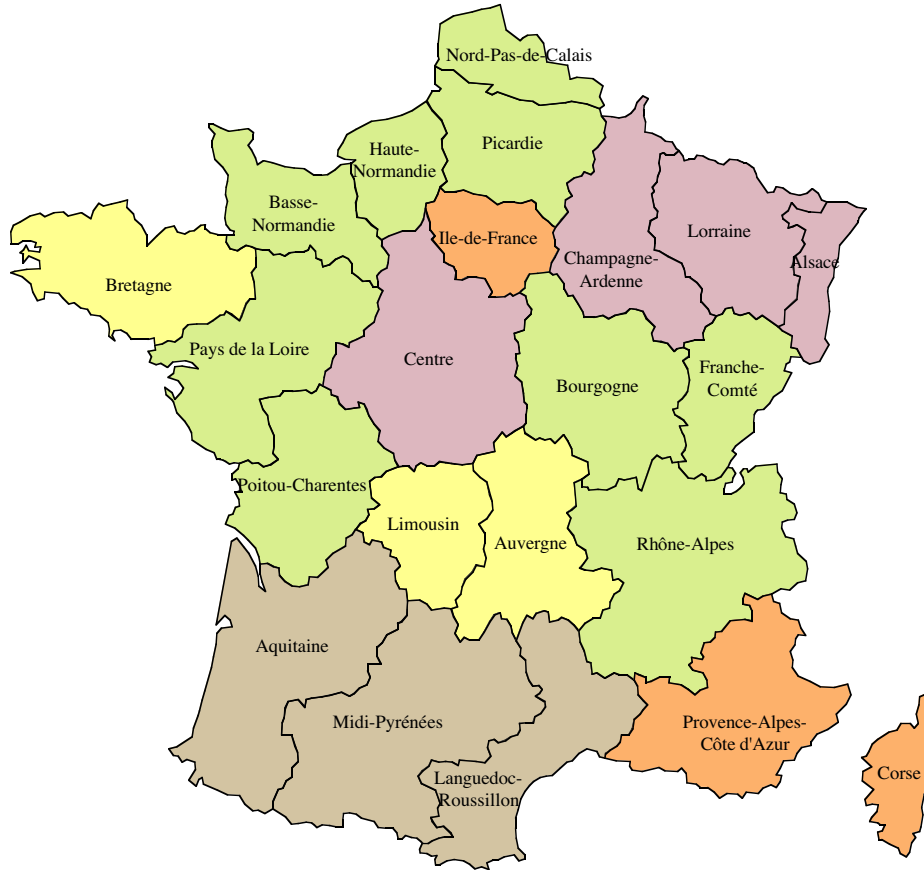
Sources : Insee, ANPE, Cnaf  
Exploitation : Fnors (1998-2000)



## CARTE 2

### Représentation cartographique de la classification réalisée sur les indicateurs d'offre de soins

Cinq groupes de régions



Typologie des régions de France métropolitaine établie à partir de l'analyse en composantes principales (indicateurs d'offre de soins)

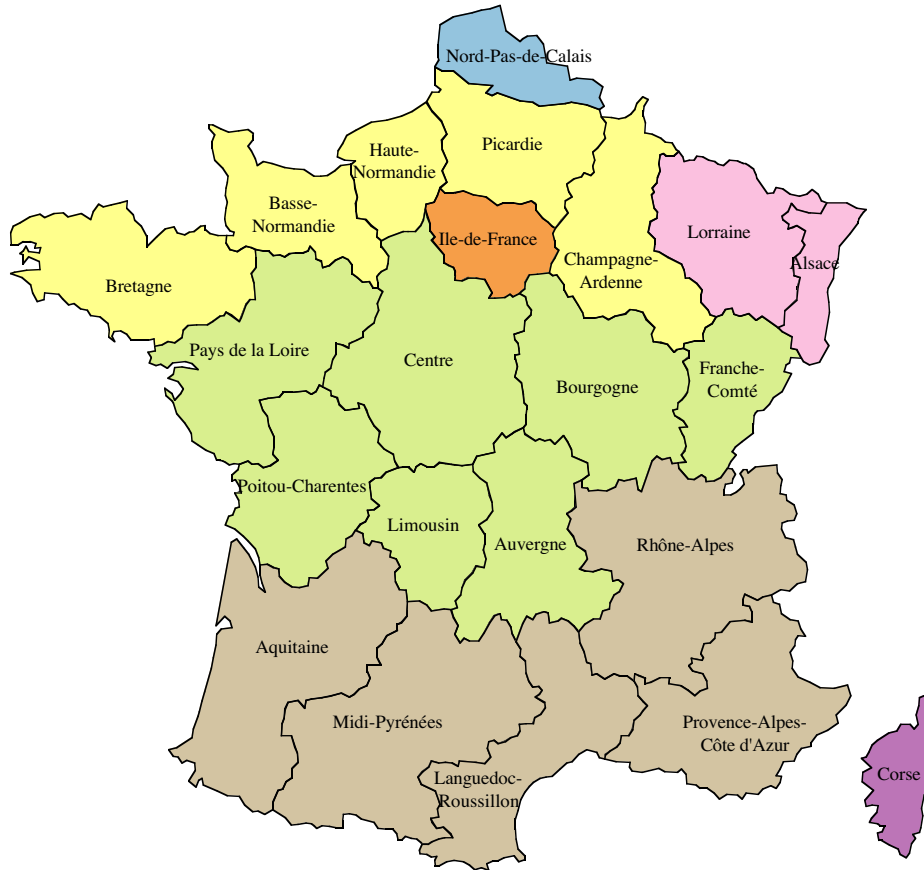
Sources : Drees, DGS, Cnav, Cnamts, Insee

Exploitation : Fnors (1998-2000)

### CARTE 3

#### Représentation cartographique de la classification réalisée sur les indicateurs d'état de santé

Sept groupes de régions



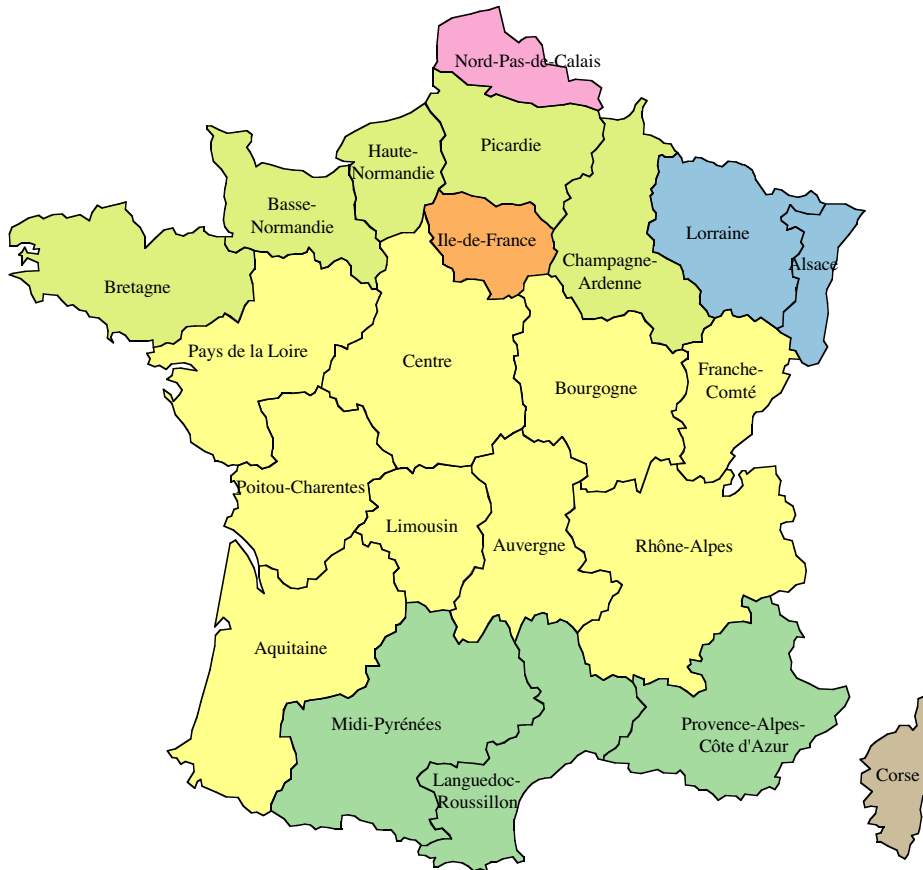
Typologie des régions de France métropolitaine établie à partir de l'analyse en composantes principales (indicateurs d'état de santé)

Sources : Inserm SC8, InVS, Insee  
Exploitation : Fnors (1998-2000)

## CARTE 4

### Représentation cartographique de la classification réalisée sur l'ensemble des indicateurs

Sept groupes de régions



Typologie des régions de France métropolitaine établies à partir de l'analyse en composantes principales (ensemble des indicateurs)

Sources : Insee, ANPE, Cnaf, Drees, MSA, InVS, Inserm SC8  
Exploitation : Fnors (1998-2000)

## GRAPHIQUE 2 : Les 22 régions de France métropolitaine et leurs caractéristiques socio-sanitaires

Sources : Insee, ANPE, Cnaf, Drees, MSA, InVS, Inserm  
SC8Exploitation : Fnoirs (1998-2000)

