

" Les espaces
géographiques ne sont
pas de simples décors,
ils façonnent les
comportements,
les cultures
et les identités
des individus."

Henri Lefebvre

GT MÉTHODOLOGIE

La géographie dans le SNDS

maternité respiratoire
périnatalité santé
métabolisme tumeur endocrinologie
traumatologie troubles
nutrition **ReDSiam** mentaux
Réseau pour mieux utiliser les Données circulatoire
du Système national des données de santé
maladies neurodégénératives
algorithmes données
infectieux ostéo-articulaire

CONTRIBUTEURS

ATTIA Vincent (Drees), BERNARD Marlène (ATIH), BIZOUARD Geoffroy (IQVIA), BONNET Caroline (ORS Bourgogne-Franche-Comté), BOSCO-LEVY Pauline (HORIANA), BROCCA Jérôme (SCN/DNUM ministère de la Santé), CHAN-WAN Guy-Noel (DRSM Réunion), CHESNEAU Julie (Santé publique France), COURTEJOIE Noémie (Drees), DALICHAMPT Marie (ORS Pays de la Loire), DE MARTINI Laure (Cnam), DESJEUX Guillaume (SANOIA), DUVER Sigolène (DRSM Normandie), ESPAGNACQ Maude (Irdes), FANTIN Sophie (Cnam), FRAGUIDEAU Fanny (HEVA), GINOUX Marine (HEVA), GODILLON Lucile (EpiDéclic), GORONFLOT Thomas (CHU de Nantes), GOUSSET Christophe (Irdes), HEUZE Guillaume (Santé publique France), KAMIONKA Julie (Drees), LARRIEU Sophie (HORIANA), LAURENT Émeline (EpiDéclic), LECUYER Anne-Isabelle (EpiDéclic), LEROY Céline (ARS Normandie), LOUACHENI Camélia (ORS Réunion), MARGUERITE Nadège (Santé publique France), MICHAULT Alain (ARS Île-de-France), NESTRIGUE Clément (ARS Île-de-France), NOURRY Annie (CCMSA), PENNEAU Anne (Irdes), REGAERT Camille (Irdes), RIVIERE Sébastien (Cnam), THOMAS Nadège (OR2S), VALLEE Cédric (CCMSA), VALLEE Nélia (CCMSA), VIARD Claude (DCGDR Normandie), VISSIERE Rémi (DGOS)

GT MÉTHODOLOGIE

Qui sommes- nous ?

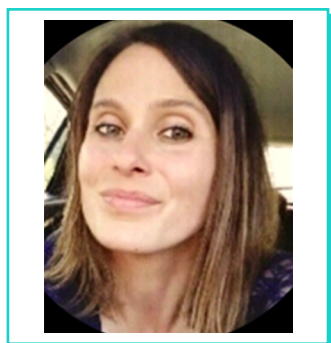
Le GT méthodologie de REDSIAM a été créé en 2018, son ambition est de répondre à des questions d'ordre méthodologique sur le SNDS. Il se veut transversal et en appui des autres GT pour traiter des questions ponctuelles.

Après avoir étudié la question du dénominateur, ce rapport traite de la thématique de la géographie et plus particulièrement de **la localisation du bénéficiaire dans le SNDS**.

Ce GT regroupe **53 membres** à ce jour qui ont choisi de travailler sur cette thématique. Pour faciliter le travail collectif, quatre sous-groupes ont été créés, chacun traitant un aspect de la géographie, à savoir :

- la géographie hospitalière ;
- le choix de la géographie dans la territorialisation de la donnée ;
- la qualité de la géographie ;
- la géographie particulière des publics spécifiques.

Présentation des coordonnateurs



Je m'appelle Céline LEROY

Je suis actuellement chargée d'études au sein de l'ARS Normandie. Depuis de nombreuses années, je nourris un vif intérêt pour les enjeux liés à la géographie des données de santé. L'approche territorialisée de ces données constitue, à mon sens, un levier essentiel pour faciliter la compréhension des territoires. Qu'il s'agisse de dresser un état des lieux socio-sanitaire ou d'évaluer l'impact des politiques publiques, je suis profondément convaincue que l'intégration de la dimension géographique permet d'élaborer une connaissance plus précise, plus pertinente et plus directement opérationnelle !



Je m'appelle Nadège THOMAS

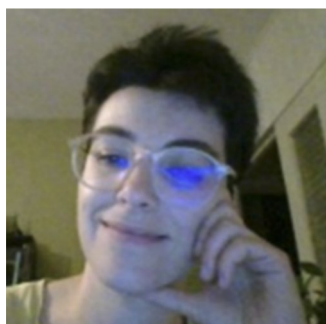
Je travaille depuis près de 20 ans au sein du réseau des observatoires régionaux de la santé (ORS), plus spécifiquement à l'OR2S qui œuvre sur le périmètre géographique des Hauts-de-France et de la Normandie. Depuis toutes ces années, je ne cesse de travailler, avec passion, sur les aspects de territorialisation. Ils fondent des préalables indispensables à la production d'une observation fiable, permettant un accompagnement des politiques publiques se voulant le plus efficient possible et, in fine, des actions à mettre en place, nécessairement adaptées aux spécificités des territoires et des populations.

Présentation des coordonnateurs des sous-groupes



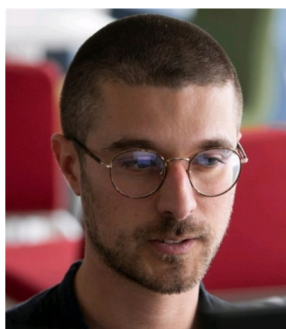
Marlène BERNARD

« *Je m'appelle **Marlène Bernard**, spécialiste des données hospitalières (dites PMSI). Actuellement statisticienne à l'Agence Technique d'Information sur l'Hospitalisation (ATIH) et formatrice sur l'utilisation des données PMSI dans le SNDS, j'ai commencé à m'intéresser à la géographie en santé à l'Observatoire Régional de la Santé (ORS) Rhône-Alpes. J'ai eu le plaisir d'animer le sous-groupe « géographie ville-hôpital », qui a travaillé à définir les notions de géographie disponibles dans le PMSI (patient et établissement de santé), pour aboutir à un cas d'usage d'utilisation de tables de passage entre ville et hôpital du SNDS. J'espère que ces travaux permettront aux utilisateurs de mieux appréhender les différents aspects de la géographie dans le SNDS ».*



Marie DALICHAMPT

« *Je m'appelle **Marie Dalichamp**, au moment où j'ai réalisé ce travail j'étais chargée d'études au sein de l'ORS Pays de la Loire. Depuis plusieurs années, je m'intéresse particulièrement aux enjeux liés à la géographie des données de santé. J'accorde une importance toute particulière à l'approche territorialisée de ces données, qui me semble constituer un levier déterminant pour mieux comprendre les réalités locales ».*



Vincent ATTIA

« *Je m'appelle **Vincent Attia**, au moment où j'ai réalisé ce travail j'étais à la Drees depuis 4 ans et demi au sein du Labsanté, et j'ai animé le sous-groupe 3 consacré à la qualité de l'information géographique et aux corrections. Ces sujets soulèvent de nombreuses questions pour les statisticiens lançant un projet. J'espère que notre travail, basé sur les expertises de nombreuses institutions, les éclairera au mieux »*



Guillaume DESJEUX

« *Je m'appelle **Guillaume Desjeux**, je travaille pour des bureaux d'études, organismes privés qui utilisent pour leurs clients les données du SNDS. Je suis réserviste pour le Service de santé des armées qui a un accès permanent au SNDS pour la population de la Caisse nationale militaire de sécurité sociale et je suis également médecin DIM. J'ai été pendant 13 ans médecin conseil de la Caisse nationale militaire de sécurité sociale en charge des études en santé via les données produites par l'assurance maladie. Dans les différentes facettes des métiers exercés la question de la géographie a été souvent importante. En tant que producteur et analyste de la donnée j'espère apporter un éclairage à ce document. »*

Quelques concepts

Assuré (ou ouvrreur de droit) : personne au nom de laquelle les droits à remboursement sont ouverts. Concrètement, c'est l'individu assuré social dont le NIR (Numéro d'Inscription au Répertoire) est utilisé pour le remboursement des soins.

Ayant droit : personne qui ne dispose pas de droits ouverts à son propre nom, mais est rattachée au compte d'un assuré (par exemple un enfant mineur dépendant de l'un de ses parents). Elle est identifiée dans les données par un pseudo-NIR construit à partir du NIR de l'assuré, de sa date de naissance et de son sexe.

Note importante : depuis le 1^{er} janvier 2016 et l'instauration de la Protection Universelle Maladie (PUMa), les jeunes de plus de 16 ans sont généralement devenus assurés à part entière, donc le statut d'ayant droit ne concerne désormais que des enfants mineurs.

Bénéficiaire (des soins) : personne qui reçoit le soin, à savoir l'assuré ou un ayant droit. Dans le SNDS, le bénéficiaire est identifié par un NIR pseudonymisé (BEN_NIR_ANO), distinct du pseudo-NIR du régime (BEN_NIR_PSA).

RNIAM : le RNIAM est un répertoire national chargé de recenser et certifier les éléments d'identification des personnes couvertes par l'Assurance maladie. Il permet aussi d'identifier l'organisme de rattachement principal et, le cas échéant, l'organisme complémentaire.

NIR certifié : un NIR certifié est un Numéro d'Inscription au Répertoire qui a été validé par l'Insee (vérification de l'état civil : nom, prénoms, date et lieu de naissance). Les individus ayant un NIR certifié par le RNIAM sont identifiables dans le SNDS *via* le champ (BEN_CDI_NIR = code '00').

En résumé

Rôle	Description
Assuré	Titulaire des droits ouverts, dont le NIR sert de référence au remboursement
Ayant droit	Personne rattachée au compte d'un assuré, sans droit personnel ouvert (souvent un enfant)
Bénéficiaire	Personne qui bénéficie effectivement d'un soin (assurée ou ayant droit)



Pour aller plus loin :

vous pouvez vous référer à la page [Glossaire](#) de la documentation du SNDS sur le site du *Health Data Hub* (HDH) entraide.

Introduction

À première vue, **la géographie** dans le Système National des Données de Santé (SNDS) est un sujet qui peut sembler trivial. Peu d'études, voire aucune, documentent précisément la notion de géographie, comme si finalement ce choix relevait d'une évidence ou pire qu'il pouvait ne pas avoir d'impact sur le résultat obtenu.

Pourtant, ce choix de la géographie revêt une importance capitale, **la géographie est un facteur déterminant** dans le développement des individus. Elle influence leur mode de vie, leur culture et leur identité. Selon leur lieu de naissance ou de résidence, les personnes sont confrontées à des environnements différents, qui façonnent leurs habitudes de vie, leurs pratiques sociales et économiques. Ces particularités géographiques influencent non seulement leur quotidien, mais aussi leur rapport à la santé, en modifiant leur perception des risques, leur accès aux soins et leur adhésion aux pratiques de prévention.

La **territorialisation des données de santé** constitue donc un enjeu crucial pour les politiques publiques et la gestion de la santé à l'échelle locale. Néanmoins, la territorialisation des données n'est pas une question simple et soulève de nombreuses interrogations :

- En premier lieu, celui de son choix : *Existe-t-il plusieurs choix possibles ? Si oui lesquels ?*
- De sa pertinence selon l'objet de l'étude : *Que cherche-t-on à montrer ?*
- De sa variabilité dans le temps : *La notion de géographie est-elle toujours la même ?*
- De la réalité à laquelle cette géographie fait référence : *Quelle géographie récupère-t-on ?*
- Et enfin, de la qualité de l'information et du degré de fiabilité : *Cette information est-elle valide ?*

L'objectif de ce rapport est bien de répondre aux questions que les utilisateurs se posent sur la géographie dans le SNDS historique¹.

Dans une **première partie**, ce rapport reviendra sur la **dimension géographique** présente dans le SNDS, que ce soit pour les assurés sociaux, les professionnels de santé, etc., et sur son sens.

Une **seconde partie** soulèvera la question de la **qualité de l'information**, explicitera les points d'attention à avoir. Un référencement sous forme d'outil, table de corrections, sera proposé aux utilisateurs pour faciliter les éventuelles corrections géographiques à opérer.

Enfin, une **troisième partie** sera consacrée aux **publics spécifiques** (les enfants, les bénéficiaires de l'aide médicale d'état...) et à leur géographie. Souvent particulière, cette géographie n'est pas toujours celle attendue.

Enfin, la question du **choix de la géographie** alimentera la **dernière partie** de ce rapport. Ce travail a été pensé comme une aide à la décision pour permettre à chaque utilisateur de faire un choix éclairé et le plus adapté à son objectif. Ainsi, ce rapport n'a pas vocation à apporter une réponse unique, mais bien à inciter les utilisateurs à se poser les bonnes questions.

Si le document a été élaboré comme un guide méthodologique à lire en entier, il est également possible de ne consulter que certaines parties.

Il est important également de préciser que la notion de lieux découle des systèmes d'information utilisés : elle est donc étroitement liée à leur conception et à leur modalité d'incorporation dans le SNDS. Il est aussi fortement conseillé de se référer également aux publications du [REDSIAM](#), aux forums de la plateforme d'accès au SNDS et aux forums du HDH entraide pour comprendre les subtilités de l'ensemble du SNDS historique.

En s'appuyant sur l'expérience de l'ensemble de ces membres, ce rapport se veut comme un guide qui vise à fournir des indications pratiques et opérationnelles pour intégrer efficacement la dimension géographique dans les travaux d'analyse.

¹ Certaines informations ne sont disponibles qu'en fonction des profils des utilisateurs et des accès définis pour leur organisme.

Sommaire

Qui sommes-nous ?.....	2
Quelques concepts.....	5
Introduction	7
Sommaire	9
PARTIE 1 - La géographie dans le SNDS.....	12
JE RETIENS.....	13
1. Concepts autour de la géographie dans le SNDS.....	14
1.1 Le lieu de résidence	14
1.2 Le lieu d’affiliation.....	15
1.3 Les lieux de consommation de soins.....	15
1.4 Le lieu de naissance	16
1.5 Le lieu de travail	16
1.6 Le lieu de décès.....	16
2. Les variables et tables du SNDS	16
2.1 Le lieu de résidence du patient.....	18
2.1.1 La dernière situation connue : le référentiel des bénéficiaires (table IR_BEN_R).....	18
2.1.2 Le lieu de résidence au moment du soin	20
2.2.3 La géographie des patients dans les autres tables du DCIR.....	22
2.3 Le lieu d’affiliation.....	23
2.4 Le lieu de décès.....	23
2.5 Le lieu de travail	24
2.6 La géographie des lieux de soins.....	25
2.6.1 Les établissements de santé	26
2.6.2 Les professionnels de santé libéraux	28
2.6.3 Les établissements médico-sociaux	29
PARTIE 2 - La qualité de la géographie.....	30
JE RETIENS.....	31
1. Constats.....	32
1.1 Insee <i>versus</i> consommateurs	32
1.2 Données Assurance maladie et données fiscales	33
2 Quelles atypies ?	33
2.1 Présence de codes postaux.....	34
2.2 Codes communes historiques ou obsolètes	37
2.2.1 Les communes avant 1975.....	37
2.2.2 Les communes obsolètes.....	37
2.2.3 Comment corriger ?.....	38
2.3 Codes imprécis.....	38
2.3.1 Zones spécifiques.....	38
2.3.2 Comment corriger ?.....	38
2.4 Anomalies de facturation.....	39
2.4.1 Erreurs communiquées par la Cnam	39
2.4.2 Comment corriger ?.....	39
2.5 Géographie spécifique, publics spécifiques	40

PARTIE 3 - La géographie des populations spécifiques.....	41
JE RETIENS	42
1. Préambule aux populations spécifiques	43
2. Les femmes enceintes et accouchements	43
2.1 Exemples de travaux reposant sur l'utilisation de distancier	43
2.3 Les accouchements à domicile	44
3. Les enfants	45
4. Les étudiants	47
5. Les personnes en situation de handicap	48
5.1 Exemple d'une étude sur les limitations fonctionnelles	48
6. Les personnes sous-main de justice	50
7. La géographie de l'aide médicale d'état et des soins urgents	51
7.1 La géographie de l'aide médicale d'état	51
7.2 Les soins urgents	52
8. La notion d'« extra-territorialité »	52
8.1 Les collectivités d'outre-mer et la Nouvelle-Calédonie	52
8.2 Les soins réalisés à l'étranger	53
8.3 Les Français résidant à l'étranger	53
8.4 Les frontaliers	53
8.5 Les étrangers séjournant en France	53
9. Les sans domicile fixe.....	54
10. La population protégée par la caisse nationale militaire de sécurité sociale	54
11. Les professionnels de santé.....	55
PARTIE 4 - Quelle géographie pour les indicateurs territorialisés?	58
JE RETIENS	59
1 Quelles études, pour quoi faire ?	60
2 Comment choisir ?	61
2.1 Données récentes	61
2.2 Historique long.....	61
2.3 Données hospitalières.....	62
2.4 Données de décès.....	63
2.5 Arbre de décision permettant de choisir la géographe	63
3 La géographie entre ville et hôpital : comment la traiter ?	64
3.1 Le passage du code géographique au code commune	64
3.2 PMSI ou DCIR quel choix ? l'exemple d'une étude	66
3.2.1 Méthode	66
3.2.2 Résultats	66
3.2.3 Limites	68
3.2.4 Conclusions	68
3.3 Le suivi des localisations des hospitalisations en tant que lieu de soins	69
Conclusion	71
Annexes.....	73

PARTIE 1



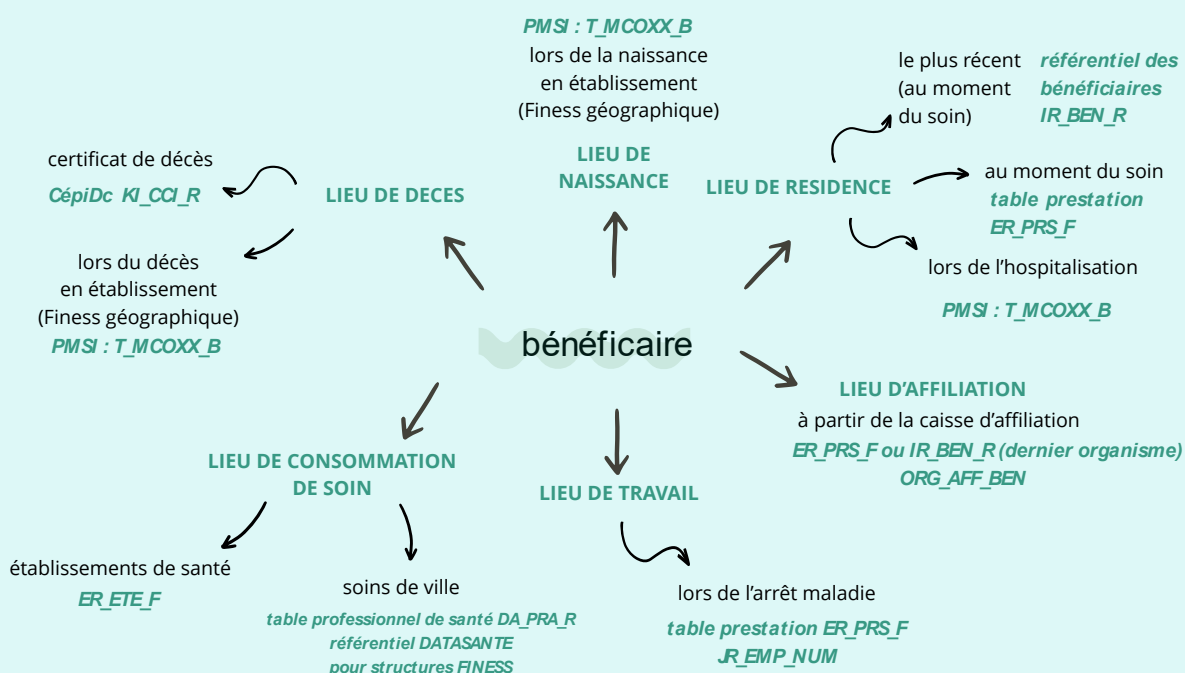
La géographie dans le SNDS

JE RETIENS

Les différentes notions de géographie présentes dans le SNDS

Le SNDS contient de nombreux éléments géographiques aussi bien au niveau des bénéficiaires des soins, du professionnel de santé, que des établissements de santé. Différents espaces géographiques sont ainsi retrouvés, tels que le lieu de résidence du patient, le lieu de travail, le lieu de naissance, le lieu de décès, le lieu de l'hospitalisation, etc.

Le tableau de synthèse ci-dessous recense l'ensemble des éléments géographiques disponibles sur le bénéficiaire ; chacun d'entre eux fait l'objet d'une section spécifique dans laquelle le concept est détaillé.



Ces lieux sont décrits par des codes qu'ils soient communaux, postaux ou départementaux.

Pour les bénéficiaires de soins, la géographie est d'abord centrée sur celle de **l'assuré (ouvreur de droit)**². La commune au moment des soins est décrite par le code officiel géographique de l'Institut national de la statistique et des études économiques (Insee). Il s'agit de l'information inscrite sur la carte vitale du bénéficiaire au moment où celui-ci l'utilise.

Des informations géographiques sont également retrouvées dans le programme de médicalisation des systèmes d'information (PMSI) lors d'une venue à l'hôpital. La géographie hospitalière n'est cependant pas construite de la même façon et repose sur le code postal, sachant qu'un code postal est le plus souvent rattaché à plusieurs communes et que quelques communes ont plusieurs codes postaux (sans parler des Cedex).

La géographie peut varier dans le temps, il est possible de récupérer l'information selon différents niveaux de temporalité (par exemple, ne pas avoir le code commune actuel, mais le code d'une commune ayant existé à un instant donné).

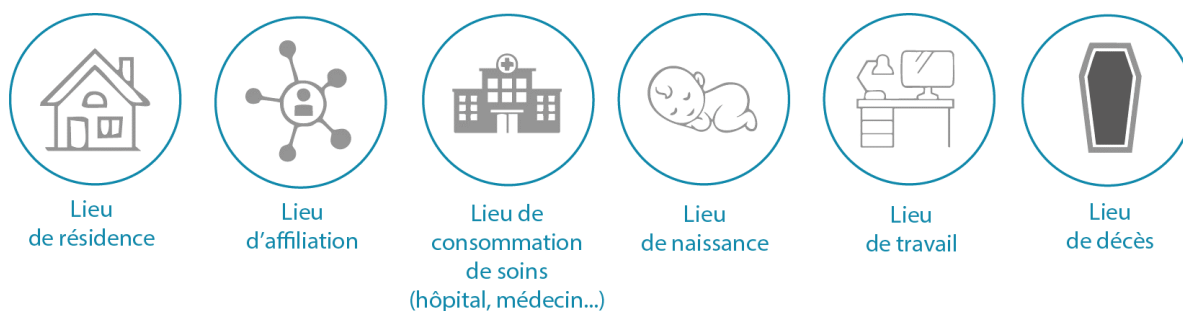
² Pour plus d'information, se référer à la partie « Quelques concepts » en début du document.

1. Concepts autour de la géographie dans le SNDS

La dimension géographique du SNDS regroupe l'ensemble des éléments géographiques disponibles dans les bases de données.

Le schéma ci-dessous reprend les six types d'informations géographiques présentes dans le SNDS.

L'objectif est de présenter la spécificité de chaque dimension et ses limites.



Il est à noter qu'à ces lieux, s'ajoute le lieu de liquidation. Il correspond au lieu de l'organisme qui traite les remboursements des soins pour un bénéficiaire à un moment donné. Cet organisme de liquidation est chargé de gérer les remboursements des bénéficiaires qui lui sont affiliés, mais peut aussi effectuer cette liquidation pour des bénéficiaires affiliés à d'autres organismes.

1.1 Le lieu de résidence

Le lieu de résidence dans le SNDS, reflète principalement la situation administrative de **la personne assurée**. L'assuré est la personne **ouvrant des droits à l'Assurance maladie**. Les bénéficiaires des soins peuvent être l'assuré lui-même, mais également les éventuels ayants droit présents sur le dossier de l'assuré (ouvrant droit) qui ne sont pas eux-mêmes assurés à titre personnel (des enfants par exemple).

L'assuré fournit à son organisme d'affiliation une adresse postale à partir de laquelle un **code commune Insee** est remonté dans les bases locales puis dans le SNDS. Si l'assuré ne fait pas de mise à jour, l'information inscrite sera la dernière connue par son organisme d'affiliation.

Pour les résidents en structure médico-sociale (EHPAD, etc.), la commune de résidence est généralement celle de l'établissement médico-social (ESMS). Néanmoins, pour les structures ou l'hébergement n'est pas à temps complet ou pour des situations particulières, l'adresse de résidence peut être encore celle de l'ancien lieu de vie.

À noter que l'Assurance maladie a besoin d'une adresse, même si les échanges avec l'assuré ou ses ayants droit sont de plus en plus dématérialisés. Si la personne n'a pas d'adresse (sans-abri, résidence mobile, hébergement temporaire, hébergement d'urgence), elle peut faire une demande de **domiciliation** au Centre Communal d'Action Sociale (CCAS) de la commune (ou autres organismes habilités). Elle aura ainsi une adresse, correspondant à la boîte postale de la commune en question.

Certaines communes peuvent donc se voir affecter des individus qui ne vivent pas réellement sur la commune. Cela peut constituer **des points d'ancrage** (ce point sera détaillé dans la partie 2 du rapport).

Le lieu de résidence du bénéficiaire est également saisi dans les bases de données du PMSI. Ce lieu de résidence est parfois différent du lieu déclaré et véhiculé par la caisse d'affiliation à l'Assurance maladie.

1.2 Le lieu d'affiliation

L'affiliation de l'individu à un organisme d'Assurance maladie obligatoire dépend en premier de la situation **professionnelle puis de la situation géographique** de ce dernier.

- Pour les salariés, travailleurs indépendants ou profession libérale, le rattachement se fait à la Caisse Primaire d'Assurance Maladie (Cpam) du département de résidence habituelle. Il s'agit du lieu où la personne déclare résider le plus longtemps durant l'année (plus de 6 mois).
- Pour les personnes sans emploi ni activité, l'affiliation se fait *via* la caisse du département de résidence sur « critère de résidence ».

En conclusion, **pour le régime général, lieu d'affiliation = lieu de résidence.**

- Pour les salariés ou exploitants agricoles, le rattachement se fait à la Mutualité Sociale Agricole (MSA).
 - **Pour les assurés de la MSA, le département de résidence peut différer du département d'affiliation.** En effet, un assuré du régime agricole va être affilié à la caisse MSA du siège social de son employeur. Par exemple, un salarié qui habite dans la Somme (département 80) et travaillant dans une coopérative agricole dans le Val-d'Oise (département 95) sera affilié à la caisse MSA d'Île-de-France (département 75).
 - Pour les salariés et non-salariés agricoles, le rattachement à un organisme d'Assurance maladie obligatoire est fonction de leur lieu d'activité et non de leur lieu de résidence.

En pratique, il peut donc exister un biais entre le département d'affiliation et le lieu réel de résidence des assurés agricoles.

- Enfin certaines professions sont rattachées à des caisses spécifiques (exemples : Caisse nationale militaire de sécurité sociale CNMSS, ENIM etc.) dont les compétences sont nationales. Les bénéficiaires y sont donc affiliés sans critère de résidence, cela signifie que les organismes en question gèrent des bénéficiaires de toutes les régions de France.

1.3 Les lieux de consommation de soins

Dans les bases de l'Assurance maladie, plusieurs types de lieux de soins permettent d'analyser la géographie du parcours de soins des bénéficiaires *via* le lieu d'exercice des professionnels consultés.

- Les établissements de santé (hôpitaux publics, cliniques privées, centres hospitaliers universitaires – CHU –, etc.), identifiés *via* le code Finess géographique³.
- Les structures de soins de premier recours, comme les cabinets de médecins généralistes et spécialistes, les maisons de santé pluriprofessionnelles (MSP), identifiés grâce aux numéros de facturations des professionnels de santé. Il existe une table spécifique qui permet de récupérer de manière cryptée ou non l'information⁴.
- Les pharmacies permettant d'analyser la délivrance des médicaments et les pratiques officinales.
- Les établissements médico-sociaux, comme les EHPAD ou d'autres ESMS, sont également identifiés dans la base RESID-ESMS, dont il est possible de retrouver des tables spécifiques dans le SNDS. Tous les soins ne sont pas identifiables et dépendent des règles de facturation.

³ Le code Finess géographique correspond à une implantation géographique (site réel où se déroulent les activités) ; il est obligatoirement relié à une entité juridique (code Finess juridique). Dans les tables du PMSI, deux types de numéro Finess sont trouvés : Finess PMSI et Finess géographique (ou Finess établissement). Le Finess PMSI (utilisé par l'établissement pour la transmission de ses données) correspond au Finess juridique pour les établissements publics et au Finess géographique pour les établissements privés ; pour les établissements psychiatriques privés qui ont un grand nombre de Finess géographiques, des transmissions cumulées sous un seul Finess PMSI sont pour tolérées. Pour en savoir plus sur le numéro Finess dans le SNDS, cf. documentation disponible sur le site du Health Data Hub : [ICI](#) et section « 2.6.1 Les établissements de santé » de ce document.

⁴ En fonction des profils des utilisateurs et des accès définis pour leur organisme. A *minima*, il est possible de récupérer le département du praticien.

1.4 Le lieu de naissance

Il n'y a pas d'information sur le lieu de naissance en tant que tel dans les bases de données du SNDS, hormis dans les bases de données hospitalières (PMSI). En effet, pour les naissances qui ont lieu en établissement, dans le PMSI, le lieu de naissance peut être approché à partir de la localisation de l'établissement (Finess géographique) qui a pratiqué l'accouchement pour les naissances à l'hôpital (voir partie population spécifique). Le guide méthodologique MCO ([ATIH](#)) précise que le « *code postal du lieu de résidence d'un nouveau-né est celui de la résidence de sa mère* ».

1.5 Le lieu de travail

Les informations sur le lieu de travail ne sont pas directement disponibles dans le SNDS. Néanmoins, dans certains contextes, il est possible d'avoir des informations sur ce lieu au niveau des prestations et indemnités journalières. En effet, ceci permet d'identifier les arrêts de travail liés à une activité professionnelle. Lorsqu'un arrêt de travail est enregistré, le numéro Siren de l'entreprise où l'assuré est employé est indiqué, permettant ainsi d'identifier l'employeur et donc de récupérer la géographie du lieu de travail. Il est toutefois à préciser que le numéro de Siren renvoie à l'adresse du siège social de l'entreprise et non au lieu de travail réel du salarié (exemple du cas où l'entreprise a différents lieux d'implantation ou encore du cas où le salarié occupe un poste où il est fortement mobile).

1.6 Le lieu de décès

Le Centre d'épidémiologie sur les causes médicales de décès (CépiDc) collecte et analyse les données relatives aux lieux de décès en France⁵. Ces informations sont cruciales pour comprendre les tendances de mortalité et les disparités géographiques en matière de santé. Le lieu de décès, tel que documenté par le CépiDc, inclut des détails précis sur la commune et le type de lieu où le décès est survenu : au logement ou à domicile, en établissement hospitalier, en hospice/maison de retraite, en clinique privée, sur la voie publique ou dans un autre lieu. Cette base alimente le SNDS.

Par ailleurs, lorsque le décès a lieu dans un établissement de santé remplissant le PMSI, le mode de sortie « décès » permet de connaître le lieu de décès avec l'identification de l'établissement (Finess géographique) ayant déclaré le décès.

Dans la table des établissements médico-sociaux, le mode de sortie « 9 » indique un décès, la localisation de l'établissement permet d'identifier le lieu de décès, en s'assurant qu'il n'y a pas de double compte avec la table des hospitalisations.

2. Les variables et tables du SNDS

Le tableau présenté à la page ci-dessous reprend l'ensemble des variables géographiques disponibles dans le SNDS. Il propose une vue d'ensemble structurée des variables géographiques disponibles ; elles sont réparties **selon quatre axes principaux : le moment de la géographie, la spécificité d'accès, les jointures entre données géographiques et identifiants bénéficiaires**, ainsi que la documentation associée. Il distingue différents moments de la géographie, tels que la localisation au moment du décès, de l'arrêt de travail ou des soins, avec des mises à jour annuelles ou mensuelles selon les cas. Concernant la spécificité d'accès, certaines données sont disponibles par défaut, tandis que d'autres nécessitent des profils d'accès spécifiques ou sont soumises à des restrictions (comme les tables du CépiDc pour les décès). Le tableau explicite également les mécanismes de jointures entre les tables et détaille les variables géographiques d'intérêt ; par exemple les codes Insee des communes, les codes postaux ou les identifiants Finess et Siret. Enfin, il renvoie à une documentation spécifique qui explicite les bases disponibles et leur fréquence de mise à jour.

Ce document constitue un support méthodologique très utile pour les utilisateurs, en leur fournissant des indications sur les tables et variables disponibles, ainsi que sur les contraintes d'accès. Il facilite l'accès aux données en expliquant comment les extraire ou les associer à d'autres informations.

⁵ Il collecte également l'information sur le lieu de domiciliation de la personne décédée.

BÉNÉFICIAIRE

Geographie de qui ?											
Geographie d'intérêt	Lieu du décès			Lieu de travail		Lieu de résidence			Lieu de d'affiliation		
Système d'information	CépiDc	PMSI		DCIR/ESND	DCIR/ESND		PMSI	CépiDc	DCIR/ESND		
Moment de la Géographie	Au moment du décès			Au moment de l'arrêt de travail	Au dernier soin de l'année		Au moment du soins	Au moment du soins	Au moment du décès	Dernière information connue / Au moment du soin	
Tables et variables Geo	- Table : table des circonstances et des causes initiales de décès ORAVUE.KI_CCL_R - Variable : DCD_DPT_COD (Département de décès du bénéficiaire) - Variable : DCD_COM_COD (Commune de décès du bénéficiaire)	- Table : Référentiel établissement de la Cnam = ORAVUE.BE_IDE_R - Variable : IDE_BDI_COD (code postal de l'établissement)	- Table : Atlas anté RFCOMMUN.DATASANTE_T_FINESS - Variable : COM_CODE (code commune Insee)	- Tables : bases Siren des entreprises = Rfcommun.Stocketablissement et Rfcommun.Stockunitelégale - variables : codeCommuneEtablissement (code Insee), codePostalEtablissement	- Tables : Tables de la cartographie des pathologies ORAMEPS.CRTO_CT_IND_Gxx_aaaa / ORAMEPS.CRTO_CT_RE_S_Gxx_aaaa et ORAMEPS.ESND_CT_IND_Gxx_aaaa / ORAMEPS.ESND_CT_RE_S_Gxx_aaaa - Variables : DPT / CODE_COMMUNE_INSEE , CODE_DEPARTEMENT_INSEE	- Tables : tables annuelles des consommants Consopat.Extractio n_PatientsAAAATR - Variables : dpt, depcom et depcom2 (code commune INSEE)	- Tables : Table des prestations (Flux) ORAVUE.ER_PRS_F et ORAMEPS.ESND_ER_PRS_F - Variables : BEN_RES_DPT, BEN_RES_COM (code commune INSEE)	- Tables : Tables d'hospitalisation : ORAVUE.T_MCOaaB, T_MCOaaFA, T_SSRaaB, T_HADaaB, T_RIPaaB et tables des ACE : T_MCOaaFASTC, T_SSRaaFASTC (ou ORAMEPS.ESND_T_xxxaaxx) - Variables : BDI_DEP , BDI_COD ou BDI_COM, POST_DEP, COD_POST	- Tables : Tables d'hospitalisation : ORAVUE.T_MCOaaB, T_MCOaaFA, T_SSRaaB, T_HADaaB, T_RIPaaB et tables des ACE : T_MCOaaFASTC, T_SSRaaFASTC (ou ORAMEPS.ESND_T_xxxaaxx) - Variables : BDI_DEP , BDI_COD ou BDI_COM, POST_DEP, COD_POST	- Table : Table des circonstances et des causes initiales de décès ORAVUE.KI_CCL_R - Variables : BEN_RES_DPT, BEN_RES_COM (commune INSEE)	- Table : table de valeurs ORAVAL_IR_ORG_V - Variables : GEO_DPT_COD / GEO_COM_COD (dans IR_ORG_V)
Jointure "géographie & identifiants bénéficiaire"	Pas de jointure	Jointure entre la variable IDE_ETA_NUM (table BE_IDE_R) et les variables ETA_NUM ou ETA_NUM_EPMSI ou ETA_NUM_GEO des tables des hospitalisations (PMSI) ORAVUE	Jointure entre la variable FINESS de la table DATASANTE_T_FINESS chaînée avec les variables ETA_NUM / ETA_NUM_EPMSI / ETA_NUM_GEO des tables du PMSI	Jointure entre la variable SIRET des bases Rfcommun.Stocketablissement et Rfcommun.Stockunitelégale et la variable JUR_EMP_NUM de la table des prestations (Flux) ORAVUE.ER_PRS_F ou ORAMEPS.ESND_ER_PRS_F	Pas de jointure (identifiant bénéficiaires et variable géographiques sont dans la même table)					Jointure entre la variable ORG_NUM (table ORAVAL_IR_ORG_V) et la variable ORG_AFF_BEN (tables ORAVUE.IR_BEN_R ou ORAMEPS.ESND_IR_BEN_R)	Jointure entre la variable ORG_NUM (table ORAVAL_IR_ORG_V) la variable ORG_AFF_BEN de la table des prestations (flux) ORAVUE.ER_PRS_F ORAMEPS.ESND_ER_PRS_F
Spécificité d'accès	- ESND : CépiDc non disponible	Table BERF : Non disponible par défaut en accès DCIR sur projet, peut l'être sur demande + la commune de décès est une donnée sensible		RFCOMMUN accessible pour tous les types d'accès dans le portail SNDS	- Accès ESND : tables non disponibles		Accès DCIR/ESND sur projet : uniquement les tables annuelles consolidées disponibles	ESND : table CépiDc non disponible	Accès DCIR/ESND sur projet : uniquement les tables annuelles consolidées disponibles	ESND : table CépiDc non disponible	Accessible pour tous les types d'accès
MAJ	Annuelle	Mensuelle	Bimestrielle	Annuelle	Mensuelle	Annuelle	Annuelle	Mensuelle	Annuelle & mensuelle	Annuelle	Mensuelle
Documentation	HDH		HDH	Documentation BO : _Documentation SNIIRAM-SNDS/DCIR/Documentation SNDS - Tables DATASANTE dans RFCOMMUN Mise à jour -	Portail SNDS/BO: Dossiers publics/SNIIRAM/_Documentation SNDS/DCIR/Documentation Référentiels/	Cartographie des pathologies Portail SNDS/BO : Dossiers publics/SNIIRAM/_Documentation SNIIRAM-	Portail SNDS/BO : Dossiers publics/SNIIRAM/_Documentation SNIIRAM-SNDS/DCIR/_TABLES ANNUELLES DE COM	HDH	HDH	HDH	

2.1 Le lieu de résidence du patient

2.1.1 La dernière situation connue : le référentiel des bénéficiaires

(table IR_BEN_R)

Le référentiel des bénéficiaires IR_BEN_R est une base qui contient les caractéristiques **administratives** de tous les bénéficiaires de l'Assurance maladie qui ont fait l'objet d'un remboursement (ville, cliniques privées ou encore actes et consultations externes des établissements publics). Elle exclut de fait les bénéficiaires qui ont reçu des soins uniquement à l'hôpital public, car le PMSI n'alimente pas le référentiel. Les bénéficiaires des soins en ESMS qui sont payés en budget global en sont exclus également (forfaits de soins).

Il est alimenté par tous les régimes d'Assurance maladie obligatoire. La mise à jour du référentiel des bénéficiaires dépend des régimes et dépend du processus d'échanges de données avec la Cnam : la plupart des régimes utilisent la « Norme d'Échanges Commune » (NEC), c'est le cas aujourd'hui du régime général, des Mines, des Marins, Cultes, Clercs de notaire et Militaires tandis que d'autres s'appuient sur un cahier des charges validé avec la Cnam : Norme technique d'échange inter-régime (NTEIR).

Depuis 2015, le référentiel des bénéficiaires est également enrichi par les informations issues des Bases de Données Opérantes (BDO) des caisses du régime général (y compris Section Locale Mutualiste – SLM –) et celles des régimes utilisant la norme NEC. Cela permet depuis cette date de remonter les non-consommateurs présents dans ces bases locales, mais aussi de récupérer les informations relatives à la date de décès du bénéficiaire et son NIR individuel pseudonymisé.

De manière globale, le référentiel est mis à jour tous les mois.

Depuis 2018, figurent les indicateurs du Répertoire National inter-régime de l'Assurance Maladie (RNIAM) pour les régimes utilisant la norme d'échanges commune (NEC ; voir partie 3).

La table IR_BEN_R est alimentée depuis 2013 ; la table IR_BEN_R_ARC correspond à l'archivage des données du référentiel de 2006 à 2012.

La notion de géographie dans IR_BEN_R

La géographie dans IR_BEN_R correspond à la **dernière situation connue** pour un même pseudo NIR (BEN_NIR_PSA). Un bénéficiaire peut ainsi avoir plusieurs pseudo NIR et donc potentiellement plusieurs géographies ; cf. illustration figurant à la page suivante à partir des enfants.

Ce schéma réalisé dans le MOOC de formation au SNDS par le *Health Data Hub* (HDH) permet de comprendre qu'un bénéficiaire peut avoir plusieurs BEN_NIR_PSA au cours de sa vie.


Sur la première ligne de l'illustration, l'enfant à gauche est l'ayant droit de sa mère (ouvreur de droit), son premier pseudo NIR est construit à partir du NIR de la mère.

Quand il devient son propre assuré (deuxième ligne de l'illustration), il a son propre NIR et donc un autre pseudo NIR.



MOOC de formation au SNDS réalisé par le HDH :
<https://health-data-hub.fr/formations/mooc-snds-001>

	Ayant droit	
NIR	Date de naissance	Sexe
2781259134086	12/06/2015	2

	Assurée (Ouvreur de droit)	
NIR	Date de naissance	Sexe
2781259134086	25/12/1978	2

	Assurée	
NIR	Date de naissance	Sexe
2150659198033	12/06/2015	2

	Assurée	
NIR	Date de naissance	Sexe
2781259134086	25/12/1978	2

Dans IR_BEN_R, c'est la dernière situation connue qui est récupérée pour chaque pseudo NIR (BEN_NIR_PSA). Par dernière situation connue, on entend celle connue à la date la plus récente (actuellement) par l'administration, c'est-à-dire celle du destinataire de règlement (celle de l'assuré = ouvreur de droit, qui correspond à celui qui reçoit le paiement). Si l'assuré ne déclare pas de changement de géographie, la dernière connue par l'administration sera donc en décalage avec la géographie réelle du bénéficiaire.

Les éléments géographiques récupérés sont les suivants :

- BEN_RES_DPT correspond au dernier département connu pour le bénéficiaire. Il est codé sur trois caractères (exemple : 014 pour le département du Calvados etc.).
- BEN_RES_COM correspond à la dernière commune connue pour le bénéficiaire.

L'identification des individus dans IR_BEN_R

Le référentiel est centré, comme évoqué précédemment, sur le pseudo NIR du bénéficiaire (BEN_NIR_PSA) qui correspond au NIR de l'ouvreur de droit couplé avec la date de naissance du bénéficiaire du soin et du sexe du bénéficiaire. **Un même individu peut donc être identifié avec plusieurs pseudo-NIR et donc potentiellement avoir plusieurs géographies différentes.**

Variable	Description
ASS_NIR_ANO	NIR pseudonymisé de l'assuré ouvreur de droit
BEN_NIR_PSA	Pseudo-NIR construit à partir du NIR de l'assuré, date de naissance et sexe du bénéficiaire
BEN_RNG_GEM	Rang gémellaire permettant de distinguer les jumeaux de même sexe
BEN_NIR_ANO	NIR pseudonymisé certifié du bénéficiaire (identifiant unique du bénéficiaire)
BEN_IDT_ANO	Identifiant global d'un bénéficiaire : BEN_NIR_ANO si existant sinon BEN_NIR_PSA + BEN_RNG_GEM
BEN_IDT_TOP	Indicateur = 1 si BEN_IDT_ANO = BEN_NIR_ANO (identifiant fiable), = 0 sinon

Le BEN_IDT_ANO permet de regrouper les différents pseudo-NIR pour un même bénéficiaire. Il est toutefois recommandé lorsque l'on travaille avec le BEN_IDT_ANO (il y a potentiellement plusieurs BEN_NIR_PSA pour un même BEN_IDT_ANO) de prendre l'information sur la ligne actualisée la plus récemment pour l'individu avec BEN_DTE_MAJ (date de mise à jour).



Pour aller plus loin :

[Identifiants des bénéficiaires | Documentation du SNDS & SNDS OMOP](#)

2.1.2 Le lieu de résidence au moment du soin

Les éléments de géographie sont retrouvés au niveau des tables de l'Assurance maladie et du PMSI. La géographie est très différente et fera l'objet d'un traitement dans deux sessions différentes de ce rapport.

Le lieu de résidence dans les tables des prestations

La géographie des bénéficiaires figure dans la table principale de prestations du DCIR : **ER_PRS_F**.

Elle peut être entendue comme la géographie qui figure sur la carte vitale au moment du soin. Une fois de plus, cette géographie s'appuie sur les éléments administratifs connus de la personne à la date du soin et cela indépendamment du lieu physique où se trouve la personne. Par exemple un assuré qui part en vacances en Corse aura un code géographique Insee qui correspond à celui de son lieu de résidence et non à celui de la commune de la Corse où il séjourne ponctuellement.

La géolocalisation du bénéficiaire s'appuie sur deux variables dans la table des prestations ER_PRS_F :

- **BEN_RES_DPT**, désignant le département du lieu de résidence, connu par l'organisme d'affiliation. Il s'exprime sur trois caractères : « 20X » pour la Corse⁶, parfois « 97X » pour les départements d'outre-mer⁷, « 99X » pour les données manquantes ou personnes à l'étranger, et « 0XX » pour la plupart des départements, sur deux caractères.
- **BEN_RES_COM**, codé sur trois caractères, qui désigne la commune à l'intérieur du département ;

Ces deux variables combinées permettent d'identifier la commune de résidence. En outre, elles permettent de former le code Insee complet de la commune (codé sur cinq caractères).

La table ER_PRS_F contient deux variables : DRG_NAT (destinataire du règlement) et DRG_NAT_AFF (destinataire du règlement affiné) qui ne sont disponibles que pour les régimes utilisant la norme NEC pour la variable DRG_NAT et uniquement le régime général pour la variable DRG_NAT_AFF.

L'association de la variable DRG_NAT en particulier en l'absence de tiers payant et de certaines valeurs de la variable DRG_NAT_AFF permet d'obtenir des informations pour le choix de la géographie (cf. ANNEXE 1).

Le lieu de résidence dans les tables du PMSI

La géographie du lieu de résidence dans les tables issues du PMSI repose sur le code postal.

Dans la majorité des cas le code géographique PMSI correspond au code postal du lieu de résidence du bénéficiaire , et dans quelques cas, il s'agit d'un groupement de codes postaux.

Au sein des établissements de santé, la donnée initiale recueillie pour déterminer le lieu de résidence du patient est le **code postal de résidence du patient**. Cette information est commune à tous les champs du PMSI, qu'il s'agisse d'une hospitalisation ou de l'activité externe des établissements publics.



Les consignes de recueil du code postal sont définies dans les guides méthodologiques, mis à jour annuellement, [MCO](#), [HAD](#), [SMR](#) et [Psychiatrie](#).⁸

⁶ Souvent 209, et quelques 201 et 202.

⁷ 971,972... mais cette codification contredit parfois le premier caractère de BEN_RES_COM ; il y a beaucoup de « 097 ».

⁸ Guide méthodologique [guide_methodo_mco_2024_version_provisoire_1.pdf](#).

En raison de contraintes de confidentialité (Cnil), le code postal n'a pas toujours été restitué dans les données pseudonymisées du SNDS.

Avant 2020

- Si un code postal compte *au moins* 1 000 habitants (données Insee), alors le code postal est conservé et le code postal correspond au code géographique PMSI.
- Si un code postal compte *moins* de 1 000 habitants (données Insee), alors il est fusionné avec un autre code postal limitrophe, pour former un **code géographique PMSI**.

Voici un exemple dans le département du Rhône :

Code géographique PMSI 2022	Code postal 2022	Population Insee
69C01	69115 Chiroubles	412
	69910 Villié-Morgon	2 048
Total 69C01		2 461

Le cas des DOM-TOM

Afin d'identifier directement les DOM-TOM à partir des deux premiers caractères du code géographique, comme pour les autres départements français, leurs codes postaux sont systématiquement transcodés dans un code géographique PMSI, quelle que soit leur population couverte :

DOM-TOM	Code postal	Code géographique PMSI
Guadeloupe	97 1XX	9A 0XX
Martinique	97 2XX	9B 0XX
Guyane	97 3XX	9C 0XX
Réunion	97 4XX	9D 0XX
Saint-Pierre-et-Miquelon	97 5XX	9E 000 (ou 9E 999)
Mayotte	97 6XX	9F 0XX
Wallis et Futuna	98 6XX	9G 0XX
Polynésie française	98 7XX	9H 0XX
Tahiti	98 7XX	9H 0XX
Nouvelle-Calédonie	98 8XX	9J 0XX
Saint-Barthélemy	97 133 (département 977)	9A033
Saint-Martin	97 150 (département 978)	9A050

Les communes multi-distribuées

Beaucoup d'études cherchent à rapporter les données PMSI à la population domiciliée. Or, les données de population sont publiées par l'Insee par code commune, et ces deux découpages géographiques (code commune / code postal) ne sont pas systématiquement superposables, avec le cas des communes multi-distribuées (une commune = plusieurs codes postaux).

Des codes géographiques peuvent donc être également créés en fusionnant au moins deux codes postaux, dans le but de pouvoir déterminer la population couverte.

Voici un exemple : en 2019, la commune avec le code Insee 61474 *Silly-en-Gouffem* est multi-distribuée sur les codes postaux 61160 et 61310.

Pour pouvoir passer d'un code géographique PMSI à un code commune Insee, des précautions sont donc nécessaires.

Certains lieux de résidence sont développés dans la partie 3 portant sur les populations spécifiques.

2.2.3 La géographie des patients dans les autres tables du DCIR

La cartographie des pathologies

La Cartographie des pathologies et des dépenses⁹ est une base de données créée par la Cnam. Elle est disponible en *Open data*¹⁰ et *via* le SNDS pour les utilisateurs ayant un accès aux données individuelles. La cartographie a pour but de faciliter les études statistiques concernant la prévalence de certaines pathologies et leurs coûts financiers pour le système de santé. Elle couvre une soixantaine de pathologies, traitements chroniques et épisodes de soins repérables dans le SNDS à partir d'algorithmes développés sur le DCIR et le PMSI. La version 12 (G12) de cette base de données est disponible en inter-régime pour la période de 2015 à 2023.

Les données restituées dans le portail du SNDS sont au niveau le plus fin : le bénéficiaire. Chaque bénéficiaire est identifié par un identifiant unique. Ce bénéficiaire est localisé au niveau de sa commune de résidence *via* le code Insee. Cette information est restituée dans la librairie **ORAMEPS** dans la table **ZZZZ_CT_RES_GXX_aaaa** (ZZZZ est soit CRTO ou ENSD, XX correspond à la version de la cartographie et aaaa à l'année). Ces informations couplées aux caractéristiques¹¹ de l'individu présentes dans la table **ZZZZ_CT_IND_GXX_aaaa**, en fait un référentiel de bénéficiaire avec un historique annuel.

L'information de la commune de résidence du bénéficiaire est produite par un algorithme qui alimente des tables mensuelles de correction. L'accès à ces tables n'est pas ouvert à l'ensemble des utilisateurs¹², mais la Cartographie des pathologies reprend ces informations dans une synthèse annuelle.

Le département de résidence est celui de la dernière date de soin de l'année. Cette variable (BEN_RES_DPT) est sélectionnée dans la table des prestations (ER_PRS_F). Lorsqu'un identifiant bénéficiaire (BEN_IDT_ANO) a plusieurs identifiants SNIIRAM (BEN_NIR_PSA||BEN_RNG_GEM), c'est le département de résidence de l'identifiant SNIIRAM ayant la date de mise à jour la plus récente dans le référentiel bénéficiaire (IR_BEN_R) qui est retenue.

Les tables CONSOPAT

On trouve dans le SNDS des tables annuelles de consommateurs. Sont considérés « consommateurs », tous les individus, hors fictifs, qui ont eu au moins une consommation dans l'année dans le DCIR. Les tables se situent dans la librairie **CONSOPAT** :

- extraction_patients2006, extraction_patients2023...
- extraction_patients2006TR, extraction_patients2023TR... pour tous régimes.

Pour les tables tous régimes (TR), on trouve une ligne par individu et régime. Il convient d'être vigilant lors de son usage.

On trouve les variables géographiques suivantes :

- **ben_res_dpt / ben_res_com** : respectivement codes département et commune de résidence du bénéficiaire correspondant **au dernier soin de l'année**,
- **depcom** : code commune obtenu par concaténation de ben_res_dpt et ben_res_com (la plupart des codes erronés/obsolètes en date du 1-1-2015 a été corrigée),
- **depcom2** : recodage de depcom du format 20XXX en 2AXXX ou 2BXXX pour la Corse et du format 97XXX en 9AXXX, 9BXXX, 9CXXX ou 9DXXX pour les DOM (non renseigné ailleurs).

top_insee : '1' si le code commune est une commune renseignée dans la table d'appartenance géographique des communes au 1-1-2015 disponible sur le site de l'Insee, '0' sinon.

⁹ Rachas, Antoine, Christelle Gastaldi-Ménager, Pierre Denis, Pauline Barthélémy, Panayotis Constantinou, Jérôme Drouin, Dimitri Lastier, et al. « *The Economic Burden of Disease in France From the National Health Insurance Perspective : The Healthcare Expenditures and Conditions Mapping Used to Prepare the French Social Security Funding Act and the Public Health Act* ». Medical Care 60, n° 9 (1^{er} septembre 2022) : 655 64. <https://doi.org/10.1097/MLR.0000000000001745>.

¹⁰ <https://data.ameli.fr/pages/data-pathologies/>. Le territoire le plus fin *via* cet accès est le département.

¹¹ Âge, sexe, ALD, régime d'appartenance, C2S...

¹² Profil 30 et sur demande.

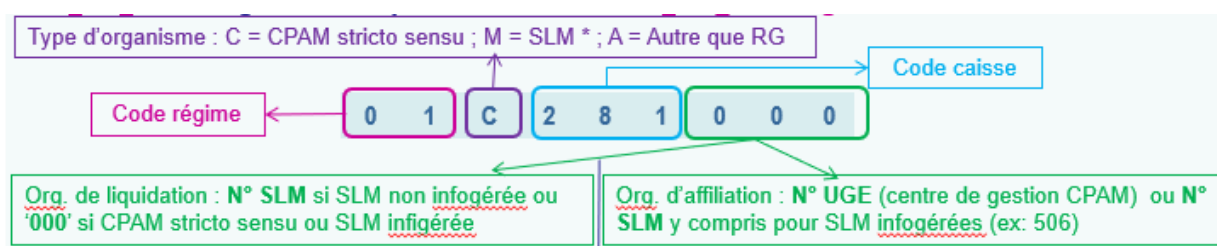
2.3 Le lieu d'affiliation

Le régime d'affiliation est constitué de neuf caractères :

- les deux premiers indiquent le régime (par exemple 01 pour le régime général, 02 pour la MSA...),
- le troisième indique le type d'organisme : Cpam, SLM ou autres caisses,
- les trois suivants correspondent au code caisse (numéro de département + numéro d'ordre ; par exemple 281 correspond à la Cpam d'Eure-et-Loir),
- les trois derniers correspondent au code de la sous-structure locale.

La variable ORG_AFF_BEN restitue l'information de l'organisme d'affiliation. Cette variable est disponible au niveau de la table des prestations (ER_PRS_F) et au niveau du référentiel des bénéficiaires (IR_BEN_R).

Elle est constituée comme suit :



Extrait du cours DIB, Cnam Janvier 2025

Pour le régime général, on peut connaître le département de résidence à partir du code caisse (caractère 4 et 5 de l'organisme d'affiliation).

2.4 Le lieu de décès

Des informations sur le lieu de décès sont disponibles à partir de plusieurs sources.

Dans les tables du PMSI

Lorsqu'un décès a lieu dans un établissement, le mode de sortie est décès (SOR_MOD='9'). Il est donc possible en utilisant le Finess géographique de l'établissement de déterminer le lieu de décès. On utilise la variable SOR_MOD pour tous les champs, dans les tables B, S ou RSA, pour cibler les séjours et la variable ETA_NUM et ses extensions pour identifier le lieu.

Dans les tables du CépiDc

Le lieu de décès peut être connu dans la table KI_CCI_R en utilisant les deux variables département du décès (DCD_DPT_COD) et commune de décès (DCD_COM_COD).

Les informations sur la nature du lieu de décès (DCD_LIE_COD) sont explicitées dans la table de valeur DS_LIE_V et permettent d'apporter les précisions suivantes :

- 0 = Non renseigné
- 1 = Domicile
- 2 = Établissement public de santé
- 3 = Établissement privé de santé
- 4 = EHPAD, maison de retraite
- 5 = Voie publique
- 6 = Autre lieu ou indéterminé
- 7 = Établissement pénitentiaire (certificat de 2017 uniquement)

La codification du lieu de décès en cas de mort violente (LIE_LMV_COD) apporte également une information qualitative. La table de valeur est DS_LMV_V et les valeurs sont les suivantes :

- 1 = Autre lieu ou indéterminé
- 2 = Domicile
- 3 = Établissement accueillant du public
- 4 = Exploitation agricole
- 5 = Lieu de sport
- 6 = Commerce
- 7 = Voie publique
- 8 = Local industriel - chantier



Une documentation est disponible dans le portail Cnam, ainsi que dans la [documentation HDH](#), la dernière version disponible est la version 6, du 20 novembre 2024. Un fichier Excel est également disponible [au même endroit](#).

2.5 Le lieu de travail

Plusieurs tables permettent d'avoir des précisions sur le lieu de travail à partir du code Siren.

RFCOMMUN.STOCKETABLISSEMENT permet d'obtenir des renseignements qui sont voisins de la base Finess. La clé est le numéro Siren.

Depuis juillet 2024, deux nouvelles tables sont mises à disposition dans la bibliothèque RFCOMMUN, accessibles à tous les utilisateurs du SNIIRAM / SNDS ayant un accès à SAS/Guide. Il s'agit de tables issues des bases SIRENE des entreprises et de leurs établissements (Siren, Siret).

Les deux tables se nomment STOCKETABLISSEMENT et STOCKUNITELEGALE.

- STOCKETABLISSEMENT : table qui représente le fichier stock des établissements (ensemble des établissements actifs et fermés dans leur état courant au répertoire).
- STOCKUNITELEGALE : table qui représente le fichier des valeurs courantes des unités légales.



Elles sont également téléchargeables *via* le lien : <https://www.data.gouv.fr/fr/datasets/base-sirene-des-entreprises-et-de-leurs-etablissements-siren-siret/>
Une description précise des variables de chacune des tables est également présente *via* le lien ci-dessus.

2.6 La géographie des lieux de soins

Le tableau ci-dessous recense l'ensemble des variables géographiques disponibles sur ce sujet.

Géographie de qui ?						
Géographie d'intérêt						
OFFREUR / PRESCRIPTEUR DE SOIN (PS / ÉTABLISSEMENT)						
Lieu du soin / de la prescription						
Système d'information						
Moment de la Géographie						
Tables et variables Géo						
Tables et variables Jointure						
Spécificité d'accès						
MAJ						
Documentation						
Référentiel des bénéficiaires de soins médico-sociaux (IR_ESM_R)	PMSI	DCIR/ESND				
Au moment de la prise en charge du patient	Au moment du soin	Au moment du soin / de la prescription			Au moment du soin ou à une date donnée	
- Tables : Référentiel des bénéficiaires de soins médico-sociaux CONSOPAT.IR_ESM_R et ORAMEPS.ESND_IR_ESM_R - variables : ESM_DPT_COD (Code du département du service ESMS consommé) (pas de code commune)	- Table : Référentiel établissement (BERF) BE_IDE_R - Variables : IDE_BDI_COD (code postal) ainsi que l'adresse de l'établissement dans cette table via les variables IDE_VOL_NUM, IDE_VOL_TYP_LRG, IDE_VOL_LIB, IDE_RSD_LIB	- Table : Atlasanté RFCOMMUN.DATASANTE_T_FINISS - Variable : COM_CODE (code Insee)	- Table : Référentiel établissement (BERF) BE_IDE_R - Variables : IDE_BDI_COD (code postal) ainsi que l'adresse de l'établissement dans cette table via les variables IDE_VOL_NUM, IDE_VOL_TYP_LRG, IDE_VOL_LIB, IDE_RSD_LIB	- Table : Atlasanté RFCOMMUN.DATASANTE_T_FINISS - Variable : COM_CODE (code Insee)	- Table Référentiel des Professionnels de Santé simplifié ORAVUE.DA_PRA_R / ORAMEPS.ESND_DA_PRA_R - Variables : PFS_COD_POS (Code postal adr. prof du P.S.) / P_FS_EXC_COM (Code commune adr. prof du P.S.) / PFS_LIB_COM (Lib. commune adr. prof. du P.S.)	
Pas de jointure (identifiant bénéficiaires et variable géographiques sont dans la même table)	Jointure entre la variable IDE_ETA_NUM (table ORAVUE.BE_IDE_R) et les variables ETA_NUM ou ETA_NUM_EPMSI ou ETA_NUM_GEO des tables des hospitalisations (PMSI) ORAVUE.T_MCOaaB, T_HADaaB, T_HADaaS, T_RIPaaRSA, T_RIPaaS, T_SSRaaB, T_SSRaaS (ou ORAMEPS.ESND_T_xxxx axx)	Jointure entre la variable FINISS de la table RFCOMMUN.DATASANTE_T_FINISS chaînée avec les variables ETA_NUM / ETA_NUM_EPMSI / ETA_NUM_GEO des tables du PMSI	Jointure entre la variable IDE_ETA_NUM8 (table ORAVUE.BE_IDE_R) et la variable ETB_EXE_FIN de la table affinée des établissements ORAVUE.ER_ETE_F (ou ORAMEPS.ESND_ER_ET E_F)	Jointure entre la variable FINISS8 de la table RFCOMMUN.DATASANTE_T_FINISS et la variable ETB_EXE_FIN de la table affinée des établissements ORAVUE.ER_ETE_F (ou ORAMEPS.ESND_ER_ETE _F)	Si jointure entre ORAVUE.DA_PRA_R et la table des prestations ORAVUE.ER_PRIS_F (ou ORAMEPS.ESND_ER_PRIS_F) , clé de jointure principale PFS_PFS_NUM (DA_PRA_R) = PFS_EXE_NUM ou PFS_PRE_NUM (ER_PRIS_F) , mais il faut une jointure supplémentaire entre date de flux FLX_DIS_DTD de ER_PRIS_F et concaténation des variables DTE_ANN_TRRT et DTE_MOI_FIN de DA_PRA_R pour avoir la situation au moment de l'étude car cette table est historisée	
- Accès permanent DCIR /ESND : disponible pour tous les profils (y compris profils régionaux) - Accès DCIR sur projet : l'extraction de cette table se fait sur demande justifiée via accès classique (via CESREES et CNIL)	Table BERF : - Non disponible pour les accès ESND - Non disponible par défaut en accès DCIR sur projet		Table BERF : - Non disponible pour les accès ESND - Non disponible par défaut en accès DCIR sur projet		Table DA_PRA_R simplifiée : - Disponible en accès permanent DCIR / ESND - Non disponible par défaut en accès DCIR sur projet, peut l'être sur demande qui doit être justifiée	
Mensuelle	Mensuelle	Bimestrielle	Mensuelle	Bimestrielle	Mensuelle	
HDH	HDH	Documentation BO : Documentation SNIIRAM-SNDS/.../ Tables DATASANTE HDH	HDH	Documentation BO : Documentation SNIIRAM-SNDS/.../ Tables DATASANTE HDH	HDH	

Chacun des éléments ci-dessus sont détaillés dans la suite de ce rapport.

2.6.1 Les établissements de santé

La base Finess

Dans le PMSI, les établissements transmettent leurs données sous un numéro issu du répertoire Finess (Fichier National des Établissements Sanitaires, Sociaux et médico-sociaux)¹³.

À chaque établissement est attribué un numéro Finess à neuf caractères, dont les deux premiers correspondent au numéro du département d'implantation. Dans le DCIR le Finess de l'établissement est codé sous huit caractères (sans la clé).

Pour tous les départements d'outre-mer, les deux premiers caractères sont « 97 » et le quatrième caractère précise dans ce cas, le département : 1 = Guadeloupe, 2 = Martinique, 3 = Guyane et 4 = Réunion. Pour Saint-Pierre-et-Miquelon, les deux premiers caractères sont 97 et le quatrième caractère est le 5. Pour Mayotte, les deux premiers caractères sont 98.

Dans le PMSI, le Finess est dit « de transmission PMSI » qui correspond au Finess juridique ou géographique selon le statut juridique de l'établissement et sa grille tarifaire.

Le Finess géographique, qui est le niveau le plus fin pour repérer un établissement, est systématiquement disponible dans les quatre champs du PMSI. Il est à noter que cette information fait l'objet d'un contrôle bloquant depuis 2013 pour les champs MCO, HAD et SMR (l'établissement est dans l'impossibilité de transmettre ses données si le Finess géographique est erroné et/ou ne correspond pas aux autorisations attendues). En psychiatrie, l'établissement est alerté si le Finess géographique ne semble pas faire partie de la liste des attendus.

Le lieu de soin, quand il est hospitalier, correspond à l'**adresse de l'établissement de santé**. Cette adresse est reliée au numéro d'immatriculation de l'établissement de santé, son « numéro Finess ». Ces données sont disponibles en libre accès (finess.esante.gouv.fr).

Dans le SNDS, des tables sont mises à disposition par l'équipe du projet AtlaSanté¹⁴ et permettent d'accéder à l'ensemble des informations¹⁵. Ces tables se trouvent dans la librairie RFCOMMUN du SNDS et sont accessibles à l'ensemble des utilisateurs et mises à jour tous les trois mois.

Sont ainsi trouvées les tables :

- **t_finess** : référentiel Finess qui fait correspondre à chaque structure Finess, géographique ou juridique, qui existe ou qui a existé depuis 2004 ses principales caractéristiques : statut juridique, catégorie, adresse, coordonnées dans les projections légales, coordonnées GPS...
- **t_actsoins_finess** : table qui recense – à une date donnée – les activités de soins (médecine, chirurgie, obstétrique, urgences, SSR, oncologie...) autorisées et mises en œuvre au sein des structures Finess.
- **t_eml_finess** : table qui recense – à une date donnée – les équipements et matériels lourds (scanner, IRM, tomographe à émission...) autorisés et mis en œuvre au sein des structures Finess.
- **t_actms_finess** : table qui recense – à une date donnée – les équipements sociaux et médico-sociaux actifs au sein des structures Finess.

¹³ [Finess.esante.gouv.fr](https://finess.esante.gouv.fr)

¹⁴ Le projet AtlaSanté est un projet mutualisé des ARS porté par la direction du numérique du ministère de la santé. Plus d'infos ici : [Accueil - AtlaSanté](#).

¹⁵ Référentiel Finess (t_finess) (DataSanté) - data.gouv.fr.

L'établissement dans le PMSI

De nombreuses variables dans le PMSI identifient l'établissement. Selon le champ et l'année, elles n'ont pas toujours la même dénomination. ETA_NUM désigne le numéro Finess servant à la transmission des données pour le PMSI. Cette variable s'appelle ETA_NUM_EPMSI pour le champ HAD et le Rim-P, depuis 2007, avec le nom de ETA_NUM_JUR en 2006 pour l'HAD.

Depuis 2013, les établissements du champ MCO ou SMR ayant plusieurs sites géographiques peuvent identifier les différents sites géographiques par la variable ETA_NUM_GEO ; cette variable avait deux noms en 2012 (ETA_NUM_GEO1 et ETA_NUM_GEO2). Pour le champ HAD, la variable ETA_NUM_GEO apparaissait en 2006 puis réapparaît depuis 2017 ; son nom était ETA_NUM_TWO entre 2007 et 2016. Elle s'appelle ETA_NUM_TWO pour le Rim-P.

Pour le champ HAD, depuis 2010, il est possible de coder le Finess de l'hébergement médico-social dans lequel la structure de l'HAD intervient. Par convention, la variable du Finess de l'hébergement s'appelait ETA_NUM_EHPA jusqu'en 2017, puis se nomme ETA_NUM_ESMS pour les années suivantes.

Le type de lieu de domicile du patient est codé, depuis 2007, par la variable PAT_TYP_DOM. En 2018 (Guide méthodologique HAD), l'intervention simultanée au domicile du patient de l'HAD et d'un SSIAD (service de soins infirmiers à domicile) ou SPASAD (service polyvalent d'aide et de soins à domicile) est autorisée sous condition. Une nouvelle modalité (6) « Type de lieu de domicile du patient » est créée.

Les codes du type de lieu sont :

Code 1	Domicile personnel du patient
Code 2	Autre domicile
Code 3	Établissement d'hébergement pour personnes âgées dépendantes (EHPAD)
Code 4	Patient hébergé en établissement médicosocial hors EHPAD
Code 5	Patient hébergé en établissement social
Code 6	Domicile personnel du patient ou autre domicile (codes 1, 2 ou 5) avec intervention concomitante d'un SSIAD ou d'un SPASAD. Un changement de type de lieu de domicile impose un changement de sous-séquence. Ainsi, si un patient pris en charge à son domicile personnel (code 1) est transféré et pris en charge dans un EHPAD (code 3), il faut clore le précédent RPSS et en ouvrir un nouveau en EHPAD.

Les modes d'entrée ou de sortie (ENT_MOD, SOR_MOD) permettent d'identifier les transferts définitifs ou provisoires, les mutations, les naissances (*cf.* partie 3), les patients entrés décédés pour prélèvements d'organe ou le domicile. La notion de mutation ou de transfert fait appel aux notions de Finess géographique, de Finess juridique. Le domicile est entendu comme domicile ou son substitut, tel un hébergement médico-social ; il inclut les entrées à partir de la voie publique.

La provenance et la destination sont étroitement liées respectivement au mode d'entrée ou de sortie. La provenance permet d'identifier le passage par une structure des urgences, y compris les unités d'hébergement de courte durée, elle prend en compte la notion d'entité géographique. Des modifications sur la provenance ont eu lieu en 2023 ; un document est disponible pour expliquer les différentes possibilités sur le site de [l'ATIH](#).

La table E issue du PMSI (T_ « champ »aaE) contient un découpage selon un regroupement en fonction des régions administratives avec la nécessité, depuis 2015, de mettre à jour le passage entre les anciennes régions et les nouvelles régions, les noms des régions ne sont pas standardisés et ont changé au cours des années.

Elle comporte également la région « Service de santé des armées ». Cette région est créée pour des raisons de gestion administrative des données du PMSI, de ces établissements. Il est cependant possible de ré-allouer les établissements en fonction de la région administrative en utilisant les deux premiers caractères du Finess. Attention, les données des établissements du service de santé ne sont disponibles que depuis 2009 pour le champ MCO, 2013 pour le SMR et le Rim-P, avec un nom non standardisé.

L'institution nationale des invalides ne transmet pas de données dans le PMSI.

L'identification de Mayotte est disponible depuis 2020 pour le MCO et 2022 pour les autres champs. Le centre hospitalier Francois Dunan de Saint-Pierre-et-Miquelon transmet le PMSI pour le MCO depuis 2016. En 2010, l'AP-HP et l'AP-HM sont considérées comme deux « régions » pour le MCO.

La variable SOC_RAI indique le libellé de l'établissement, il s'agit d'un référentiel fourni par l'ATIH qui n'est pas exactement celui de la base de données Finess.

Depuis 2016, le fichier supplémentaire du SMUR, qui est repris dans la table T_MCOaaSUP_SMUR, contient le Finess du SMUR, donc la localisation du SMUR.

Le fichier des unités médicales en MCO (T_MCOaaSUPUM) permet par exemple d'identifier la localisation des maternités selon leur niveau en utilisant le ETA_NUM_GEO.

Pour les établissements de psychiatrie, le lieu de soins peut être affiné sans que la localisation du lieu de soins en termes de données « GPS » soit toujours possible ; variable ACT_LIEU de T_RIPaaR3A.

Code lieu	Libellé du lieu	Commentaire / précision
L01	Centre médico-psychologique (CMP)	
L02	Lieu de soins psychiatriques de l'établissement	hors CMP, CATTP et urgences psychiatriques
L03	Établissement social ou médicosocial sans hébergement	
L04	Établissement scolaire ou centre de formation	
L05	Protection maternelle et infantile	
L06	Établissement pénitentiaire	
L07	Domicile du patient ou substitut du domicile	Inclut visites à domicile, logement des proches si résidence habituelle
L08	Établissement social ou médicosocial avec hébergement	Inclut EHPAD
L09	Unité d'hospitalisation ou de consultation (MCO, SMR, USLD)	Intervention psy auprès d'un patient pris en charge dans ces unités
L10	Structure des urgences autorisée	Urgences, UHCD, ZSTCD - psychiatrie de liaison
L11	Centre d'accueil thérapeutique à temps partiel (CATTP)	
L12	Unité d'accueil d'urgences psychiatriques	Urgences psychiatriques hors structure d'urgences autorisée
L13	Autres lieux d'accueil ou structures de prise en charge	Centres de santé, cabinets libéraux, MSP, etc.
L14	Autre établissement de psychiatrie	
LNA	Lieu non attendu	

De même, dans les établissements SMR, les actes du CSARR (CSAR) comprennent des modulateurs de lieu permettant d'identifier des lieux de soins qui ne sont pas dans les établissements (lieu de vie, hors établissements, etc.). La prise en charge dans ces lieux est temporaire et ne dépasse pas la journée.

La Variable MODUL_LIEU_COD de la table T_SSRaaCCAR et les modalités suivantes permettent de localiser la prise en charge du patient sans pour autant connaître son emplacement :

- HW : réalisation de l'acte hors établissement en milieu urbain,
- LJ : réalisation de l'acte hors établissement en milieu naturel,
- XH : réalisation de l'acte sur le lieu d'activité du patient.

La table RFCOMMUN.ETA_NUM_AP_GEO_ANC indique les 40 établissements de l'assistance publique de Marseille et Paris.

La table RFCOMMUN.ETA_NUM_AP_VIA_UM permet de relier la variable eta_num_geo de la table UM, disponible dans le PMSI MCO à un libellé, une raison sociale et à l'eta_num_jur.

2.6.2 Les professionnels de santé libéraux

Cette section vise à décrire une méthode permettant de localiser les professionnels de santé (PS) libéraux dans le SNDS à un niveau départemental ou à un niveau communal, à partir du référentiel des professionnels de santé libéraux (DA_PRA_R).

La table **DA_PRA_R** est un référentiel de PS, accessible aux profils d'utilisateurs de données individuelles du SNDS, rassemble des informations sur les professionnels de santé libéraux et leurs lieux d'exercice. Elle est issue du Fichier National des Professionnels de Santé (FNPS), alimenté par les Caisses primaires d'assurance maladie.

La table **DA_PRA_R** est cumulative et alimentée mensuellement à partir **du flux de facturation**. Elle contient une ligne par professionnel, lieu d'exercice et mois d'intégration des données dans le SNDS.

Localisation du lieu d'exercice des professionnels de santé libéraux

Sous réserve de disponibilité relative au statut du profil, les variables de localisation du cabinet contenues dans la table *DA_PRA_R* sont les suivantes :

- *PFS_COD_POS* : le code postal du cabinet,
- *PFS_EXC_COM* : les trois derniers caractères du code commune du cabinet (si accès PS en clair, crypté sinon),
- *PFS_LIB_COM* : le libellé de la commune du cabinet.

Attention : aucune variable ne contient le code département du lieu d'exercice, mais il peut être récupéré à partir de son code postal (*PFS_COD_POS*).

2.6.3 Les établissements médico-sociaux

La table *IR_ESM_R* est un référentiel alimenté à partir de RESID-ESMS.

RESID-ESMS (Résidents des Établissements et Services Médico-Sociaux) est un traitement de données inter-régimes mis en œuvre par la Caisse Nationale de l'Assurance Maladie (Cnam). Il vise à collecter et centraliser des informations sur les personnes accueillies ou accompagnées par les établissements et services médico-sociaux (ESMS), tels que les EHPAD, SSIAD, SPASAD, IME, MAS, FAM, ESAT, etc.

Son objectif est de :

- Suivre l'activité et la consommation de soins des personnes accueillies ou accompagnées par les ESMS ;
- Analyser la globalité des dépenses d'Assurance maladie rattachables à ces personnes, en rapprochant les données relatives aux budgets des établissements et services avec celles relatives aux remboursements de soins de ville et hospitaliers ;
- Suivre les parcours de soins des personnes concernées ;
- Contrôler les facturations présentées au remboursement des organismes d'Assurance maladie, en vue de la récupération des indus ;
- Répartir l'activité globale de soins entre les régimes d'Assurance maladie.

Les données collectées dans le cadre de RESID-ESMS incluent :

- Informations sur les bénéficiaires : identifiants pseudonymisés, dates de début et de fin de prise en charge, motifs d'entrée et de sortie, etc. ;
- Données relatives aux établissements ou services : catégorie d'ESMS, modalités tarifaires (tarif global ou partiel), présence ou non d'une pharmacie à usage intérieur (PUI), etc. ;
- Données sur les prises en charge : dates de début et de fin, motifs de fin de prise en charge, etc.

Il est important de noter que certaines informations, telles que le numéro Finess de l'établissement ou la localisation précise (au niveau communal), ne sont pas restituées afin de préserver la confidentialité des données. Aussi, la localisation des EHPAD et autres ESMS est disponible uniquement au niveau du département, mais pas au niveau de la commune.

NOTA BENE : cette table n'est pas encore exhaustive de tous les EHPAD, ni de l'ensemble des ESMS, elle sera complétée au cours du temps.

Il importe également de tenir compte du fait qu'il est impossible de retrouver :

- les soins inclus dans l'option tarifaire (ni dans cette table, ni dans le DCIR),
- les médicaments utilisés, si l'établissement possède une pharmacie à usage intérieur (PUI).

PARTIE 2

La qualité de la géographie

JE RETIENS

La qualité de la donnée géographique dans le SNDS est essentielle pour la réalisation d'analyses territoriales fiables, qu'il s'agisse de caractériser des populations à l'échelle communale, de calculer des indicateurs à des niveaux infrarégionaux, qui reposent sur l'agrégation de valeurs communales (EPCI, CLS, CLSM, zone d'emploi, territoire de démocratie sanitaire...), ou aux niveaux départemental et régional.

De manière globale, il est estimé que **95 % des codes géographiques retrouvés dans le SNDS sont « fiables »**. Par fiable, on entend existants et concordants entre les différentes sources de données géographiques (PMSI, DCIR).

Une étude de comparaison des données du DCIR et des données fiscales faite par la Drees a montré 96 % de convergence entre les départements. Un projet d'étude au niveau communal est en cours.

Cependant, les informations de localisation du bénéficiaire (variables BEN_RES_DPT et BEN_RES_COM) peuvent présenter un certain nombre d'atypies qui nécessitent des traitements correctifs spécifiques. C'est le cas des territoires d'outre-mer ou des Corses pour lesquels des retraitements sont nécessaires, mais aussi il peut y avoir des erreurs de liquidation qui génèrent des situations atypiques (point d'ancrage par exemple). Enfin, la géographie des publics spécifiques peut contribuer à afficher une donnée géographique faussée, car partielle ou tronquée (cf. partie 3).

DES CORRECTIONS SONT POSSIBLES

– Usage de la géographie corrigée de la table de la Cartographie des pathologies.

La Cnam fait tourner un programme de correction chaque année sur l'ensemble des bénéficiaires. Pour plus de détail, se reporter à la section cartographie des pathologies de la partie 1 du document. À noter que cette table ne permet pas d'avoir les informations corrigées pour les données les plus récentes : par exemple, à partir de juillet 2025, il est possible de disposer des informations corrigées pour les années 2015 à 2023, mais pas pour les données 2024.

– Usage de table de correction.

La base T_FIN_GEO_LOC, initiée par Michel SEGUIN (ARS Île-de-France) et disponible dans le dossier RF_COMMUN, permet de corriger ces erreurs par appariement avec les bases Hexaposte enrichies.

Pour automatiser ce traitement, **la macro SAS macro_corrections_code_commune_DCIR_DCIRS**, développée par Jérôme BROCCA (DNUM), est largement utilisée. Elle intègre les anomalies connues et peut être enrichie localement, notamment pour des territoires complexes comme la Corse ou les DOM.

Les codes Insee évoluent au fil des années en raison de fusions ou de créations de communes. L'utilisation de référentiels historiques est donc nécessaire pour restituer correctement les données à une année donnée.

La base T_GEO_COM, basée sur le projet **AtlaSanté** (projet mutualisé inter-ARS porté par la DNUM du ministère de la Santé) et mise à jour trimestriellement, permet de convertir les codes communes dans une nomenclature donnée sans devoir importer manuellement les fichiers Insee. Pour des usages dynamiques, l'API COG de l'Insee peut également être mobilisée.

La qualité de l'information géographique est une dimension essentielle dans l'exploitation du SNDS, en particulier lorsqu'il s'agit d'étudier les disparités territoriales de santé ou de relier des données individuelles à leur environnement. Or, les variables de localisation du bénéficiaire présentent parfois des imprécisions ou des biais, liés aussi bien aux sources administratives qu'aux processus de mise à jour. Ces limites peuvent affecter la robustesse des analyses menées et conduire à des interprétations erronées. D'où l'importance de questionner la fiabilité de la localisation disponible et d'explorer des méthodes de correction adaptées. Cette section propose un état des lieux des principales difficultés rencontrées, puis décrit les approches envisageables pour améliorer la qualité de la donnée géographique dans le SNDS.

1. Constats

En 2022, la Cnam a mené une étude interne sur la qualité des codes communes dans le SNDS. Cette étude a recensé l'ensemble des anomalies rencontrées par régimes et par année. Il est ressorti de cette dernière que 95 % des codes communes dans le SNDS étaient corrects.

Voici une comparaison des effectifs par commune avant l'apport de la correction et après l'ajout de celle-ci. Cette comparaison est effectuée pour les individus ayant consommé au moins un soin liquidé en 2020 en prenant la commune de résidence à la dernière date de liquidation de l'année pour les individus vivants au 31 décembre 2020. Pour 95 % d'entre eux, la commune avant et après correction est identique.

	Effectif	Pourcentage
Code commune manquant après correction mais présent avant correction	25 119	0,04%
Code commune présent après correction mais manquant avant correction	2 560 148	3,78%
Code commune différent après correction	771 998	1,14%
Code commune similaire avant et après correction	64 327 313	95,04%
Total	67 684 578	100%

Source Cnam

1.1 Insee versus consommateurs

La Cnam a comparé les effectifs de consommateurs par commune et au nombre de consommateurs dans l'année vivants au 31 décembre avec la population du recensement de l'Insee par commune.

Certaines raisons peuvent s'expliquer ces écarts :

- Entre 2,5 et 3 millions de déménagements sont dénombrés par an. Un décalage temporel est lié à la déclaration du déménagement et/ou du temps que l'information parvienne dans les bases de données de l'Assurance maladie *via* une consommation de soins. Aussi, cela peut faire varier les effectifs des codes communes SNDS : avec des effectifs plus élevés dans certaines communes que ce que fournit l'Insee, et inversement.
- La présence de doublons de NIR, notamment sur les données les plus anciennes. Le NIR unique n'étant pas renseigné pour certains bénéficiaires, il n'est pas possible de faire le lien entre leurs différents identifiants, notamment entre différents régimes d'affiliation.
- La population des étudiants est atypique, car en fonction du lieu d'étude, les étudiants sont amenés à se déplacer régulièrement d'une année sur l'autre. Certaines villes comportant des universités, cela peut impacter le nombre de résidents. Étant donné que les étudiants sont jeunes et qu'ils consomment moins de soins que la population générale, les informations reçues sont moins fréquentes et il se peut que cela impacte les effectifs.
- Il ne s'agit que de la population ayant consommé des soins. Il est estimé qu'au moins 6 % de la population ne consomme pas de soins chaque année. Cette population varie sensiblement en fonction des classes d'âge.
- Les personnes vivant exclusivement en EHPAD avec pharmacie à usage interne ne sont pas comptabilisées.

- L'attribution de NIR provisoires pour les étrangers ayant des soins sur le territoire français peut augmenter les effectifs de certaines communes. Lors de l'obtention d'un NIR permanent, le lien entre les différents NIR n'est pas réalisé. Cela peut aussi augmenter le nombre de doublons.

1.2 Données Assurance maladie et données fiscales

L'EDP-Santé (Échantillon Démographique Permanent) fait le lien entre les données fiscales et celles de santé du SNDS. L'objectif est de croiser les informations fiscales et les informations de santé pour mieux analyser, par exemple, l'impact de certaines caractéristiques socio-économiques sur la santé, les dépenses de santé ou l'accès aux soins. Cet échantillon représente 4,4 % de la population.



Une comparaison entre les données du SNDS et les données fiscales a été faite par la Drees. Ainsi, en 2016, 96 % des individus présents dans les données fiscales et dans le SNDS avaient le même département dans les deux sources. L'analyse des 4 % restants donne les résultats suivants :

- beaucoup de Franciliens partis à l'étranger sans en informer leur caisse ;
- beaucoup de confusions entre les départements d'Île-de-France et d'autres départements proches Pas-de-Calais et Nord, Rhône et Isère ;
- 14 % des jeunes femmes de 22 ans ont des départements SNDS et EDP différents, contre 3 % des filles de 10 ans et 1 % des femmes de 80 ans.

En conclusion, la correspondance entre les sources dépend de l'âge.

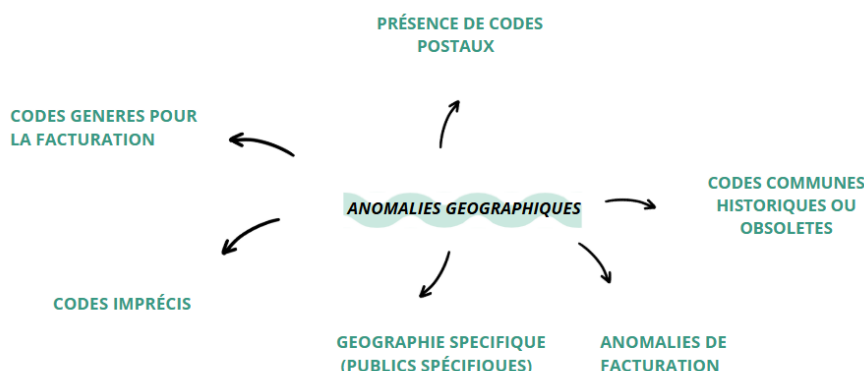
Tableau 4 : Correspondance du code commune entre la déclaration de la Taxe d'Habitation Principale 2017 et les remontées du DCIR 2017 (en pourcentage).

Code commune DCIR	population présente dans les deux sources	classes d'âge							
		0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 -
Premier code commune dans les remontées	85,9	89,5	85,9	66,0	83,9	88,5	88,0	90,8	92,2
Commune la plus fiable remontée	86,9	89,5	86,1	67,0	84,6	89,7	90,7	92,4	92,5
Commune corrigée	87,2	89,6	86,5	69,2	84,6	89,7	90,8	92,4	92,5

Champ : Individus EDP présents dans les deux sources (2,6 millions d'individus)

2 Quelles atypies ?

Les principales atypies sont présentées dans le schéma ci-dessous :



2.1 Présence de codes postaux

Une anomalie fréquente retrouvée dans les bases de données opérantes des caisses d'affiliation est l'usage de « codes postaux » au lieu du code commune.

Il est à noter que **près de la moitié des suites de chiffres constituant un code postal peut aussi correspondre à un code commune** : on ne peut donc dissocier *a priori* les individus concernés. Les corrections réalisées *via* l'algorithme de la Cnam permettent de cibler les caisses d'affiliation et périodes concernées.

La base permettant de convertir les codes postaux en code commune est hexaposte¹⁶, mais son usage pose question : un code postal couvre bien souvent plusieurs communes, notamment dans les espaces ruraux.

Plusieurs méthodes pour estimer l'information à la commune peuvent être mises en place : se ramener arbitrairement à une des communes, en prenant par exemple la commune la plus peuplée selon le recensement, ventiler les individus sur les communes en utilisant une pondération (calcul du poids de la population, totale ou par sous-groupe d'âge et sexe) ...

Au-delà de la commune, dans de rares cas, le département peut être affecté, car certaines communes sont couvertes par le système postal d'un département limitrophe, *cf.* tableau ci-dessous.

Codes communes et codes postaux dans des départements différents

Commune	Code commune	Code postal
CURBANS	04066	05110
PONTIS	04154	05160
ST PIERRE	04194	06260
FOUGUEYROLLES	24189	33220
VILLEBOIS LES PINS	26374	05700
LAJOUX	39274	01410
CHANGE	71085	21340
ELOISE	74109	01200
PARAY VIEILLE POSTE	91479	94390
ST PIERRE LAVAL	03250	42620
LA ROCHETTE	04170	06260
CLARET	04058	05110
PIEGUT	04150	05130
VENTEROL	04234	05130
LAVEYRUNE	07136	48250
LEYVAUX	15105	43450
PORT STE FOY ET PONCHAPT	24335	33220
CHANCIA	39102	01590
LAVANCIA EPERCY	39283	01590
ST EULIEN	51478	52100
SAPIGNICOURT	51522	52100
BUXEUIL	86042	37160
RIBOUX	83105	13780

À noter aussi : Monaco a le code postal 98000 et le code commune 99138.

¹⁶ <https://datanova.laposte.fr/datasets/laposte-hexasmal>. Quant aux cedex, on peut les trouver [ici](#).

Le choix le plus rigoureux consiste à se baser sur la fusion BEN_RES_DPT || BEN_RES_COM si l'on souhaite redresser le département en détectant les erreurs de codes postaux. Il arrive notamment que BEN_RES_COM soit correct, mais que le département utilisé par une caisse d'affiliation soit celui du code postal, aboutissant à une suite de chiffres qui correspond à un code commune existant ou non.

Comment corriger ?

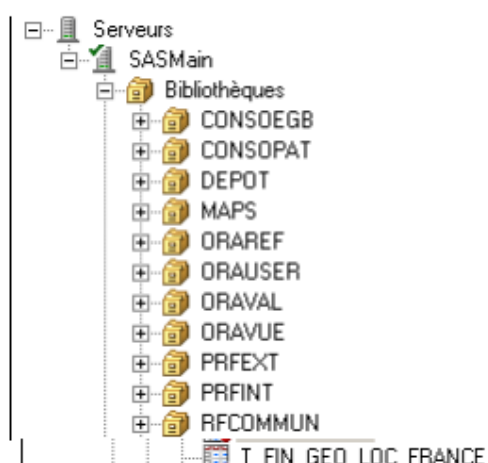
Michel SEGUIN de l'ARS Île-de-France est à l'initiative de la base **T_FIN_GEO_LOC** permettant la conversion code postal / code commune, disponible dans le dossier RFCOMMUN. Le principe est celui-ci :

Exemple 1 : Nous trouvons dans les tables de flux ou référentiels bénéficiaires BEN_RES_DPT = '075' et BEN_RES_COM = '001', ce qui donne '75001' après la concaténation (deux derniers caractères de la variable BEN_RES_DPT et des trois caractères de la variable BEN_RES_COM). Si nous cherchons le code Insee avec cette valeur ('75001') nous ne trouvons aucune correspondance. En revanche, si nous cherchons le code postal de ce CODE_JOINTURE ('75001'), nous le trouvons une SEULE fois. Nous pouvons donc faire correspondre le Code Insee ('75101') et le LIBELLE_DE_LA_COMMUNE ('PARIS 01').

Exemple 2 : Nous trouvons dans les tables de flux ou référentiels bénéficiaires BEN_RES_DPT = '077' et BEN_RES_COM = '871', ce qui donne '77871' après concaténation. Si nous cherchons le code Insee avec cette valeur ('77871'), nous ne trouvons aucune correspondance. En revanche, si nous cherchons le code postal de ce CODE_JOINTURE ('77871'), nous le trouvons CINQ FOIS, c'est-à-dire sur 5 codes Insee différents ! Dans ce cas-là, nous attribuons le code Insee de la commune où la population est la plus élevée, soit pour notre exemple le code Insee '77305' et le LIBELLE_DE_LA_COMMUNE ('Montereau Fault Yonne').

Présentation de la table « T_FIN_GEO_LOC_FRANCE »

Située dans la bibliothèque RFCOMMUN, la table « T_FIN_GEO_LOC_FRANCE » donne la correspondance entre les codes communes INSEE et les codes postaux présents dans les données de flux du DCIR. Elle est accessible à tout utilisateur SNIIRAM / SNDS à partir d'un profil accédant à SAS/Guide.



La macro SAS « **macro_corrections_code_commune_DCIR_DCIRS** » développée par Jérôme BROCCA (DNUM, ministère de la Santé) fait appel à cette base et la complète manuellement, notamment en Corse et Centre-Val de Loire.



Cette macro est disponible tous les profils.
Chacun peut s'en saisir et la compléter à sa guise.

exemple

```
/** On charge la Macro de correction des Codes Communes **/  
/** MACRO CORR_LOC **/  
* Paramètres en entrée ;  
*TAB_IN= Table initiale;  
*TAB_OUT= Table de sortie;  
/** Permet de réaffecter les codes communes à des Codes INSEE affectation des Départements --> Régions **/  
/** Correction des codes communes avec le DCIR/DCIRS :  
Nécessité de disposer dans la Table en entrée TAB_IN des Variables BEN_RES_DEPT et BEN_RES_COM **/  
/**Appel de la Macro par la fonction **/  
/* %corr_loc(TAB_IN=BEN_CONS_SPE,TAB_OUT=BEN_CONS_SPE_CORR); */  
%let Fichiers=%sysget(HOME)/sasdata;  
%include "&Fichiers/pfext/Macros/MacroCorrection Commune DCIR DCIRS.sas";
```

2.2 Codes communes historiques ou obsolètes

Une autre problématique concerne le Code Officiel Géographique (COG) de l'Insee : chaque année, la nomenclature est légèrement modifiée afin de tenir compte des mouvements de communes (fusions, créations, changements de rattachement, etc.). Ces évolutions, même marginales à l'échelle nationale, peuvent avoir un impact non négligeable sur la qualité des analyses territoriales menées à partir du SNDS, notamment lorsqu'il s'agit de suivre des populations ou des indicateurs de santé sur une longue période.

2.2.1 Les communes avant 1975

Comme précisé dans la partie 1 du rapport, l'adresse connue par l'Assurance maladie est celle qui a été fournie par l'assuré. Le code de la commune peut donc ne plus exister, en particulier pour les communes antérieures à 1975. Les nombreuses fusions et réorganisations administratives survenues à cette période ont entraîné la disparition de certains COG. Lorsque l'assuré n'a pas actualisé son adresse ou que l'information a été conservée telle quelle dans les bases, la commune déclarée peut correspondre à un code obsolète.

Exemple

La commune Commelles (60158), dans l'Oise, a été supprimée en 1974 pour fusionner avec Orry-la-Ville (60482), le code 60158 a donc cessé d'exister après 1975.

De même, dans le Rhône, la commune Saint-Rambert-l'Île-Barbe (69234) a été rattachée à Lyon en 1963. Son ancien code n'existe plus dans le COG actuel.

2.2.2 Les communes obsolètes

La présence de communes disparues dans les bases du SNDS illustre un problème plus large lié aux adresses obsolètes. Lorsque les assurés n'actualisent pas leur information auprès de l'Assurance maladie ou lorsque les fichiers conservent des références anciennes, il devient difficile de rattacher correctement ces individus à une commune existante. Ces anomalies peuvent avoir plusieurs conséquences :

- une perte de précision géographique, si le code est invalidé et remplacé par une information plus agrégée (ex. au niveau du département),
- un biais dans les analyses territoriales, certaines communes pouvant apparaître artificiellement sous-représentées ou sur-représentées.

Exemple

Période	Département	Code Insee	Statut	Remarques
Avant 1973	Calvados (14)	14509	Commune simple	Pont-Farcy existait seule, dans le Calvados
1973 - 2017	Calvados (14)	14509	Commune associée (fusion avec Pleines-Cœuvres)	Pont-Farcy absorbe Pleines-Cœuvres le 1 ^{er} mai 1973
1 ^{er} janvier 2018 à aujourd'hui	Manche (50)	50649	Commune déléguée de Tessy-Bocage	Rattachement à la Manche par décret du 26 décembre 2017 ; intégration dans Tessy-Bocage

Le code 14509 est obsolète après le 31/12/2017.

Depuis 2018, toutes les adresses de Pont-Farcy doivent être reclassées en 50649 (Tessy-Bocage, commune déléguée).

Attention aux données SNDS ou Insee couvrant plusieurs décennies : un assuré déclaré « Pont-Farcy (14509) » avant 2018 peut apparaître artificiellement dans le Calvados, alors que la commune se situe aujourd'hui dans la Manche.

2.2.3 Comment corriger ?

La table T_GEO_COM est le référentiel géographique de la base de données DataSanté : elle recense toutes les communes et arrondissements municipaux qui existent ou qui ont existé depuis 2010 et leur associe les codes officiels géographiques antérieurs, les découpages Insee (unités urbaines, bassins de vie...) et les découpages nationaux définis ou utilisés par les ARS (territoires de démocratie sanitaire, codes géographiques PMSI).

DataSanté est une base de données mise à disposition par le projet Atlasanté. Cet outil a été construit pour stocker et partager des informations provenant de différentes sources de données, afin d'alimenter les cartes interactives produites par Atlasanté (SIRSé, CartoSanté...).

Cette table est disponible dans la librairie RFCOMMUN du SNDS. Elle permet d'éviter que chaque utilisateur importe les tables de l'Insee dans son espace de travail.



Pour aller plus loin :
[Référentiel géographique \(t_geo_com\)](#)

2.3 Codes imprécis

2.3.1 Zones spécifiques

Départements d'outre-mer (DOM)

Le filtre sur les individus des DOM s'effectue sur la variable **BEN_RES_DPT**, mais aussi à l'aide de **BEN_RES_COM**.

BEN_RES_DPT peut avoir pour modalité « 097 », mais aussi « 971 », « 972 », « 973 » selon que la norme d'échange soit la NEC ou la NTEIR... Il est préconisé de convertir tous les départements « 97X » (dont la validité est questionnée), en « 097 » et se fier ensuite au premier chiffre de la variable BEN_RES_COM.

Il existe une subtilité concernant la Guadeloupe (« 971 ») : les codes communes de Saint-Martin et Saint-Barthélemy sont respectivement 97801 et 97701 alors que leurs codes postaux sont 97150 et 97133.

La base IR_GEO_V peut permettre de comprendre, grâce au libellé de la commune, la zone géographique associée à un code.

Corse

S'il peut arriver de rencontrer des BEN_RES_DPT valant « 201 » ou « 202 », la plupart des individus en Corse sont codés en « 209 », ce qui ne permet pas de distinguer la Haute-Corse et la Corse-du-Sud. Il faut dans ce cas concaténer la valeur « 20 » à la variable BEN_RES_COM pour obtenir le code commune, qu'on pourra alors classer grâce à T_FIN_GEO_LOC (dans RF_COMMUN).

2.3.2 Comment corriger ?

Les erreurs retrouvées peuvent concerner plus spécifiquement certaines caisses d'affiliation (ORG_AFF_BEN), éventuellement durant une période définie. Certaines ont été relevées directement par des utilisateurs ou d'autres communiquées par la Cnam ; elles donnent lieu à des corrections.

La Cnam a ainsi répertorié un nombre important d'erreurs et s'appuie sur un algorithme de ciblage et de correction, abordé dans la suite de ce rapport. Il alimente le code commune dans la Cartographie des pathologies et des dépenses ; table CRTO_CT_RES_G10 (dans ORAMEPS).

2.4 Anomalies de facturation

2.4.1 Erreurs communiquées par la Cnam

La Cnam publie régulièrement *via* un fichier anomalie disponible dans le portail, la liste des problèmes rencontrés. Ce fichier est historisé et permet de retrouver l'ensemble des anomalies rencontrées.

Erreurs liées à certains organismes de liquidation

Exemple du communiqué publié le 24 mai 2023 sur les communes Corse (il est consultable dans la doc du SNDS dans la partie BO du portail Cnam, rubrique Documentation) :

- « Par moment, les caisses d'affiliation commençant par « 01C201 » et « 01C202 » accordent respectivement à tort les départements « 037 » et « 038 » aux individus au lieu de « 209 » (« 20 » pour Corse) ».

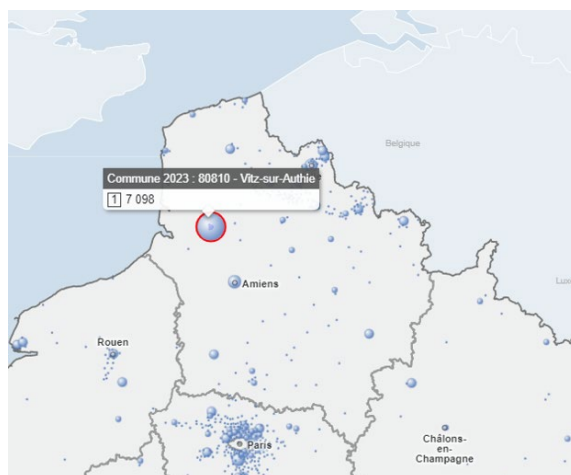
Attention, cette anomalie de codage des communes corses ne concerne pas tous les bénéficiaires de Corse.

Erreurs d'affectation du code commune

Ces erreurs d'affectation peuvent être importantes et constituer des points d'ancrage. On désigne par « point d'ancrage » une situation dans laquelle un grand nombre d'individus sont rattachés administrativement à une même commune, alors même qu'ils n'y résident pas nécessairement. Cette commune se retrouve artificiellement sur-représentée dans le SNDS, avec un effectif de « consommateurs » de soins qui apparaît atypiquement élevé par rapport à sa population réelle. Ce phénomène peut résulter d'anomalies de codage, d'adresses de gestion utilisées par certaines caisses d'Assurance maladie ou encore de défauts de mise à jour des adresses déclarées.

Exemple de la commune Vitz-sur-Authie (88810)

Le point d'ancrage est bien visible sur la carte ci-dessous. Cette anomalie provient d'une génération automatique d'un code commune pour des individus qui mutaient d'un organisme à un autre.



2.4.2 Comment corriger ?

Une correction en utilisant l'organisme d'affiliation du bénéficiaire est possible.

Exemple de la commune Vitz-sur-Authie (88810)

```
If BEN_RES_DPT in('080') AND BEN_RES_COM='810' AND SUBSTR(ORG_AFF_BEN,1,3) in('01C')
AND SUBSTR(ORG_AFF_BEN,1,5) not in('01C97') then BEN_RES_DPT='0'||SUBSTR(ORG_AFF_BEN,4,2) ;
```

```
If BEN_RES_DPT in('080') AND BEN_RES_COM='810' AND SUBSTR(ORG_AFF_BEN,1,3) in('01C')
AND SUBSTR(ORG_AFF_BEN,1,5) in('01C97') then BEN_RES_DPT=SUBSTR(ORG_AFF_BEN,4,3) ;
```

2.5 Géographie spécifique, publics spécifiques

Certaines populations présentes dans le SNDS ne sont pas localisées selon leur lieu de résidence réelle, mais rattachées administrativement à une entité particulière. Cela conduit à une géographie « conventionnelle » qui ne reflète pas la réalité territoriale des individus concernés.

Un exemple emblématique concerne les NIR fictifs utilisés pour la facturation des interruptions volontaires de grossesse (IVG). Dans ce cas, l'assuré n'est pas identifié par son propre numéro, mais par un numéro technique dit « NIR fictif » créé pour les besoins de la prise en charge. La localisation géographique n'est donc pas celle du domicile de la patiente, mais celle de la Cnam du département qui assure la gestion administrative de l'acte.

D'autres cas sont développés dans la partie 4 de ce rapport.

Ces spécificités entraînent plusieurs conséquences :

- une surestimation artificielle de l'activité de soins ou du nombre d'assurés dans certains départements ou certaines communes,
- une perte d'information territoriale empêchant l'analyse fine des inégalités de santé par lieu de résidence,
- des limites d'interprétation pour les études territoriales, qui doivent tenir compte de ces biais liés à des logiques administratives et non géographiques.

Ces exemples montrent que, pour certains publics spécifiques (patients avec NIR fictif, bénéficiaires de l'AME, assurés rattachés à des dispositifs exceptionnels), **la géographie observée dans le SNDS ne correspond pas à la réalité des territoires de vie**. Leur prise en compte nécessite donc des précautions méthodologiques, et parfois l'exclusion ou le traitement particulier de ces populations dans les analyses territoriales.

L'algorithme de correction de la commune de résidence de la Cnam

L'algorithme de correction de la commune de résidence de la Cnam alimente le code de la commune du référentiel annuel de la cartographie des pathologies à partir de la version G10¹⁷.

Les travaux de la Cnam s'appuient sur la connaissance fine de l'origine de problèmes de remontée *via* les caisses d'affiliation (ORG_AFF_BEN) à partir de 2013.

Dans un premier temps, les corrections s'effectuent au niveau de chaque ligne de la table des prestations (ER_PRS_F).

Une note de fiabilité décroissante comprise entre 1 et 5 est donnée pour chaque correction.

À chaque date de remontée de flux (FLX_DIS_DTD) de la période étudiée, l'algorithme de la Cnam met à jour la commune des consommateurs, en gardant la plus « fiable »¹⁸.

Dans les travaux réalisés par la Cnam, la commune de résidence des individus consommateurs en 2020 a été comparée à celle attribuée par l'algorithme appliqué à la période 2013-2020 ; près de 5 % des individus ont une commune de résidence différente. De plus, il est constaté une diminution de l'écart entre le nombre de consommateurs dans une commune et le nombre d'habitants selon l'Insee.

Un exemple de code de correction, réalisé au sein du sous-groupe « la qualité de la géographie », est présenté en annexe.

¹⁷ Commune attribuée au mois de décembre de l'année.

¹⁸ Ainsi, un changement de commune entériné par la caisse d'affiliation qui aboutirait à une commune moins fiable est ignoré. Une autre solution serait de conserver la commune obtenue (et son département) en signalant sa moindre fiabilité.

PARTIE 3

La géographie des populations spécifiques

JE RETIENS

La géographie des populations spécifiques est à la fois le motif et la conséquence des objectifs d'une étude de santé et donc de la définition même de la population spécifique.

L'exploitation du SNDS permet de conduire des études ciblées sur des populations spécifiques dont les comportements de recours aux soins ou les trajectoires de santé diffèrent sensiblement de ceux observés dans la population générale. Ces différences peuvent résulter de facteurs géographiques (lieu de résidence du bénéficiaire, localisation de l'offre de soins), mais aussi d'éléments liés à l'affiliation à un régime de sécurité sociale ou à des dispositifs de prise en charge particuliers.

Le groupe méthodologique SNDS a ainsi exploré plusieurs sous-populations à titre d'illustration (par exemple, résidents de DOM, bénéficiaires de régimes spécifiques, zones rurales ou insulaires, etc.). Ces choix sont contextuels et non exhaustifs, guidés par les travaux des membres du groupe et les problématiques de santé publique identifiées.

Il est important de souligner que la définition d'une population spécifique est toujours liée aux objectifs propres à chaque étude. En conséquence, la manière d'appréhender la dimension géographique peut varier selon les projets : la géographie n'est pas un invariant, mais un modulateur des résultats, à interpréter au regard du cadre méthodologique adopté.

1. Préambule aux populations spécifiques

En utilisant le SNDS, des études sont proposées sur des populations particulières dont les consommations de soins impliquent des résultats parfois différents de la population générale. Les résultats peuvent différer en raison de la géographie du bénéficiaire, de la localisation de l'offre de soins, de l'appartenance à un organisme de prise en charge, etc.

Le groupe méthodologique s'est intéressé à certaines populations et aux conséquences ou aux causes de la géographie sur les résultats. Les populations spécifiques traitées dans ce document sont la résultante de sujets d'intérêts des membres du sous-groupe. Le nombre de populations particulières a été volontairement limité. Cependant, la géographie d'autres populations pourra être étudiée par la suite, en fonction des demandes et intérêts des membres et futurs membres du groupe méthodologique.

La définition des populations spécifiques est un choix réalisé par les auteurs de l'étude en fonction de leurs objectifs. Par conséquent, les résultats concernant la « géographie » de ces études peuvent varier au gré des études et des objectifs.

En raison de la construction historique et légale de la Sécurité sociale en France, l'affiliation à un régime de sécurité sociale dépend principalement des ressources liées au travail. Cependant, depuis la création de la protection universelle maladie (PUMa), cette affiliation s'est élargie à des conditions de résidence (Métropole et Départements d'outre-mer). Ainsi en France, l'affiliation ne dépend pas de la nationalité de l'individu.

2. Les femmes enceintes et accouchements

2.1 Exemples de travaux reposant sur l'utilisation de distancier

Deux travaux ont été menés par la Drees sur la « géographie » des femmes enceintes dans le SNDS (Jeanne Fresson, Gladys Baudet et Alice Bergonzoni). Le premier travail concernait l'impact sur les accouchements de la distance entre la résidence et la maternité: accouchement hors établissement, mortalité périnatale et recours à la césarienne. Le second travail évaluait la population susceptible de bénéficier de la prise en charge de prestation hôtelière en raison d'une distance à la maternité supérieure à 45 min et de l'effet potentiel des regroupements de maternités en raison d'une faible activité.

Ces travaux ont étudié particulièrement l'impact de l'utilisation de distancier sur les études.

- Le premier distancier est METRIC ([MEsure des TRajets Inter-Communes / Carreaux](#)), distancier dépendant de l'Insee, qui est probablement en cours d'abandon. Il utilise le chef-lieu de commune comme point central. Cette localisation n'est cependant pas toujours fiable. Il propose des calculs de temps paramétrable en heures de trafics importants et heures de trafics creux. Le temps de parcours est mis à 0 si le départ et l'arrivée sont dans la même commune.
- Le second distancier est l'OSRM open source routing machine (<http://project-osrm.org>), sur R, utilisant openstreetMap de 2018, distance entre deux coordonnées précises, sans prise en compte du trafic, ni des virages. OSRM calcule les distances par carroyage pondérées par la population.

L'utilisation de la distance plutôt que du temps semble être plus robuste et permet des comparaisons avec les autres études. Cependant les organismes d'état se fondent plutôt sur des notions de temps.

Une correspondance a été faite entre le code géographique/code postal du PMSI et le code de la commune du DCIR au moment du remboursement du soin (date la plus proche de l'accouchement). Cette comparaison nécessite un nettoyage de la base et une prise en compte de l'historique des communes. L'avantage de la géographie du PMSI est qu'elle est déclarative et semble être un meilleur reflet de la localisation géographique habituelle de la femme au moment d'accoucher.

Les localisations de la Guyane, de Mayotte, mais également des petites îles peuvent entraîner des distorsions pour le calcul des distances en temps ou en mètres.

Pour les travaux sur les maternités, la distance a été calculée à partir du centroïde du carreau. Les coordonnées des établissements ont été prises en compte. Cependant il y a une approximation de la distance qui calcule à la route la plus proche, celle-ci pouvant être, par exemple, une autoroute. Il est possible que l'approximation ne soit pas juste. De même, n'est pas prise en compte la circulation « interne » dans l'établissement.

L'utilisation du distancier fourni par l'ATIH est également possible. Cependant, il apparaît que les vitesses de déplacement sont plus élevées et les informations sur les sources et les mises à jour ne sont pas évidentes à trouver.

D'autres distanciers existent mais ils sont parfois payants (Google Maps). Ils fournissent également des informations non stables dans le temps, car elles sont actualisées très fréquemment.

En résumé, une étude sur la géographie des femmes enceintes et des accouchements comprenant un aspect de distance doit :

- procéder à un nettoyage de la base de travail avec une prise en compte de l'historique des communes,
- privilégier des données issues du PMSI (déclaration de l'adresse au moment des soins),
- utiliser un distancier nécessitant de préciser le type de fonds de carte, la date, la vitesse ou la distance en mètre, le lieu pris en compte (centre du carreau, chef-lieu, prise en compte de particularité de commune (Arles, Guyane, petites îles...)),
- émettre l'hypothèse de la répartition homogène des femmes enceintes dans la commune.

La géographie concernant la femme enceinte et l'accouchement peut être abordée *via* deux approches, dépendant de l'objectif à traité, elles sont élaborées par l'Agence régionale de santé dans son *open data* (<https://opendata-perinat.sante-idf.fr/app/>) :

- Les données domiciliées par territoire (par exemple toutes les patientes domiciliées dans un département particulier) ;
- Les données enregistrées dans les établissements où s'est déroulé le séjour (un séjour peut être enregistré dans un établissement en Île-de-France même si la patiente n'y réside pas).

2.3 Les accouchements à domicile

Le repérage de l'accouchement à domicile est identifiable de plusieurs manières. Comme toutes les données du SNDS, il dépend de la qualité et de l'exhaustivité de la saisie. L'accouchement est repéré par différents codes Cim 10 et actes (voir Blotière P-O, Weill A, Dalichampt M, et al. *Development of an algorithm to identify pregnancy episodes and related outcomes in health care claims databases : An application to antiepileptic drug use in 4.9 million pregnant women in France*. Pharmacoepidemiol Drug Saf. 2018;27: 763–770. <https://doi.org/10.1002/pds.4556>). Il est important de se poser la question sur la différence entre le lieu de l'accouchement hors établissement de santé et celui de l'accouchement à domicile qui peut ne pas être le lieu de résidence habituel de la personne.

L'identification des naissances à domicile et hospitalisation ensuite

Il s'agit des accouchements sans actes d'accouchement dans le PMSI, mais avec un séjour pour la mère dans un établissement. Ils sont identifiés également *via* le code z3900 (*cf.* ATIH - Consignes de codage - Gynécologie Obstétrique, Fascicule V - Créé le 1^{er} mars 2011). À compter de 2011, un code de la catégorie Z39 doit être employé pour les hospitalisations du post-partum. Le code Z39.0 fait l'objet d'une extension :

- Z39.00 : Soins et examens immédiatement après un accouchement hors d'un établissement de santé. C'est le cas en particulier lorsque la patiente a accouché à domicile ou dans un véhicule sanitaire et est hospitalisée pour surveillance ;
- Z39.08 : Soins et examens immédiatement après l'accouchement, autres et sans précision. Ce code est employé lorsqu'après accouchement dans un établissement de santé « A » une mère est transférée dans un établissement de santé « B » pour les soins du post-partum.

Le mode d'entrée permet également de faire un repérage de ces accouchements hors établissements

Pour rappel, en 2019, le mode d'entrée « N » a été créé pour identifier les séjours de naissance lors de naissance dans un établissement de santé. Ainsi :

- les séjours concernant un nouveau-né, né à domicile ou hors d'un établissement de santé doivent être enregistrés avec un mode d'entrée « 8 »,
- les séjours concernant les nouveau-nés transférés doivent être enregistrés avec un mode d'entrée « 7 »,
- les nouveau-nés nés en établissement de santé doivent être enregistrés avec un mode d'entrée « N »,
- il est également indiqué, dans le guide méthodologique, que le mode d'entrée « N » doit aussi être enregistré dans le RUM du nouveau-né, dans le cas de naissance hors d'un établissement et en présence d'un SMUR.

La consigne concernant le mode d'entrée « N » génère une différence de statut entre le RUM de la mère, qui est considéré comme un séjour de post-partum, et celui du nouveau-né qui est considéré comme un séjour de naissance dans l'établissement.

Afin de rendre cohérents ces deux séjours, il est dorénavant demandé depuis 2025 de coder un mode d'entrée « 8 » (domicile) pour tous les nouveau-nés, nés hors de l'établissement, y compris si la naissance a lieu en présence du SMUR.

La notion d'accouchement à domicile ou hors établissement de santé et sans intervention du SMUR peut être repérée par un acte CCAM dans le DCIR sans hospitalisation

Les nouveau-nés : le code postal du lieu de résidence d'un nouveau-né est celui de la résidence de sa mère, il répond aux mêmes règles de saisie que pour les autres codes de lieu de résidence.

3. Les enfants

L'affiliation à un organisme d'Assurance maladie se fait par l'ouvreur de droit. En France, le modèle de protection sociale étant de type Bismarckien, l'ouverture de droits est principalement liée au travail. Par conséquent, les enfants sont affiliés en tant qu'ayant droit de l'ouvreur de droit. Depuis 2016, la Protection Universelle Maladie (PUMa) permet l'affiliation à un régime d'Assurance maladie sous condition de résidence et permet de bénéficier de l'Assurance maladie à titre personnel pour les mineurs ayant au moins 16 ans.

Le Répertoire National inter-régime de l'Assurance Maladie (RNIAM) contient des informations d'identification des assurés et des ayants droit. Les informations contenues dans le RNIAM sont confrontées à celles provenant respectivement du Répertoire national d'identification des personnes physiques (RNIPP) tenu par l'Insee et des organismes gestionnaires des régimes d'Assurance maladie. Le RNIAM contribue aussi aux procédures de délivrance et de mise à jour des cartes électroniques Vitale. Le répertoire sert enfin à des études statistiques dans les domaines sanitaire, social et démographique. La gestion technique du RNIAM a été confiée à la Caisse nationale d'assurance vieillesse (Cnav) pour le compte, et sous le contrôle, des organismes d'Assurance maladie.

Les mineurs peuvent être ayants droit d'un ou plusieurs ouvriers de droits. Cependant, pour faciliter certains traitements, un ouvrier de droit est défini prioritairement, les autres sont « créés pour ordre ». Les ouvriers de droits peuvent ou non appartenir à des organismes d'affiliation différents. Les organismes d'affiliation utilisant la norme d'échanges commune (NEC), indiquent les « créer pour ordre » par le code RNIAM W ou X (lorsque la demande est en cours).

La distribution des codes W et X (MAX_TRT_DTD > 01JAN2022) par organisme d'affiliation figure dans le tableau ci-dessous. L'âge médian des bénéficiaires est autour de 10 ans et 95 % des bénéficiaires ont moins de 18 ans.

BEN_CDI_NIR représente approximativement le nombre d'affiliés certifiés dans le régime, le pourcentage est la fréquence de RNIAM W ou X parmi le nombre de BEN_CDI_NIR.

Distribution des codes W et X* par organisme d'affiliation

	ORG	W	X	BEN_CDI_NIR	%
Régime général	01C	8 185 501	12 039	78 890 536	10 %
Régime général/Mutuelles	01M	75 446	169	5 738 701	1 %
MSA	02A	3 740	75		
RSI	03A	78	12		
SNCF	04A	103	1		
RATP	05A	65	2		
ENIM	06A	6 486	1	139 718	5 %
Régime Minier	07A	115	2		
CNMSS**	08A	146 897	31	1 452 080	10 %
Clerc de notaire	10A	7 988	8	171 754	5 %
Assemblée nationale	14A	11			
Sénat	15A	8			
Culte	90A	623	3	64 884	1 %
	91C	9			
	99I	45			

* MAX_TRT_DTD > 01JAN2022

** cf. section « 10. La population protégée par la caisse nationale militaire de sécurité sociale »

Les personnes ayant un code RNIAM créé pour ordre ont une commune différente, dont la fréquence varie selon le régime (cf. tableau ci-dessous). Cette fréquence de communes différentes augmente lorsque les ouvriers de droit n'appartiennent pas au même régime d'affiliation. Au vu de ces données, il est nécessaire lors d'une étude sur les enfants, de décider si on prend la dernière commune connue ou celle au moment de l'exécution des soins et si l'on prend en considération l'éventuelle double affiliation, avec un organisme d'affiliation pouvant être différent. La date de la dernière mise à jour et de l'identifiant BEN_IDT_ANO sont des variables à utiliser pour ce type d'études.

Fréquence des communes différentes selon le régime d'affiliation de l'ouvrier de droit

	Nombre de communes différentes	Nombre de communes identiques	% de communes différentes
Régime général	2 007 390	6 104 819	25 %
Régime général/Mutuelles	27 720	40 665	41 %
ENIM	2 561	3 818	40 %
Régime Minier	35	52	40 %
CNMSS*	109 228	32 198	77 %
Clerc de notaire	2 214	4 996	31 %
Culte	120	489	20 %

* cf. section « 10. La population protégée par la caisse nationale militaire de sécurité sociale »

4. Les étudiants

La Section locale mutualiste étudiante (SLM, qui gère la part légale Assurance maladie pour le compte du régime général) a pris fin en septembre 2019. Avant cette date, les étudiants étaient affiliés à un centre de Sécurité sociale étudiante (LMDE et mutuelles régionales) qui avait une délégation de gestion de la part légale Assurance maladie.

Il était alors possible de repérer les étudiants dans les bases du SNDS *via* leur organisme d'affiliation selon les codifications suivantes :

SLM N°601 - LMDE - La Mutuelle des étudiants

Codification	Jusqu'en septembre 2015 (en 303)	À partir d'octobre 2015 : SLM en infogérance = caisse 909
Organisme de liquidation	01Mxxx601	01C909000
Organisme de rattachement	01Mxxx601	01Mxxx601

Anomalie de codification de l'organisme d'affiliation : de décembre 2015 à juin 2016, une partie des données (40 % des données de décembre 2015 en date de traitement) était codée à tort sous le code organisme d'affiliation = 01M909601 au lieu de 01Mxxx601. Ceci implique que l'affectation géographique selon la Cnam d'affiliation est impossible pour ces données.

SLM N°617 - Sociétés d'étudiants mutualistes

(SMEBA, SMERRA, SMESO, SMENO, MEP, MGEL, SMEREB, SMECO, SMEREP, SMERE ...)

Codification	Jusqu'en août 2019
Organisme de liquidation	01Mxxx617 ou 01Mxxx618
Organisme de rattachement	01Mxxx617 ou 01Mxxx618

Anomalie de codification de l'organisme d'affiliation : de 2013 à septembre 2020, il y a un fort taux de troncature du département de résidence du bénéficiaire pour la SLM 617 (à la fois dans ER_PRS_F et dans IR_BEN_R : le département de résidence commence par deux 0 puis le 1^{er} chiffre du département de résidence du bénéficiaire).

Par exemple : un bénéficiaire étudiant de la SLM 617 résidant à Paris (département 075) se retrouve dans ER_PRS_F et dans IR_BEN_R avec un département de résidence BEN_RES_DPT = 007.

Correction proposée : prendre le département de la Cnam d'affiliation (le xxx du code organisme) ou le département d'exercice du professionnel de santé.

Aujourd'hui, les étudiants sont affiliés au régime général, il n'existe pas de code petit régime pour les repérer.

Jusqu'à leurs 24 ans, les étudiants rentrants à l'université conservent le régime qu'ils avaient auparavant (le plus souvent celui des parents). À partir de 24 ans, chaque étudiant bascule vers le régime général obligatoire de l'Assurance maladie, quel que soit le régime dont il relevait antérieurement (généralement celui de ses parents). Chaque étudiant doit ainsi faire la démarche auprès de l'Assurance maladie afin d'être affilié à la Cnam du département dont il dépend en fonction de son lieu de résidence. Leurs département et commune de résidence sont alors relativement bien renseignés.

5. Les personnes en situation de handicap

5.1 Exemple d'une étude sur les limitations fonctionnelles

La géographie concernant l'étude sur les limitations fonctionnelles a été réalisée par une équipe de l'Institut de recherche et de documentation des études en santé (Irdes) : Maude Espagnacq, Catherine Sermet, Camille Regaert, Fabien Daniel.

Le repérage des personnes avec des limitations motrices ou organiques sévères présenté ici est le résultat de l'indicateur Fish (Faisabilité d'identification des personnes en situation de handicap), construit à partir des données du SNDS. Source administrative déjà issue des remboursements de soins par l'Assurance maladie et échappant aux biais déclaratifs des enquêtes, cet indicateur utilise des données accessibles en continu et sur longue période, sur l'ensemble de la population, quels que soient l'âge ou le lieu de vie permettant ainsi d'améliorer la connaissance des populations dans le domaine de l'autonomie.

Ces travaux sont aussi les premiers à proposer une comparaison départementale, tous âges et lieux de vie, du risque de handicap à la suite de ce type de limitations. En 2019, la population ayant des limitations motrices ou organiques, identifiée par l'indicateur, est légèrement plus large que dans les sources déclaratives (14 % contre 12 %, soit près de 10 millions de personnes). Si globalement la population souffrant de ces limitations est plutôt âgée et féminine, cet indicateur identifie aussi une population non repérée dans les enquêtes : population plus jeune et plutôt masculine. Cet indicateur met également en lumière des disparités départementales qui s'accroissent avec l'avancée en âge.

Un indicateur sur les limitations psychiques intellectuelles et cognitives a également été construit. En 2019, cette population est composée de 4,5 millions de personnes (soit 7 % de la population). L'algorithme repère 4,5 millions de personnes ayant des troubles psychiques, intellectuels ou cognitifs conduisant à des limitations dans les activités de la vie quotidienne, soit 6,6 % de la population identifiée dans le SNDS. Cette population est en moyenne âgée de 55 ans, avec une grande différence selon le genre : un âge moyen de 59,6 ans pour les femmes et 49,8 ans pour les hommes. La majorité d'entre elles a des limitations à la suite d'un trouble psychique seul (2,8 millions) ou à la suite de troubles psychiques accompagnés de troubles intellectuels ou cognitifs (1,3 million), et seulement 266 000 du fait de troubles intellectuels ou cognitifs sans troubles psychiques associés.

Les algorithmes sont disponibles dans la bibliothèque ouverte des algorithmes de santé (BOAS) du HDH.

La méthodologie utilisée pour la géographie est la suivante :

L'identification du lieu de vie est réalisée à partir de la variable BEN_RES_DPT dans la table IR_BEN_R en reprenant les instructions de la note et en restant seulement sur le niveau départemental.

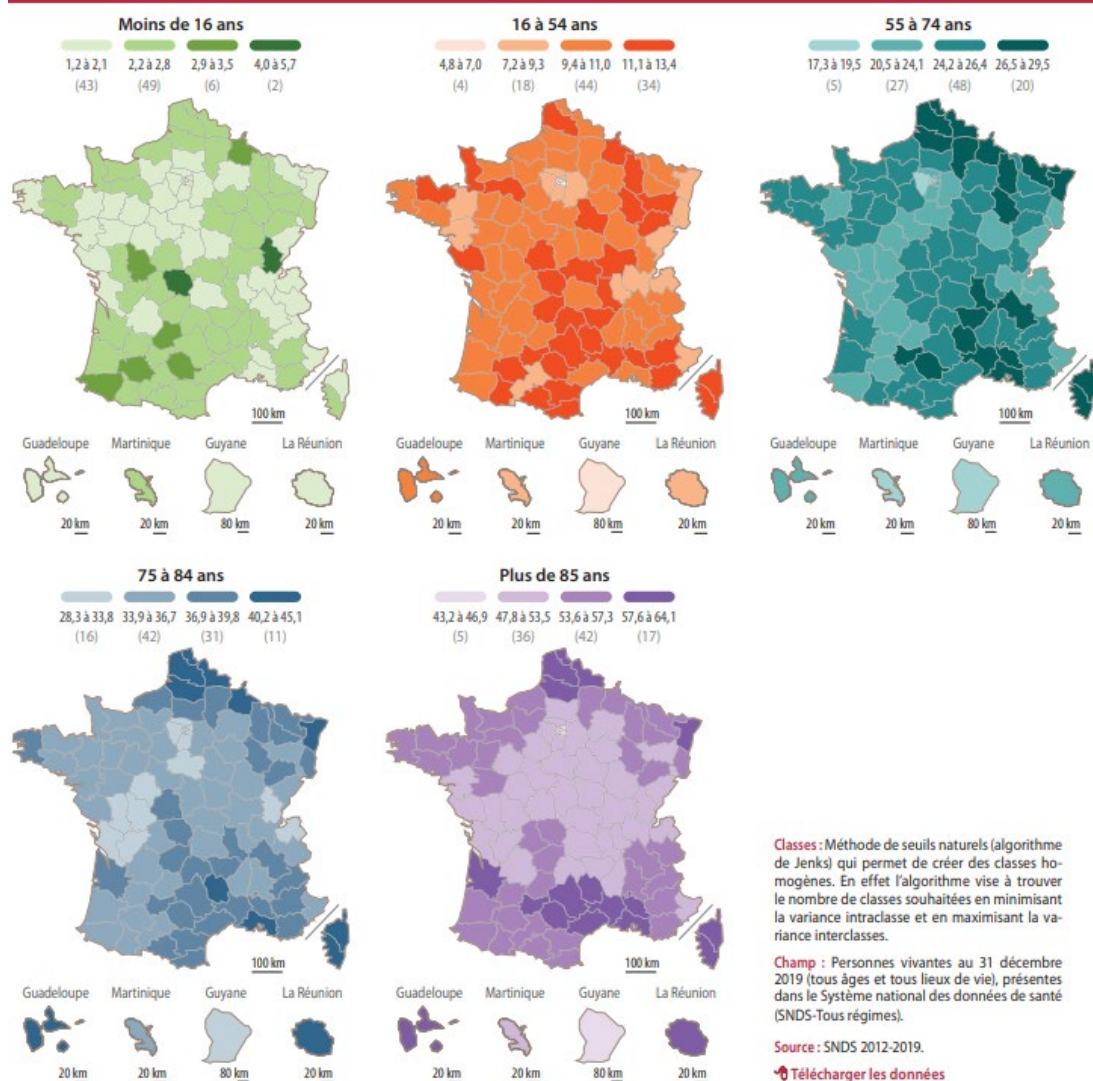
L'algorithme est créé en 2023, dans un premier temps pour l'année 2019 et en utilisant les années 2012-2019. Il y a ainsi un biais, car c'est le lieu de vie (BEN_RES_DPT dans IR_BEN_R) de 2023 et non celui de 2019 qui a été utilisé ; les changements de lieu de vie intervenus entre temps ont pu entraîner un biais. L'objectif de l'algorithme étant d'être produit en routine tous les ans, ce biais se réduira donc à terme, car l'année d'étude correspondra au lieu de vie (BEN_RES_DPT) mis à jour dans IR_BEN_R lors de l'année d'étude.

De plus, les travaux de l'algorithme sont basés sur la consommation de soins des individus. Ainsi, dans certains départements (notamment la Seine-Saint-Denis et la Guyane) la sous-consommation et le non-recours aux soins de la population impliquent un repérage légèrement plus faible provoquant un taux de limitation fonctionnelle sous-estimé.

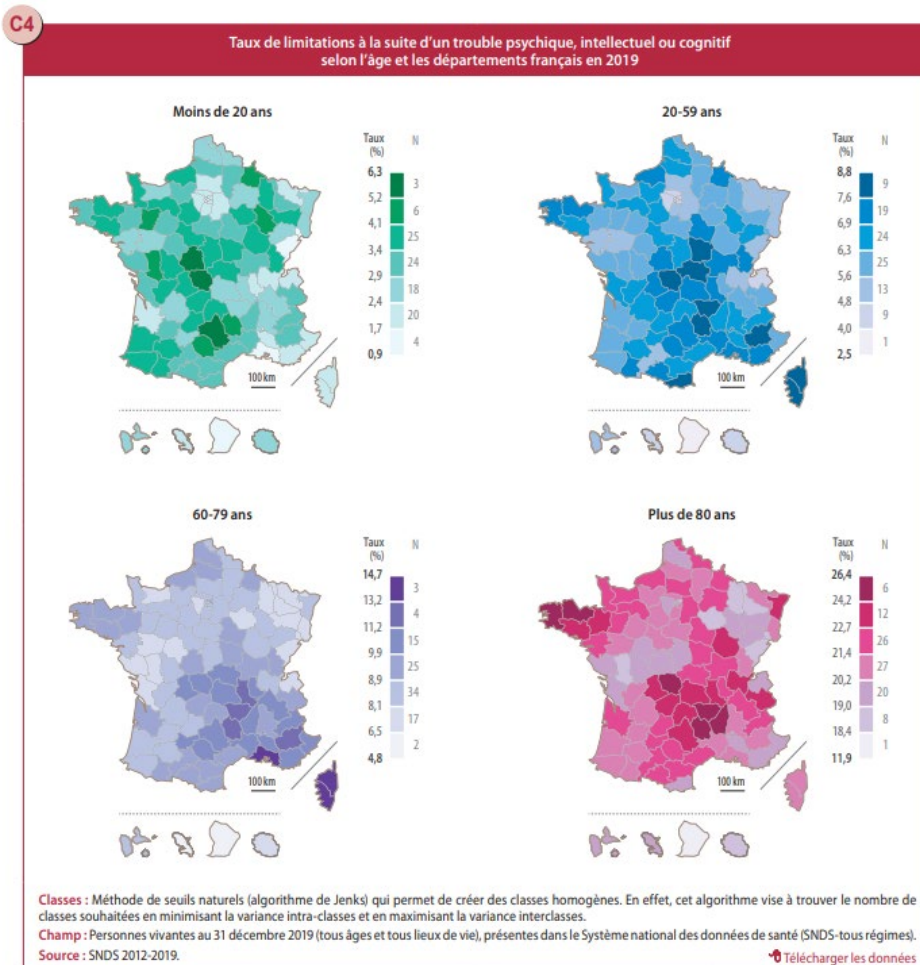
Inversement, d'autres départements ont des taux d'équipements en établissements et services médico-sociaux très supérieurs à la moyenne nationale ainsi que de nombreux EHPAD, ce qui implique un repérage de la limitation fonctionnelle plus important dans ces départements (comme la Creuse ou la Lozère).



Les rapports de l'Irdes
[n°587 \(avril 2023\)](#) et [n°592 \(septembre 2024\)](#)



Questions d'économie de la santé - Irdes
n° 276 (Mars 2023) et n° 294 (Novembre 2024)



6. Les personnes sous-main de justice

Les personnes sous main de justice dès leur entrée en détention sont affiliées à un petit régime spécifique du régime général de l'Assurance maladie. Une partie d'entre eux, notamment les personnes sous écrous, est affiliée au CNPE (centre national de gestion de la protection sociale des personnes écrouées) tandis que les personnes en aménagement de peine (bracelet électronique, etc.) sont rattachées à leur Cnam de résidence.

Le tableau ci-dessous issu du guide du CNPE illustre parfaitement les différents cas de figure.

TYPE DE PEINE	ACTIVITÉ PROFESSIONNELLE OU FORMATION RÉMUNÉRÉE OU INDEMNITÉS POLE EMPLOI	CPAM D'AFFILIATION
Incarcération	NON	CNPE
Aménagement de peine	OUI	CPAM DE RÉSIDENCE
Aménagement de peine	NON	CNPE



Pour plus d'information sur le ciblage des personnes sous-main de justice, il est possible de se référer à la fiche disponible sur la documentation collaborative du SNDS. [Détenus | Documentation du SNDS & SNDS OMOP \(health-data-hub.fr\)](#)

La question de la géographie est ici intéressante, le code géographique disponible dans le SNDS correspond-il au lieu de détention ? Ou alors au lieu de vie de la personne sous-main de justice ?

L'analyse des codes géographiques disponibles, aussi bien pour les personnes rattachées au CNPE que les autres, ne permet pas d'apporter une réponse unique. Cela permet néanmoins de comprendre d'où elles proviennent.

Concernant les personnes sous écrous, la géographie dépendrait de la situation administrative de la personne sous main de justice. Si elle ne peut communiquer d'adresse au centre de détention, c'est l'adresse du centre qui figurera alors. Dans d'autres cas, il s'agira du lieu de vie.

Pour les personnes en aménagement de peine et en l'absence d'adresse précise, il s'agira de l'adresse de résidence de ce dernier, nécessaire pour établir les droits en lien avec la Cnam du lieu de résidence.

Dans le PMSI l'adresse est celle du centre de détention (*cf.* guide méthodologique).

7. La géographie de l'aide médicale d'état et des soins urgents

7.1 La géographie de l'aide médicale d'état

L'Aide Médicale de l'État (AME) est destinée à permettre l'accès aux soins des personnes en situation irrégulière en France. Elle est attribuée sous conditions de résidence.

Ce dispositif s'inscrit dans le cadre de la lutte contre les exclusions. Elle s'adresse aux ressortissants étrangers en situation irrégulière et précaire.

Il existe plusieurs façons d'identifier les bénéficiaires de l'aide médicale d'état.

Dans le DCIR, pour identifier les bénéficiaires de l'AME dans la table des prestations du DCIR ER_PRS_F, il faut utiliser la variable RGM_COD (code petit régime).

En considérant les modalités suivantes :

- 95 : aide médicale gratuite 100 % état,
- 96 : aide médicale hospitalière gratuite 100 % état,
- 652 : détenu en situation irrégulière.

Cette information est présente au niveau :

- du référentiel des bénéficiaires IR_BEN_R,
- de la table des prestations.

Il est aussi possible d'utiliser la variable BEN_CMU_CAT (catégorie d'organisme complémentaire) dont la modalité 5 correspond à l'AME.

En outre, joindre la table prestation avec la table affinée ER_ARO_F en posant le filtre ARO_REM_TYP=7 permet de retracer les prestations des bénéficiaires de l'AME. Tous les montants remboursés sont dans la table ER_ARO_F tandis que les quantités d'actes sont dans ER_PRS_F et tables affinées.

Pour rappel également, l'AME relève d'un fond état et non de l'Assurance maladie (non valorisation par l'AM, se retrouve au niveau du motif de non facturation).

Dans le PMSI, il est aussi possible de repérer les séjours et prestations hospitalières pour les bénéficiaires de l'AME grâce à la table VALO qui permet de savoir qu'un séjour est non valorisé pour motif AME (86 % ont un NIR fictif).

7.2 Les soins urgents

Les demandeurs d'asile (hors Mayotte) venant d'arriver sur le territoire peuvent bénéficier de la prise en charge des soins urgents, durant le délai de trois mois au cours duquel ils ne relèvent pas encore de la protection universelle maladie. Les personnes en situation irrégulière non admises à l'AME bénéficient d'une prise en charge des soins urgents à l'hôpital (hospitalisation ou consultation en établissement de santé)¹⁹.

Les soins pris en charge sont :

- les soins dont l'absence mettrait en jeu le pronostic vital ou pourrait conduire à l'altération grave et durable de l'état de santé de la personne ou de celui d'un enfant à naître,
- les soins destinés à éviter la propagation d'une maladie à l'entourage ou à la collectivité (par exemple, la tuberculose),
- tous les soins d'une femme enceinte et d'un nouveau-né (les examens de prévention réalisés pendant et après la grossesse, l'accouchement),
- les interruptions de grossesse (volontaires ou pour motif médical),
- les soins dispensés à des mineurs.

Le repérage des soins urgents dans la partie PMSI du SNDS se fait en utilisant les tables STC avec des variables différentes dans les divers champs.

MCO	=> FAC_SEJ_AM='0' and NON_FAC_AM='4'
HAD	=> FAC_SEJ_AM='0' and PEC_SOI_URG='1'
SSR ou RIP	=> SEJ_FAC_AM='0' and NON_FAC_AM='4'

Le code de la commune du bénéficiaire (issu des données déclaratives du PMSI) est différent du code officiel géographique de l'entité géographique de l'établissement dans 44 % à 50 % des séjours, selon l'année prise en considération. Cette correspondance varie également selon l'établissement prenant en charge les séjours. Cette variation s'étend de 0 % à 98 %. Lorsque le code de la commune du bénéficiaire diffère de celui de l'établissement, c'est le plus souvent un code inconnu, un code Insee de pays étranger ou un code postal non fiable à une commune de résidence qui est reporté dans le résumé standardisé. Les établissements prenant le plus fréquemment des soins urgents sont en Guyane, en région Île-de-France, à Mayotte, en Guadeloupe, dans les Alpes-Maritimes, le Pas-de-Calais, la Marne, la Loire-Atlantique, l'Île-et-Vilaine et le Rhône.

Les soins les plus fréquents (groupe homogène de malades - GHM -) pris en charge sont les séances dont principalement la dialyse, les grossesses pathologiques, les accouchements et affections du post-partum.

8. La notion d'« extra-territorialité »

Cette notion est inspirée du livre blanc de l'AFCCRO.

8.1 Les collectivités d'outre-mer et la Nouvelle-Calédonie

L'organisation de la Sécurité sociale dans les collectivités d'outre-mer et en Nouvelle-Calédonie varie d'un territoire à l'autre. Ainsi :

- Saint-Barthélemy dispose de la Caisse de Prévoyance sociale de Saint-Barthélemy pour l'ensemble de la population (Décret n° 2017-992 du 10 mai 2017). Saint-Martin relève de la Guadeloupe pour la Sécurité sociale ;
- Saint-Pierre-et-Miquelon dispose d'une Caisse de Prévoyance sociale qui gère toutes les branches de la Sécurité sociale selon des règles fixées par les autorités nationales ;

¹⁹ Source : ameli aide-medicale-etat-soins-urgents.

- En Nouvelle-Calédonie, en Polynésie française et dans les îles Wallis-et-Futuna, la Sécurité sociale relève de la compétence de la collectivité, et il existe des caisses placées sous la tutelle de la collectivité qui fixe l'ensemble des règles législatives et réglementaires en la matière.
- Les données de l'hôpital de Saint-Pierre-et-Miquelon sont disponibles dans le PMSI.

8.2 Les soins réalisés à l'étranger

Selon le pays ou le territoire où les soins sont exécutés, il peut exister des conventions avec la Sécurité sociale française ; ces conventions, qui fixent les modalités de prise en charge, varient en fonction du pays ou du territoire. Pour être pris en charge, les soins doivent être inopinés et le séjour à l'étranger ne doit pas excéder six mois. Cependant, certains soins prévus peuvent être pris en charge à la condition d'un accord préalable du service médical de la caisse gestionnaire. À noter que tous les frais engagés ne sont pas remboursables.

Les remboursements des frais peuvent se faire par flux, si la carte européenne d'Assurance maladie a été utilisée, ou après le retour en France, sur demande. Les consommations de soins sont renseignées dans le système d'Assurance maladie sous forme de prestations en nature, dont le libellé permet de distinguer de grandes catégories de prise en charge.

8.3 Les Français résidant à l'étranger

La Caisse des Français à l'Étranger (CFE) est la caisse de Sécurité sociale des Français résidant à l'étranger. Financièrement autonome, la CFE est un organisme privé chargé d'un service public. Elle est régie par le Code de la Sécurité sociale et placée sous la tutelle des ministères en charge de la Sécurité sociale et du budget. Elle ne transmet pas d'information au SNDS.

8.4 Les frontaliers

Le frontalier est une personne qui exerce son activité professionnelle sur le territoire d'un État membre et réside sur le territoire d'un autre État membre, il doit revenir au moins une fois par semaine dans le territoire de sa résidence. Il bénéficie des droits du lieu d'exécution du contrat. Pour ouvrir les droits aux prestations des frais de santé en France, il est nécessaire de s'inscrire auprès de l'organisme de gestion français.

Les remboursements s'effectuent sur la base des tarifs appliqués dans le pays dans lequel les soins ont été dispensés. Des différences de prise en charge existent entre les différents pays frontaliers vis-à-vis des bénéficiaires. Pour la Suisse, un droit d'option doit être émis dans les trois mois suivant le début du contrat de travail pour bénéficier des prestations de frais de santé en France selon la législation et le règlement français.

8.5 Les étrangers séjournant en France

Les étudiants étrangers, ainsi que ceux de Nouvelle-Calédonie, Wallis et Futuna ou les Français nés à l'étranger, venant étudier en France, doivent s'inscrire à la Sécurité sociale française. Il existe quelques situations particulières : résidents monégasques, bénéficiaires de fonctionnaires internationaux, membres de l'Union européenne.

Les expatriés bénéficient de l'Assurance maladie française. Les travailleurs détachés restent affiliés à leur pays d'origine, mais peuvent bénéficier des remboursements de soins en France dépendant du pays d'origine et d'éventuels accords bilatéraux.

9. Les sans domicile fixe

Dans le PMSI le codage des sans domicile fixe (code Z59.0) peut être réalisé s'il répond aux conditions du guide méthodologique. Pour mieux décrire les situations de précarité rencontrées par les patients et susceptibles d'avoir un impact sur leur prise en charge, des consignes d'utilisation de codes existants ont été apportées et des extensions nationales ont été créées en 2015 : Z59.0 « Sans abri ». Ce code s'utilise pour une personne vivant dans la rue au moment de l'admission, ou hébergée dans un centre d'hébergement d'urgence, ou dans un centre d'hébergement et de réinsertion sociale (CHRS).

Pour les personnes sans domicile, il convient de retenir l'information concernant le lieu, la plus précise connue (pour coder le lieu de résidence) et d'enregistrer le code Z59.0 de la Cim 10. Le nombre de résumés (RSA, RHA, RIPNUM, RHAD) contenant un code Cim 10 Z59.0 varie selon le champ et l'année ; il varie entre 4 300 et 5 600 en SMR (ex SSR) entre 2017 et 2023, avec un taux de codes de lieu imprécis variant entre 1 % et 4 %. Depuis 2011, ce nombre augmente continuellement pour le Rim-P, atteignant près de 900 résumés, le taux de codes de lieux imprécis varie entre 9 % et 13 % entre 2016 et 2023. Pour le MCO, le nombre de RSA est supérieur à 20 000 depuis 2018, le taux de codes de lieux imprécis varie entre 3 % et 5 %.

Le recueil du code Z59.0 peut entraîner une augmentation de la sévérité dans le groupage des groupes médico-économiques des SMR, ce qui a pour effet d'améliorer le recueil de l'information et donc une meilleure connaissance de la géographie du bénéficiaire.

La qualité du code de résidence est un des éléments du financement par la qualité des établissements appartenant au champ psychiatrique.

10. La population protégée par la caisse nationale militaire de sécurité sociale

La caisse nationale militaire de sécurité sociale (CNMSS) est le régime auquel sont affiliées les personnes ayant contracté un engagement dans les armées, les personnes bénéficiant d'une pension retraite et quelques autres cas particuliers. Les contrats d'engagement étant plus ou moins longs, la CNMSS a un turnover important de ses affiliés. Il s'agit d'une population jeune, beaucoup plus masculine (70 % d'hommes) que la population totale des affiliés de l'Assurance maladie. La compétence de la CNMSS est nationale, mais elle couvre également, sous conditions, ses affiliés résidant à l'étranger. La CNMSS utilise la norme d'échange commune. La particularité de cette population, est qu'il y a un effet du travailleur sain important. De plus, des soins sont délivrés par le service de santé des armées sans que ceux-ci soient transmis dans le DCIR, ou que le chaînage par le PMSI soit réalisable (non facturable à l'Assurance maladie et code retour du NIR différent de 0).

D'un point de vue « géographique », le militaire change régulièrement de résidence. Il est parfois envoyé pour des périodes de temps variables dans des pays étrangers. Il bénéficie de missions ou de formations avec des déplacements hors résidence, pour des temps longs, parfois plusieurs mois, ce qui entraîne la déclaration d'« adresse de repli ». Il s'agit d'adresses où le militaire ne réside pas habituellement pendant ces périodes. La couverture nationale, le peu de consommation de cette population et la dématérialisation des échanges font que l'adresse déclarée n'est pas toujours la résidence actuelle.

Pour une étude réalisée sur la période 2018 à 2023, il y avait environ 1 million de consommateurs, ayant un âge moyen de 39 ans, avec 69 % d'hommes pour 84 millions de dates de délivrances. Le délai de changement de code officiel géographique entre deux dates pour un même BEN_IDT_ANO était dans 95 % des cas inférieurs à 69 jours. Près de 12 % des bénéficiaires avaient plusieurs ouvriers de droits, il s'agissait principalement d'enfants de moins de 20 ans, mais également de femmes de plus de 60 ans.

La résidence correspondait au code officiel géographique dans 92 % des cas. Le code Insee des pays étrangers a été utilisé pour qualifier les codes départementaux commençant par 99. Les adresses erronées correspondaient le plus souvent à un « code postal ». Sans prendre en compte la qualité du code, pour la période 2018 à 2023, le nombre de résidences différentes variait de 1 à 9. Environ 40 % des personnes ont changé de résidence pendant la période, 33 % avaient un ouvrier de droit unique, 7 % un ouvrier de droit multiple. Lorsqu'il y a changement de résidence, il n'y a pas d'anomalies dans 70 % des cas, c'est-à-dire que les dates de délivrance pour la résidence 1 sont toutes antérieures aux dates de délivrance pour la résidence 2. Dans le cas d'anomalies, il s'agit souvent de périodes de chevauchement inférieures à 70 jours entre la première date de la résidence 2 et la dernière date de la résidence 1.

Particularités liées à l'organisme :

- le code commune du référentiel des bénéficiaires est délibérément codé « 999 »,
- le code 99100 et 99200 sont des codes utilisés par les armées sans possibilité d'identifier la localisation géographique,
- le code 999000 est fréquemment utilisé avec les prestations concernant les indemnités journalières, or celles-ci ne sont pas versées habituellement par la CNMSS,
- le code commune du NIR générique employé pour le remboursement de l'interruption volontaire de grossesse est la commune 83137 (Toulon).

11. Les professionnels de santé

La géographie des professionnels de santé est identifiable à partir de la table DA_PRA_R (pour plus d'information se reporter à la partie 1 de ce document, sur la géographie des professionnels de santé).

Comment utiliser la table DA_PRA_R pour retrouver la localisation des professionnels de santé ?

Pour utiliser la table DA_PRA_R sans doublon, il faut considérer la dernière situation connue de la période étudiée à partir du mois de flux et de l'année.

- Les variables *DTE_ANN_TRT* (année de traitement) et *DTE_MOI_FIN* (mois de traitement) renseignent sur le mois d'intégration des données dans le SNDS.
- La variable *PFS_PFS_NUM* est un identifiant du lieu d'exercice du professionnel de santé²⁰. Elle est unique par professionnel et lieu d'exercice.

La table DA_PRA_R contient donc une ligne par *PFS_PFS_NUM* x *DTE_ANN_TRT* x *DTE_MOI_FIN*.

La variable *STA_PFS_NUM* est un identifiant du cabinet principal du professionnel de santé. Lorsqu'un professionnel de santé libéral exerce ou a exercé dans plusieurs cabinets, plusieurs identifiants de cabinet *PFS_PFS_NUM* peuvent être rattachés à un même identifiant de cabinet principal *STA_PFS_NUM*. Pour un mois de traitement donné (*DTE_ANN_TRT* x *DTE_MOI_FIN*), l'identifiant de cabinet principal est unique par professionnel de santé. En revanche, il n'est pas stable dans le temps : si un professionnel de santé change de cabinet principal, son identifiant de cabinet principal change. L'identifiant de cabinet principal peut être utilisé pour localiser un professionnel de santé à son cabinet principal et éviter les doubles comptes de professionnels exerçant dans plusieurs cabinets.

En pratique, pour utiliser la table *DA_PRA_R*, il est utile de commencer par figer le mois d'intégration des données à partir duquel travailler (variables *DTE_ANN_TRT* et *DTE_MOI_FIN*). Par exemple, les informations les plus récentes peuvent être récupérées (cf. méthode détaillée dans la suite de ce rapport).

Certaines variables, renseignant sur le lieu d'exercice, peuvent changer d'un mois de flux à un autre (c'est le cas par exemple de la variable *EXC_EXC_NAT* qui indique, entre autres, si un cabinet a fermé). Chaque mois,

²⁰ Selon la profession de santé, le lieu d'exercice peut être un cabinet libéral, une officine, un laboratoire...

la table est ainsi mise à jour avec les dernières informations du professionnel de santé qui peuvent n'avoir pas changé ou, au contraire, peuvent s'être modifiées (nature d'exercice, cabinet principal, etc.)

Localisation du lieu d'exercice des professionnels de santé libéraux

Sous réserve de disponibilité relative au statut du profil, les variables de localisation du cabinet contenues dans la table *DA_PRA_R* peuvent être les suivantes :

- *PFS_COD_POS* : le code postal du cabinet ;
- *PFS_EXC_COM* : les trois derniers caractères du code commune du cabinet ;
- *PFS_LIB_COM* : le libellé de la commune du cabinet.

Département

Description de la méthode

1. S'assurer que la variable *PFS_COD_POS* est bien une chaîne de cinq caractères ;
2. Lorsque les trois premiers caractères de *PFS_COD_POS* sont « 200 » ou « 201 », le département est la Corse-du-Sud (« 2A ») ;
3. Lorsque les trois premiers caractères de *PFS_COD_POS* sont « 202 » ou « 206 », le département est la Haute-Corse (« 2B ») ;
4. Lorsque les deux premiers caractères de *PFS_COD_POS* sont « 97 », le code département est constitué des trois premiers caractères de *PFS_COD_POS* (DROM) ;
5. Sinon, le code département est constitué des deux premiers caractères de *PFS_COD_POS*.

Les professionnels de santé installés à Saint-Pierre-et-Miquelon, Saint-Barthélemy et Saint-Martin peuvent être exclus du champ. Pour cela, une méthode peut être d'exclure les cabinets dont le code postal commence par « 97 », mais pas « 974 » (attention à la commune de Saint-Pierre à la Réunion), et dont le libellé de la commune (*PFS_LIB_COM*) semble correspondre à ces territoires : SAINTMARTIN, STMARTIN, SAINTPIERRE, STPIERRE, SAINTBARTHELEMY, STBARTHELEMY... Pour faciliter le traitement des chaînes de caractères, les libellés des communes peuvent être passés en majuscules et les caractères spéciaux et les espaces peuvent être supprimés.

Les cabinets dont le code postal commence par « 98 » ou « 99 » qui ne sont peut-être pas localisés en France peuvent également être exclus.

Pour une vingtaine de communes, le département associé au code postal ne correspond pas au département de la commune. Elles sont énumérées plus haut.

Commune

Le code commune d'un cabinet peut être reconstitué à partir du code postal (*PFS_COD_POS*) et des trois derniers caractères du code commune (*PFS_EXC_COM*).

Description de la méthode

1. Le département associé au code postal est récupéré de la façon décrite dans la section précédente ;
2. Si le code département alors constitué contient trois caractères (DROM), il est concaténé avec les deux derniers caractères du code commune (deux derniers caractères de la variable *PFS_EXC_COM*) ;
3. Sinon, il est concaténé avec les trois derniers caractères du code commune (variable *PFS_EXC_COM*).

Le code commune reconstitué n'est pas forcément correct. Une façon minimale de s'en assurer est de vérifier si ce code commune existe ou a déjà existé.

Pour cela, les codes communes constitués peuvent être confrontés à la liste des codes communes existants dans la géographie la plus récente (communes au 1^{er} janvier de l'année en cours). Comme depuis quelques années il y a davantage de créations de communes nouvelles (fusions de communes) que de rétablissements de communes (séparations de communes), il est souvent plus simple d'utiliser la géographie la plus récente, ou une géographie plus récente que les données utilisées : lorsqu'un ancien code d'une commune qui a fusionné est repéré, il est plus simple de lui attribuer son code commune actuel (celui de la commune résultant de la fusion) que de faire le chemin inverse.

Les codes communes non retrouvés dans la géographie peuvent être corrigés de la façon suivante :

1. Par le libellé de la commune et le département : s'il existe une commune dans le même département ayant le même libellé, son code commune est récupéré ;
2. Par l'historique des changements communaux : si le code commune constitué est un ancien code commune, son nouveau code commune lui est attribué ;
3. Pour les codes communes encore non corrigés, une correction manuelle à partir du code postal et du libellé de la commune peut être appliquée et complétée au fur et à mesure des exploitations, comme sur le dépôt public de la Drees, [ici](#).

Cette méthode permet de récupérer uniquement des codes communes existants dans la géographie souhaitée. Pour autant, cela ne signifie pas que les codes communes récupérés correspondent effectivement aux communes d'exercice des professionnels de santé. Parfois, les codes communes reconstitués ne semblent effectivement pas correspondre aux communes d'exercice. Par exemple, un médecin traitant dont les patients résident à proximité les uns des autres, mais dans des communes éloignées de la commune d'exercice de leur médecin.

Étude de la Drees

Cette méthode de localisation des cabinets des professionnels de santé libéraux a été appliquée dans le cadre de [l'Étude et Résultats 1249](#) portant sur l'activité des médecins généralistes libéraux et salariés en centre de santé en 2020 et 2021. Le code de production de cette étude peut être consulté à [cette adresse](#). Le script ayant servi à constituer les codes communes des cabinets des médecins généralistes libéraux peut être retrouvé [ici](#).

En 2021, une activité de médecine générale libérale est exercée dans 9 362 communes (communes au 1^{er} janvier 2022).

Avant correction des codes communes, 133 des codes communes constitués ne sont pas retrouvés dans les communes existantes au 1^{er} janvier 2022. Cela représente 0,7 % de l'activité de consultation en cabinet, visite à domicile et téléconsultation de médecine générale libérale en 2021.

- Après application des corrections manuelles, 114 codes communes constitués ne sont toujours pas existants.
- Après correction par le libellé de la commune et le département, il reste 94 codes communes non existants au 1^{er} janvier 2022.
- Après correction par l'historique des changements communaux, il ne reste aucun code commune non retrouvé dans les communes au 1^{er} janvier 2022.

PARTIE 4

Quelle géographie pour les indicateurs territorialisés ?

JE RETIENS

Le choix de la géographie dépend de l'objectif de l'étude

La pertinence géographique varie selon le type d'études : d'épidémiologie, de consommation de soins ou d'études sur les parcours. L'étude peut s'appuyer sur des géographies fixes (dernier lieu connu) ou sur celles du moment du soin (dans les tables de prestations). Il convient de garder en tête la spécificité de chaque table pour faire un choix éclairé.

Une étude sur les consommateurs de soin en 2019 par exemple, ne pourra pas s'appuyer sur le référentiel des bénéficiaires qui contient la dernière géographie connue. Le décalage entre le lieu de résidence en 2019 et le lieu de vie actuelle est en effet trop grand et risque de biaiser l'étude.

Le choix de la géographie dépend des données mobilisées

Le choix de la géographie peut se faire de manière pragmatique, si les données mobilisées portent uniquement sur les soins hospitaliers, l'usage de la géographie hospitalière constitue le choix le plus évident.

La géographie entre PMSI et DCIR est de manière générale cohérente. Toutefois, environ 8 % des patients ont des incohérences entre les géographies PMSI et DCIR, souvent liées à l'âge, au régime ou à des délais entre soins.

Toutefois quand les soins portent à la fois sur la ville et sur hôpital, il convient d'intégrer les différences de géographie entre bases du DCIR et PMSI.

Le PMSI utilise les codes postaux tandis que le DCIR se base sur les codes Insee. Pour homogénéiser les données, des traitements de correspondance sont nécessaires (exemple de la table DATASANTE_T_GEO_COM).

1 Quelles études, pour quoi faire ?

Le sous-groupe de travail s'est attaché à réaliser une typologie de la géographie présente dans les études de la dizaine d'organismes qu'il représente.

Ce travail a abouti à une typologie des études suivantes :



Le choix géographique est conditionné à **plusieurs facteurs**.

L'objectif de l'étude

Les études épidémiologiques vont déterminer la zone géographique la plus pertinente pour répondre aux exigences des recherches épidémiologiques. Ces dernières visent à identifier des régions présentant une proximité maximale avec l'exposition à un facteur déterminé, permettant ainsi une analyse précise et fiable des corrélations entre les variables étudiées et les effets observés.

Les études sur la consommation et le recours aux soins nécessitent également une territorialisation rigoureuse, car une localisation erronée des bénéficiaires pourrait entraîner des taux de consommation incohérents ou peu plausibles. De la même façon, une mauvaise affectation géographique de l'offre conduirait à des distances incohérentes, voire très atypiques.

La temporalité de l'étude : plusieurs années ou à un moment précis (instant « t »)

Comme vu dans la partie 1, il existe plusieurs codes géographiques pour un même individu : le code commune Insee sur une période donnée, le dernier code géographique connu à l'instant « t », donc celui de 2025 au 1^{er} janvier 2025 ; on peut également retrouver le code postal (ou regroupement de codes postaux) de résidence au moment d'une hospitalisation ou soins externes à l'hôpital.

Sur une période donnée plus ou moins longue, un bénéficiaire peut avoir plusieurs communes/départements. Se pose alors la question du choix géographique à opérer.

Les études épidémiologiques chercheront à déterminer la géographie la plus proche du phénomène observé alors qu'à contrario les taux de recours au médecin généraliste pourront très bien intégrer les consommations sur les multiples communes sans que cela aboutisse à un choix.

Il existe plusieurs méthodes pour déterminer la géographie à considérer :

- soit sur une date fixe dans les tables de prestations (ER_PRS_F) ;
- soit la géographie du référentiel qui correspond à la dernière situation du bénéficiaire (par exemple 2025 au 6 janvier 2025).

Le type de soin recherché dans l'étude

Suivant le type de soin recherché, le choix géographique peut être différent. La géographie hospitalière repose sur le code postal et non le code commune Insee comme vu dans la partie 1. De ce fait, il peut être intéressant d'utiliser le code postal du PMSI lorsque les études portent sur le champ hospitalier et la géographie du DCIR lorsque le sujet porte plutôt sur les soins de ville.

Néanmoins, les études sur le parcours de soins peuvent rechercher des épisodes de soin aussi bien en ville qu'à l'hôpital. Là aussi, un choix peut être fait, ce point sera développé dans la partie sur les taux de recours à l'imagerie notamment.

2 Comment choisir ?

L'objectif de cette partie est d'aider au choix géographique en proposant **un arbre de décision**.

Cet arbre repose sur quatre questions :

- Est-ce que je travaille sur des données récentes (année en cours) ?
- Est-ce que je travaille sur un historique long ?
- Est-ce que je travaille uniquement sur des données hospitalières ?
- Est-ce que je travaille uniquement sur des données de décès ?

2.1 Données récentes

La première question porte sur la période étudiée : est-ce que l'analyse se concentre sur les données de l'année en cours ?

Si oui, il est préférable de privilégier l'usage de IR_BEN_R, qui contient la dernière situation connue et donc la géographie la plus récente.

2.2 Historique long

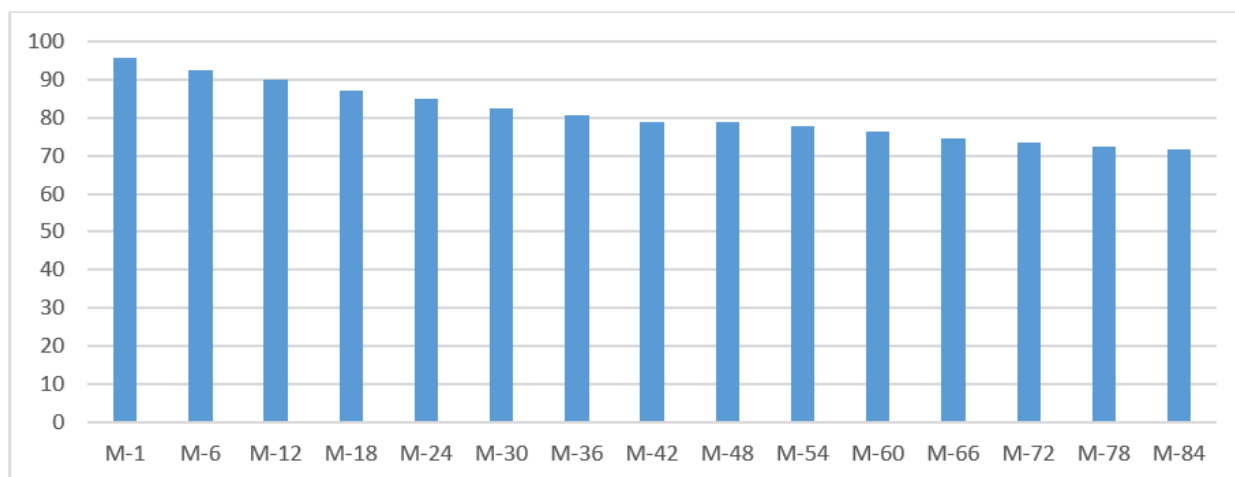
La deuxième question interroge l'étendue temporelle des données : souhaite-t-on travailler sur une série historique longue ?

Afin de se faire une idée du risque à utiliser le référentiel des bénéficiaires pour des périodes d'études éloignées, un travail spécifique est présenté ci-dessous.

L'information géographique des BEN_NIR_PSA||BEN_RNG_GEM (février 2024) a été comparée avec celle issue de la table des prestations (ER_PRS_F), prestations remontées à plusieurs mois d'intervalle (FLX_DIS_DTD).

Le taux de BEN_NIR_PSA||BEN_RNG_GEM associé à une même commune, non manquante, s'élève à 95,6 % dans les prestations remontées de janvier 2024 par rapport au référentiel en février 2024 (cf. graphique ci-dessous). Il chute à 89,8 % sur le flux de janvier 2023 et à 84,9 % sur le flux de janvier 2022. Par rapport aux prestations de janvier 2017, le taux BEN_NIR_PSA||BEN_RNG_GEM ayant une commune non manquante et concordante est de 71,7 %.

Évolution de la part des individus concordants entre le référentiel (février 2024) et les prestations d'un mois de flux dans le temps (avec un code commune non manquant), en pourcentage



Lecture : Dans les remontées de janvier 2024, 95,6 % des individus ont un code commune conforme et identique au référentiel de février 2024, contre 71,7 % des individus des remontées de janvier 2017 (M-84).

Il est à noter que la concordance diffère selon les régimes. Par exemple, en janvier 2024, comparé au référentiel de février 2024, la part de codes commune non manquants et concordants s'élève à 96,3 % pour le régime général, 98,4 % pour la MSA, 99,1 % pour la SNCF, 98,4 % pour la RATP, 92,1 % pour les marines, et 98,7 % pour les mineurs. Les militaires, la MGEN, la MGP et les autres mutuelles ont un taux de code commune manquant dépassant les 70 %²¹.

2.3 Données hospitalières

La troisième question concerne la nature des données de santé mobilisées.

Si l'on travaille uniquement sur des données hospitalières (séjours, admissions, diagnostics...), il est plus simple de s'appuyer sur les données du PMSI. La géographie hospitalière repose principalement sur le code postal et n'est pas construite de la même manière que la géographie centrée sur l'assuré.

De plus, ce code est fourni par le patient et est donc potentiellement plus à jour.

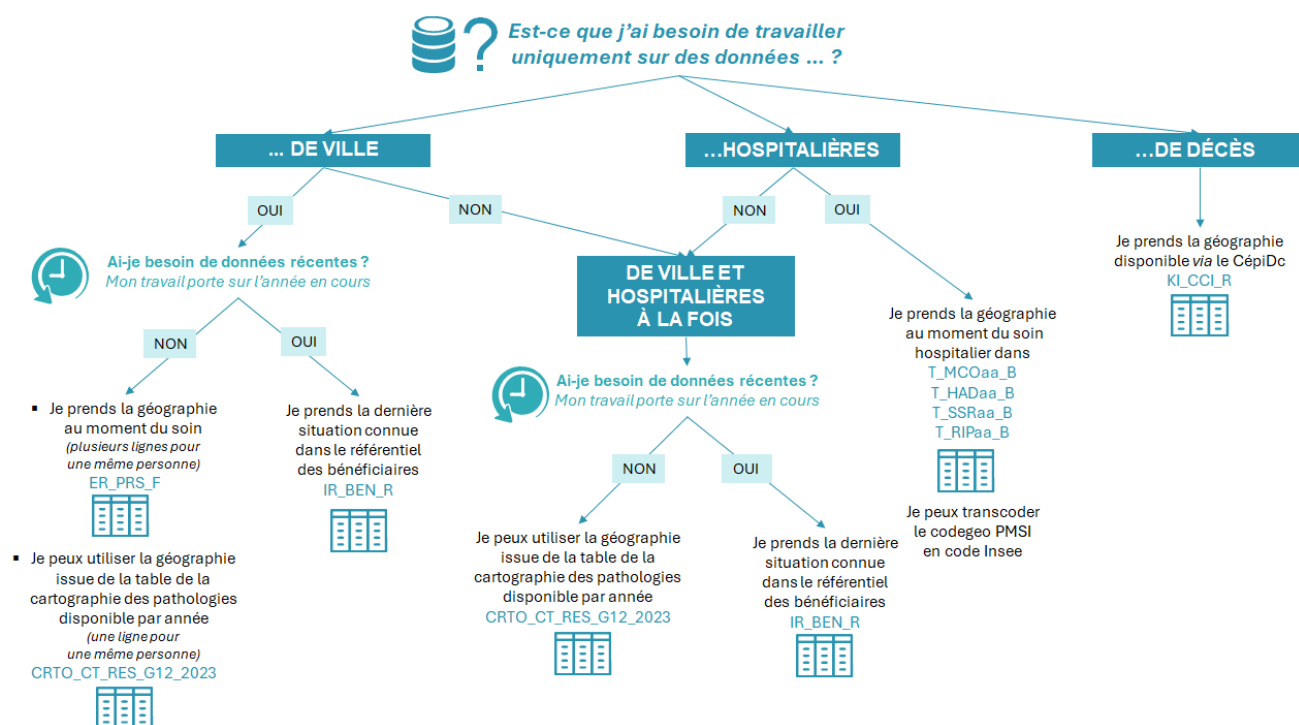
²¹ La commune des militaires n'est pas renseignée dans le référentiel, ce sont les prestations qui sont principalement affectées pour les mutuelles.

2.4 Données de décès

Enfin, si l'analyse repose sur les données de mortalité, il est crucial de travailler avec la commune de décès issue des fichiers CépiDc. Contrairement au code postal, la commune permet une localisation précise et standardisée des décès, indispensable pour des analyses fines à l'échelle locale. Négliger cette information peut entraîner des biais géographiques significatifs, car les données hospitalières ne reflètent pas toujours correctement le lieu de résidence du défunt.

2.5 Arbre de décision permettant de choisir la géographie

Voici une version simplifiée d'un arbre de décision permettant d'aider au choix géographique. D'autres cas de figure auraient pu être ajoutés (analyses nécessitant des données de villes et de décès, hospitalières et de décès...), l'objet est ici d'avoir un exemple relativement simple.

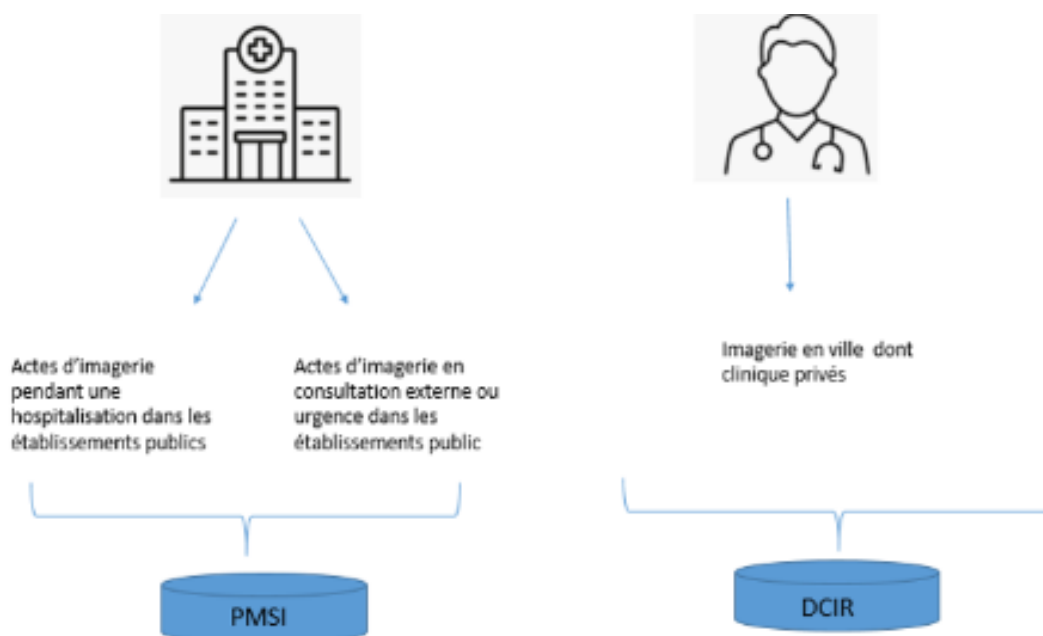


3 La géographie entre ville et hôpital : comment la traiter ?

Certaines thématiques de travail nécessitent d'avoir recours aux données du DCIR et du PMSI. La géographie (comme détaillé dans la première partie du rapport) y est différente, car elle repose sur le code commune Insee pour le DCIR et sur le code postal dans les bases hospitalières. Pour traiter ce sujet, plusieurs solutions existent. La section ci-après illustre une méthode.

3.1 Le passage du code géographique au code commune

La problématique dans le cas présent consiste à homogénéiser la localisation du bénéficiaire sachant qu'une partie des actes est réalisée en ville (dont les cliniques privées), l'autre dans les établissements publics que ce soit au cours d'un séjour ou lors d'un acte ou une consultation externe.



En effet, les bases à mobiliser sont, d'un côté, le DCIR et, de l'autre, le PMSI dans le portail SNDS, qui disposent de niveaux de granularité géographiques différents : le DCIR restitue l'information sur le code Insee de la commune du bénéficiaire et, le PMSI le code géographique PMSI (correspondant soit au code postal ou à une agrégation de codes postaux, sur des années antérieures à 2020).

Par exemple, dans le cas d'une étude portant sur une analyse au niveau des EPCI, il n'y a pas de superposition exacte entre le périmètre d'un code postal et celui des EPCI.

Pour pouvoir fournir des informations à l'EPCI, il va donc s'agir d'estimer une information au niveau de code Insee de la commune quel que soit le lieu de l'acte, puis de la réagréger par EPCI.

Ce type de travail est ici explicité, avec la nécessité d'avoir un traitement différencié selon que l'acte relève du DCIR ou du PMSI.

Traitement dans le PMSI :

Lorsque l'acte est issu du PMSI (séjour ou ACE), on récupère la géographie du patient au moment du soin.

La correspondance entre codes géographiques et codes communes est obtenue par l'intermédiaire de la table DATASANTE_T_GEO_COM de la librairie RFCOMMUN. Cette table contient la correspondance pour les années 2020 à 2023.

Dans les données du PMSI, la dernière hospitalisation de l'année est utilisée.

Ainsi, si un acte est trouvé sur un code PMSI, celui-ci est réparti et proratisé en fonction de la population de chaque code Insee.

Par exemple, si le code PMSI contient trois codes Insee (Code_Insee 1, Code_Insee 2, Code_Insee 3) avec des populations respectives de 1 000, 5 000 et 4 000 habitants, la quantité de l'acte en question se ventilerait de la manière suivante :

Code_Insee 1	0,1
Code_Insee 2	0,5
Code_Insee 3	0,4

Traitement dans le DCIR :

Lorsque l'acte est issu du DCIR (ville ou cliniques privées), on récupère la géographie du patient au moment du soin BEN_RES_DEPT, BEN_RES_COM qui se trouve dans la table ER_PRS_F.

Celle-ci correspond au code Insee de la commune du patient modulo les erreurs potentielles à la liquidation.

L'application de la macro SAS « Macro_Correction_Commune_DCIR_DCIRS.SAS » élaborée par le SCN SI mutualisé des ARS permet d'améliorer la localisation, en s'appuyant notamment sur la table T_FIN_GEOLOC construite par l'ARS Île-de-France. Cette macro réaffecte les codes postaux trouvés aux codes Insee correspondants et permet de mieux réaffecter les codes de la Corse et des DOM.

Fusion des traitements :

Une fois les traitements réalisés dans le PMSI et le DCIR, il est possible de rassembler les deux dans une même table.

En effet, il n'y a pas de redondance puisque :

1. Dans le PMSI ont été sélectionnés :
 - les actes liés à un séjour dans un établissement public uniquement ; condition sur le statut établissement de la table E du PMSI STA_ETA in ('STC', 'DGF'),
 - les actes CCAM des consultations externes du public.
2. Dans le DCIR :
 - exclusion des actes pour information ; DPN_QLF not in (71) et FIDES (IND_TAA NE 1).

Concernant les actes du DCIR, un coefficient 1 est donné puisqu'il est toujours affecté à un seul code Insee après correction.

En termes de résultat, un patient au cours de la même année peut contribuer au calcul du taux de recours sur plusieurs communes/EPCI dès lors qu'il a des actes qui sont retrouvés sur plusieurs communes.

La correspondance entre codes géographiques et codes communes a été obtenue par l'intermédiaire de la table DATASANTE_T_GEO_COM de la librairie RFCOMMUN. Cette table contient la correspondance pour les années 2020 à 2023.

Dans les données du PMSI, la dernière hospitalisation de l'année a été utilisée.

3.2 PMSI ou DCIR quel choix ? l'exemple d'une étude

Étant donné qu'il est possible de chaîner les bénéficiaires entre le DCIR et le PMSI, il est envisageable de réaliser un enrichissement mutuel DCIR/PMSI, avec comme objectif l'utilisation des données du PMSI à la commune et non au code postal, en faisant le lien avec les informations du DCIR, ou à l'inverse, récupérer le code géographique du PMSI lorsque la commune est absente du DCIR.

Ce type d'approche nécessite au préalable de vérifier l'homogénéité entre les codes communes dans le DCIR et les codes géographiques dans le PMSI, afin d'avoir une meilleure connaissance des biais inhérents à une utilisation de ce type.

Dans ce cadre, un travail réalisé au sein du groupe « choix de la géographie dans la territorialisation de la donnée » et présenté ci-dessous a consisté à analyser si la commune mentionnée dans le fichier des prestations du DCIR était compatible avec le code géographique du PMSI. Le but de ce travail n'était pas d'étudier les raisons de ces divergences, mais leur importance et leur répartition suivant les départements et l'âge des bénéficiaires et le régime.

3.2.1 Méthode

Ce travail a été réalisé sur les années 2020 et 2021, pour les bénéficiaires dont la commune était renseignée et ayant eu au moins une hospitalisation au cours de l'année étudiée.

Cette analyse a été réalisée à partir de la table ER_PRS_F, car les données sont d'accès général par rapport à correction_commune.

La correspondance entre codes géographiques et codes communes a été obtenue par l'intermédiaire de la table DATASANTE_T_GEO_COM de la librairie RFCOMMUN. Cette table contient la correspondance pour les années 2020 à 2023.

Dans les données du PMSI, la dernière hospitalisation de l'année a été utilisée.

3.2.2 Résultats

À partir de la table er_prs_f

Sur l'année 2020, sur les 11 567 192 bénéficiaires concernés, 915 161 (7,9 %) ont un code commune qui n'était pas en lien avec le code géographique du PMSI, dont 325 451 (2,8 %) avait un département différent.

Sur l'année 2021, sur les 12 820 045 bénéficiaires concernés, 1 000 441 (7,8 %) ont un code commune qui n'était pas en lien avec le code géographique du PMSI, dont 370 177 (2,9 %) avait un département différent.

Hétérogénéité suivant le délai entre les prestations et l'hospitalisation

Une des raisons pouvant expliquer cette divergence pourrait être un délai plus important entre la date de soin dans le DCIR et l'hospitalisation chez les personnes pour lesquelles il y a hétérogénéité. L'analyse montre cependant que la divergence ne dépend pas directement de l'augmentation du délai (cf. tableau ci-dessous). Elle est importante pour les délais inférieurs à 30 jours, et la plus faible pour les délais entre 31 et 90 jours, avant de réaugmenter. Elle reste néanmoins du même ordre de grandeur, quel que soit le délai.

Pourcentage de divergences en fonction du délai entre l'hospitalisation et l'exécution du soin

	2020	2021
≤ 30 jours	8,6 %	8,3 %
31-90 jours	6,9 %	6,8 %
91-185 jours	7,5 %	7,6 %
≥ 186 jours	8,3 %	8,1 %

Hétérogénéité suivant l'âge

Cette divergence est âge-dépendant (cf. tableau ci-dessous). Les classes d'âge avec une divergence plus importante étant les adolescents et jeunes adultes, ce point pouvant être expliqué par la poursuite d'étude. Cette divergence réaugmente ensuite après 80 ans (lien avec le logement en institution, probablement).

Pourcentage de divergences en fonction de l'âge (PMSI)

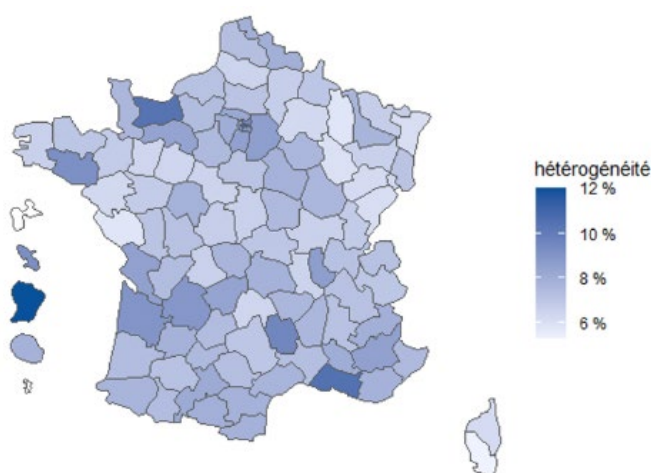
	2020	2021
0-4 ans	8,0 %	8,0 %
5-9 ans	7,4 %	7,2 %
10-14 ans	7,0 %	7 %
15-19 ans	10,1 %	10,0 %
20-24 ans	18,3 %	18,2 %
25-29 ans	15,1 %	15,1 %
30-34 ans	11,1 %	10,8 %
35-39 ans	8,8 %	8,6 %
40-44 ans	7,4 %	7,2 %
45-49 ans	6,7 %	6,5 %
50-54 ans	6,2 %	6,0 %
55-59 ans	6,0 %	5,8 %
60-64 ans	5,4 %	5,5 %
65-69 ans	4,7 %	4,7 %
70-74 ans	4,3 %	4,3 %
75-79 ans	4,6 %	4,6 %
80-84 ans	6,2 %	6,0 %
95-89 ans	9,5 %	9,3 %
90 ans et plus	14,9 %	14,8 %

Hétérogénéité géographique

Cette divergence n'est pas constante suivant les départements. La carte ci-dessous montre le gradient d'hétérogénéité pour les départements français²².

Les divergences vont de 5,3 % à 21,7 %, avec une médiane à 7,4 %. Quatre départements ont une divergence de plus de 10 %.

Gradient d'hétérogénéité entre les codes communes du DCIR et les codes géographiques du PMSI en 2020



²² La Guadeloupe (16,9 %) et Mayotte (21,7 %) ont été exclus de l'échelle, étant donné leurs valeurs respectives.

Hétérogénéité suivant les régimes

Cette divergence dépend des régimes (cf. tableau ci-dessous). La CNMSS a un taux autour de 15 %, avec un effectif assez important (entre 90 et 100 000 personnes).

Pourcentage de divergences en fonction des régimes

	2020	2021
RG	8,1 %	7,9 %
MSA	6,5 %	6,6 %
SLM 506 ²³	5,7 %	5,7 %
CNMSS	14,8 %	15,4 %
SNCF	5,5 %	5,2 %
SLM 603 ²⁴	5,9 %	6,1 %
SLM 537 ²⁵	7,6 %	7,6 %
Mines	8,1 %	7,9 %
CRPCEN	6,8 %	6,9 %
SLM 604 ²⁶	8,3 %	8,3 %
RATP	7,9 %	8,1 %
ENIM	6,6 %	7,3 %
BDF	5,8 %	6,1 %
CAVIMAC	14,5 %	15,0 %
SLM 532 ²⁷	25,5 %	26,3 %
SLM 501 ²⁸	37,2 %	35,1 %
SLM 619 ²⁹	4,8 %	13,7 %
SLM 602 ³⁰	15,1 %	88,8 %

3.2.3 Limites

Ce travail n'a été effectué que pour deux années, en raison de la disponibilité de certaines données. Les résultats sont cependant cohérents sur celles-ci. Les deux années analysées sont des années « Covid », mais il n'y a *à priori*, pas de raison que cette pandémie ait eu un impact sur la non-cohérence des codes communes et codes géographiques. Il pourrait être intéressant de reproduire ce travail pour une année à distance des deux années ici étudiées, afin de confirmer les résultats. Par ailleurs, le même travail a été effectué avec les codes communes `correction_commune_dcir_20xx12` et donnait des résultats similaires.

L'analyse a été réalisée sur les personnes ayant une hospitalisation et une prestation remboursée en ville au cours de la même année, excluant de l'analyse les personnes ayant eu l'un et l'autre assez proches, mais sur deux années différentes. Il n'y a pas de raison d'envisager que ces personnes aient des données différentes des personnes analysées dans cette étude.

3.2.4 Conclusions

Ce croisement montre que, pour environ 8 % des personnes ayant eu une hospitalisation et une prestation, il existe une incohérence entre le code commune du DCIR et le code géographique du PMSI. Cet élément doit être pris en compte lorsque les données du DCIR (code commune), sont utilisées *via* un appariement avec le PMSI.

²³ Mutuelle générale de l'éducation nationale (MGEN).

²⁴ Mutuelle de l'EDF (CMCAS).

²⁵ Mutuelle générale de la police (MGP).

²⁶ Mutuelle du ministère de l'Intérieur (MMI) et mutuelle de la police nationale de Strasbourg (MPN).

²⁷ Mutuelle civile de la guerre y compris mutuelle du ministère des anciens Combattants et victimes de guerre.

²⁸ Mutuelle générale de l'éducation nationale - section extra-métropolitaine (SEM).

²⁹ Mutuelle nationale des hospitaliers (MNH) y compris certaines mutuelles des personnels municipaux.

³⁰ Mutuelle complémentaire de la ville de Paris de l'Assistance publique et des administrations annexes.

3.3 Le suivi des localisations des hospitalisations en tant que lieu de soins

Chaque établissement prenant en charge des patients et transmettant ses données dans le PMSI est tenu de remplir le mode d'entrée et le mode de sortie, ainsi que ce qui est appelé la provenance pour le mode d'entrée ou la destination pour le mode de sortie. Il est donc possible, sous réserve d'avoir un patient chaînable au sens du PMSI (les codes retours du NIR, du sexe et de l'âge ont la modalité « 0 »), de suivre les différents lieux de prise en charge. Cependant les Finess des établissements de prise en charge antérieures ou ultérieures, s'ils sont renseignés dans les dossiers administratifs hospitaliers, ne sont pas transmis dans le résumé de sortie ou le fichier VIDHSOP. Ainsi, il est tout à fait possible d'avoir deux établissements de provenances le même jour. Il faut donc étudier plus finement la trajectoire et prendre comme premier établissement celui qui a déclaré un « domicile » comme mode d'entrée. Le suivant aura des modes d'entrée ou de sortie « 7 » ou éventuellement « 6 ». La qualité de la saisie du Finess de l'établissement de destination ou de provenance n'est pas contrôlée dans les systèmes d'information hospitaliers. Ces numéros Finess font appel à des référentiels locaux.

La provenance d'une structure d'urgence d'un autre établissement est également indiquée, cependant la connaissance du Finess de cette structure, et donc sa localisation géographique, est plus compliquée à identifier puisqu'elle ne transmet pas de résumé de sortie, mais seulement un forfait *via* le DCIR ou le PMSI, selon son statut. Il est donc nécessaire de vérifier l'absence de collision ou de non-chaînage des informations.

La localisation d'un établissement d'hospitalisation à domicile (HAD) correspond à la localisation de la structure administrative ou géographique, mais par définition pas au lieu de soins puisqu'il s'agit d'un domicile, sauf en cas d'accord conventionnel avec un établissement médico-social où l'on peut identifier le lieu de soins.

La qualité des données des établissements dans le PMSI est à prendre en compte, par exemple :

09	MCO	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
09	SSR	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
10	MCO	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
10	SSR	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
11	MCO	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX (PLUS DE MCO DEPUIS LE 01/02/2011)
11	SSR	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX (PLUS DE MCO DEPUIS LE 01/02/20)
12	SSR	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
13	SSR	010780112RHÔNE-ALPES	HÔPITAL LOCAL DU PAYS DE GEX
14	SSR	010780112RHÔNE-ALPES	HL DU PAYS DE GEX
15	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
16	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
17	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
18	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
19	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
20	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
21	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
22	SSR	010780112AUVERGNE-RHÔNE-ALPES	HL DU PAYS DE GEX
23	SSR	010780112AUVERGNE-RHÔNE-ALPES	CH DU PAYS DE GEX

Le code Finess 010780112 donne les informations suivantes sur DATASANTE :

en 2023

010780112	CH DU PAYS DE GEX	CENTRE HOSPITALIER DU PAYS DE GEX
-----------	-------------------	-----------------------------------

en 2015

010780112	CH DU PAYS DE GEX
-----------	-------------------

et en 2009

010780112	HÔPITAL LOCAL DU PAYS DE GEX	HOPITAL LOCAL DU PAYS DE GEX
-----------	------------------------------	------------------------------

Dans le cumul depuis 2009 jusqu'en 2023, des tables T_champaaE il y a 4 176 codes ETA_* distincts, 2 130 de ces établissements ont plus d'un libellé de raison sociale différente pendant la même période.

On trouve par exemple :

590780128	CENTRE DE REEDUCATION FONCTIONNELLE "HELENE BOREL" RAIMBEAUC
590780128	CENTRE DE REEDUCATION FONCTIONNELLE "HELÈNE BOREL" RAIMBEAUC
590780128	CENTRE DE REEDUCATION FONCTIONNELLE HELÈNE BOREL RAIMBEAUCOU
590780128	CENTRE DE REEDUCATION FONCTIONNELLE HELÈNE BOREL RAIMBEAUCOURT
590780128	CRF HELENE BOREL
590780128	CRF HELÈNE BOREL RAIMBEAUCOURT

Depuis 2012, lors d'un séjour dans un établissement en tarification ex-DG, les établissements ayant plusieurs sites géographiques différents peuvent fermer leur séjour en cas de changement de site géographique. Ils indiquent alors dans leur résumé un transfert vers un établissement de court séjour (7.1) cependant cette pratique conforme au guide méthodologique n'est pas toujours respectée.

En 2023, 44 établissements, le plus souvent des CHU, émettent un seul séjour avec plusieurs entités géographiques différentes. L'utilisation de la combinaison des codes d'entrée et de provenance 6 et 1 (mutation d'un court séjour de la même entité géographique) est parfois utilisée (96 séjours en 2023), de même, la combinaison 6.1 pour le mode de sortie et la destination est parfois utilisée (6 séjours en 2023)

Raison sociale	Nombre de RSA comportant plusieurs UM avec un site géographique différent
CHRU DE RENNES	1 479
CHRU DE LILLE	716
CHRU DE NANCY	308
CHR METZ THIONVILLE	277
CHR TOULOUSE	230
CHU ROUEN	230
CHU DE TOURS	216
CHU BREST	209
CH DU HAUT ANJOU	185
CHR ORLEANS	139

Conclusion

L'analyse conduite par le groupe de travail a montré que la géographie n'est pas un simple attribut secondaire de la donnée, mais un élément structurant qui influence directement la qualité des analyses, la pertinence des indicateurs et la portée des résultats.

Ce rapport a permis de mettre en évidence et recenser la diversité des informations géographiques disponibles dans le SNDS, qu'il s'agisse du lieu de résidence, de l'affiliation, du lieu de soins, de naissance, de travail ou de décès. Chacun de ces ancrages apporte une perspective différente, mais soulève également des interrogations méthodologiques quant à leur fiabilité et à leur adéquation avec les objectifs poursuivis. Les travaux présentés ont aussi montré l'importance des traitements correctifs, indispensables pour améliorer la qualité de la donnée, réduire les biais et harmoniser les pratiques, notamment dans les cas particuliers liés aux DOM, à la Corse ou encore aux populations spécifiques.

La réflexion conduite a également souligné la nécessité d'un choix raisonné de la géographie selon la problématique étudiée. Il n'existe pas de solution unique : le rôle du chercheur ou de l'analyste est donc de poser clairement les hypothèses, de connaître les limites des variables mobilisées et de sélectionner le périmètre le plus cohérent avec son objet d'étude.

En offrant un panorama détaillé des sources disponibles, des limites rencontrées et des pistes de correction envisageables, ce rapport se veut un guide méthodologique, conçu pour accompagner les utilisateurs du SNDS dans leurs démarches.

En définitive, ce travail n'apporte pas une réponse unique, mais propose un cadre de réflexion et d'action. Il invite chaque utilisateur à se poser les bonnes questions, à mobiliser les ressources mises à disposition et à partager les pratiques au sein de la communauté.

Enfin, nous espérons que ce rapport sera un outil pratique et inspirant pour l'ensemble des acteurs amenés à travailler avec les données du SNDS. Qu'il s'agisse de chercheurs, d'analystes, d'institutions publiques ou de partenaires de santé, notre souhait est qu'il contribue à faciliter les usages, à renforcer la qualité des travaux et, plus largement, à développer une culture commune autour de la géographie en santé.

Annexes

ANNEXE 1 : LISTE DES CODES DESTINATAIRES DE RÈGLEMENT

DRG_NAT	DRG_NAT_LIB	DRG_TIE_PAY
0	SANS OBJET	SANS OBJET
10	SANS TIERS PAYANT	SANS TIERS PAYANT
31	CORRESPONDANT D ENTREPRISE	AVEC TIERS PAYANT
32	MUTUELLE	AVEC TIERS PAYANT
33	SUBROGATION	AVEC TIERS PAYANT
34	TITRE MEDECIN	AVEC TIERS PAYANT
35	TIERS PAYANT (PART OBLIGATOIRE -RO)	AVEC TIERS PAYANT
36	Tiers Payant Intégral (RO+OC)	AVEC TIERS PAYANT
37	SUBROGATION MULTIPLE	AVEC TIERS PAYANT
99	VALEUR INCONNUE	INCONNU
38	TP RO + OC C2S	AVEC TIERS PAYANT

DRG_AFF_NAT	DRG_NAT	DRG_AFF_LIB
0	0	SANS OBJET
11	10	Assuré
12	10	Adresse Provisoire
13	10	Conjoint divorcé
14	10	Conjoint séparé
15	10	Procuration
16	10	Tutelle
17	10	Veuvage
18	10	conjoint
19	10	Enfant
21	10	Autre ayant droit
22	10	Concubin
23	10	Ascendant, descendant
24	10	Concubin en rupture
25	10	Bénéficiaire Pacsé
26	10	Pacs fin
31	31	Correspondant d'entreprise
32	32	Mutuelle
33	33	Subrogation
34	34	Titre Médecin
35	35	Tiers Payant
36	36	Tiers Payant Intégral (RO+OC)
37	37	Subrogation multiple
99	99	VALEUR INCONNUE
27	10	A-D- RESIDENCE UE
38	38	TP RO + OC C2S

ANNEXE 2 : LES ZONAGES INSEE

Mise à disposition dans le portail SNDS de deux nouvelles tables de valeurs dans ORAVAL :

– **Table IR INS V**

Cette table contient des informations de zonage Insee, obtenues depuis le fichier n_geo de C@rtosanté.

Contenu des 17 variables de cette table :

Nom de variable	Format de la variable	Contenu de la variable
GEO_COM_COD	Char(3)	Code commune
GEO_COM_LIB	Char(50)	Libellé de la commune
INS_ARR_COD	Char(5)	Code arrondissement départemental
INS_ARR_LIB	Char(50)	Libellé de l'arrondissement départemental
INS_CAN_COD	Char(5)	Code canton
INS_CAN_LIB	Char(50)	Libellé du canton
INS_ZEM_COD	Char(5)	Code de la zone d'emploi
INS_ZEM_LIB	Char(50)	Libellé de la zone d'emploi
INS_BVI_COD	Char(5)	Code du bassin de vie
INS_BVI_LIB	Char(50)	Libellé du bassin de vie
EPC_SRN_COD	Char(9)	Code Siren de l'EPCI
EPC_SRN_LIB	Char(70)	Libellé Siren de l'EPCI
SAN_TVI_COD	Char(5)	Code territoire de vie santé
SAN_TVI_LIB	Char(50)	Libellé du territoire de vie santé
CAN_BVI_COD	Char(5)	Code Bassin de vie ou code Canton
CAN_BVI_LIB	Char(50)	Libellé du bassin de vie ou du canton ville
GEO_DPT_COD	Char(3)	Code département = dep du fichier n_geo de C@rtosanté

Cette table sera mise à jour une fois par an, en annulant et remplaçant la précédente.

Table IR ZCN V

Cette table contient des informations de zonage conventionnel, destinées à déterminer si la commune ciblée est située en zone déficitaire, et obtenues depuis le fichier zonage_conv de C@rtosanté.

Contenu des 8 variables de cette table :

Nom de variable	Format de la variable	Contenu de la variable
GEO_COM_COD	Char(3)	Code commune
GEO_DPT_COD	Char(3)	Code département
GEN_ZCN_COD	Char(15)	Code zonage conventionnel généraliste
INF_ZCN_COD	Char(20)	Code zonage conventionnel infirmier
KIN_ZCN_COD	Char(20)	Code zonage conventionnel kinésithérapeute
DEN_ZCN_COD	Char(20)	Code zonage conventionnel dentiste
ORH_ZCN_COD	Char(20)	Code zonage conventionnel orthophoniste
SGF_ZCN_COD	Char(20)	Code zonage conventionnel sage-femme

Cette table sera mise à jour une fois par trimestre, en annulant et remplaçant la précédente.

ANNEXE 3 : EXEMPLE D'UN PROGRAMME DE CORRECTION, FAIT AU SEIN DU SOUS-GROUPES « LA QUALITÉ DE LA GÉOGRAPHIE »

```
DA_PRA_R = spark.sql("""SELECT
    PFS_PFS_NUM AS PS,
    CASE
    WHEN PFS_COD_POS='01200' AND PFS_EXC_COM='109' THEN '74109'
    WHEN PFS_COD_POS='01410' AND PFS_EXC_COM='274' THEN '39274'
    WHEN PFS_COD_POS='01590' AND PFS_EXC_COM='102' THEN '39102'
    WHEN PFS_COD_POS='01590' AND PFS_EXC_COM='283' THEN '39283'
    WHEN PFS_COD_POS='05110' AND PFS_EXC_COM='066' THEN '04066'
    WHEN PFS_COD_POS='05110' AND PFS_EXC_COM='058' THEN '04058'
    WHEN PFS_COD_POS='05130' AND PFS_EXC_COM='150' THEN '04150'
    WHEN PFS_COD_POS='05130' AND PFS_EXC_COM='234' THEN '04234'
    WHEN PFS_COD_POS='05160' AND PFS_EXC_COM='154' THEN '04154'
    WHEN PFS_COD_POS='05700' AND PFS_EXC_COM='374' THEN '26374'
    WHEN PFS_COD_POS='06260' AND PFS_EXC_COM='194' THEN '04194'
    WHEN PFS_COD_POS='06260' AND PFS_EXC_COM='170' THEN '04170'
    WHEN PFS_COD_POS='13780' AND PFS_EXC_COM='105' THEN '83105'
    WHEN PFS_COD_POS='21340' AND PFS_EXC_COM='085' THEN '71085'
    WHEN PFS_COD_POS='33220' AND PFS_EXC_COM='189' THEN '24189'
    WHEN PFS_COD_POS='33220' AND PFS_EXC_COM='335' THEN '24335'
    WHEN PFS_COD_POS='37160' AND PFS_EXC_COM='042' THEN '86042'
    WHEN PFS_COD_POS='42620' AND PFS_EXC_COM='250' THEN '03250'
    WHEN PFS_COD_POS='43450' AND PFS_EXC_COM='105' THEN '15105'
    WHEN PFS_COD_POS='48250' AND PFS_EXC_COM='136' THEN '07136'
    WHEN PFS_COD_POS='52100' AND PFS_EXC_COM='478' THEN '51478'
    WHEN PFS_COD_POS='52100' AND PFS_EXC_COM='522' THEN '51522'
    WHEN PFS_COD_POS='94390' AND PFS_EXC_COM='479' THEN '91479'
    WHEN PFS_COD_POS='98000' AND PFS_EXC_COM='138' THEN '99138'
    ELSE SUBSTR(PFS_COD_POS,1,2)||PFS_EXC_COM
    END AS PS_CC5
FROM DA_PRA_R loc""")
```

- « modif » désigne la correction et « prio » la fiabilité de la correction, servant à les prioriser.

```
ERPRSF.createOrReplaceTempView("ERPRSF_adapted")
ERPRSF = spark.sql("""SELECT ORG_AFF_BEN, PS, ID, BEN_NAI_ANN, BEN_SEX_COD, EXE_SOI_DTD, FLX_DIS_DTD,
RGM_GRG_COD,
    FLX_TRT_DTD, PRS_ORD_NUM, ETB_PRE_FIN, BEN_RES_DPT, BEN_RES_COM,
    INDIV_CC5, PS_CC5, CODGEO_2022, CC_INSEE, CORR15B, prio,
    CASE
    WHEN (ORG_AFF_BEN='01M751501' AND BEN_RES_DPT='99' AND (SUBSTR(BEN_RES_COM,1,1) IN ('6','7','8')))
THEN '1a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='512' AND
    FLX_TRT_DTD<TO_DATE('2014-06-23','YYYY-MM-DD')) AND
    (CC_INSEE IS NOT NULL))
    THEN '2a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='602' AND BEN_RES_DPT='99'
    AND BEN_RES_COM=SUBSTR(PS_CC5,3,3))
```

```

THEN '3a1'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='602' AND BEN_RES_DPT='99'
AND BEN_RES_COM=SUBSTR(PS_CC5,1,3))
THEN '3a2'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='602' AND BEN_RES_DPT='99'
AND SUBSTR(BEN_RES_COM,2,2)=SUBSTR(PS_CC5,1,2))
THEN '3a3'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,9)='602' AND BEN_RES_DPT='99'
AND SUBSTR(BEN_RES_COM,3,1)=SUBSTR(PS_CC5,1,1))
THEN '3a4'
WHEN SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='604' AND (CODGEO_2022 IS NOT
NULL) THEN '4a1'
WHEN SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='604' AND (CODGEO_2022 IS NULL)
THEN '4a2'
WHEN SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='609' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD')
THEN '5a'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='610' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2014-01-01','YYYY-MM-DD'))
THEN '6a'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='612' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2013-02-01','YYYY-MM-DD'))
THEN '7a'
WHEN (ORG_AFF_BEN='01M381613' AND BEN_RES_DPT=SUBSTR(BEN_RES_COM,1,2)
AND BEN_RES_COM=SUBSTR(PS_CC5,3,3) AND
FLX_TRT_DTD<TO_DATE('2015-10-01','YYYY-MM-DD') AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD'))
THEN '8a'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='613' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD'))
THEN '8b'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='614' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2014-01-01','YYYY-MM-DD')
AND INDIV_CC5=PS_CC5)
THEN '9a1'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='614' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2014-01-01','YYYY-MM-DD')
AND INDIV_CC5 <> PS_CC5 AND (CODGEO_2022 IS NOT NULL))
THEN '9a2'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='614' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2014-01-01','YYYY-MM-DD')
AND INDIV_CC5 <> PS_CC5 AND (CODGEO_2022 IS NULL) AND (CC_INSEE IS NOT NULL))

```

```

THEN '9a3'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='617' AND
SUBSTR(BEN_RES_DPT,1,1)='0' AND
SUBSTR(BEN_RES_DPT,2,1)=SUBSTR(PS_CC5,1,1) AND
(SUBSTR(ORG_AFF_BEN,4,2) IN ('04','05','06','13','83','84',
'11','30','34','48','66',
'20','2A','2B',
'14','27','50','61','76',
'02','59','60','62','80',
'75','77','78','91','92','93','94','95') OR
SUBSTR(ORG_AFF_BEN,4,3) IN ('971','972','973'))
THEN '10a1'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='617' AND
SUBSTR(BEN_RES_DPT,1,1)='0' AND
SUBSTR(BEN_RES_DPT,2,1) <> SUBSTR(PS_CC5,1,1) AND
(SUBSTR(ORG_AFF_BEN,4,2) IN ('04','05','06','13','83','84',
'11','30','34','48','66',
'20','2A','2B',
'14','27','50','61','76',
'02','59','60','62','80',
'75','77','78','91','92','93','94','95') OR
SUBSTR(ORG_AFF_BEN,4,3) IN ('971','972','973'))
THEN '10a2'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='617' AND
(SUBSTR(ORG_AFF_BEN,4,2) IN ('24','33','40','47','64',
'19','23','87',
'09','12','31','32','46','65','81','82') OR
SUBSTR(ORG_AFF_BEN,4,3)='974'))
THEN '10b'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='617' AND
SUBSTR(ORG_AFF_BEN,4,2) IN ('08','10','51','52','54','55','57','67','68','88') AND
BEN_RES_COM='395')
THEN '10c'
WHEN (ORG_AFF_BEN='01M371651' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD'))
THEN '11a1'
WHEN (ORG_AFF_BEN='01M861651' AND
FLX_TRT_DTD>TO_DATE('2013-05-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2019-08-01','YYYY-MM-DD'))
THEN '11a2'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='652' AND
SUBSTR(BEN_RES_COM,1,2)='86' AND BEN_RES_COM=SUBSTR(PS_CC5,1,3) AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
FLX_TRT_DTD<TO_DATE('2016-01-01','YYYY-MM-DD'))
THEN '12a'
WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='652' AND
SUBSTR(BEN_RES_COM,1,2)<> '86' AND
FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND

```

```

        FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD'))
        THEN '12b'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01M' AND SUBSTR(ORG_AFF_BEN,7,3)='654' AND
        FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
        FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD'))
        THEN '13a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,7)='01C7590' AND
        FLX_TRT_DTD>TO_DATE('2014-02-28','YYYY-MM-DD') AND
        FLX_TRT_DTD<TO_DATE('2017-10-01','YYYY-MM-DD'))
        THEN '14a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,6)='01C976' AND SUBSTR(INDIV_CC5,1,2)='97' AND
SUBSTR(BEN_RES_COM,1,1)='5')
        THEN '15a'
    WHEN CORR15B=1 AND PFS_EXE_NUM='99999999' THEN '15b'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='01C' AND SUBSTR(ORG_AFF_BEN,7,3)='895' AND
PFS_EXE_NUM='99999999')
        THEN '15c'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='02A' AND SUBSTR(ORG_AFF_BEN,6,1)='2' AND SUBSTR(ORG_AFF_BEN,4,2)
<> '97' AND
        FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
        FLX_TRT_DTD<TO_DATE('2013-07-01','YYYY-MM-DD'))
        THEN '16a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='03A' AND BEN_RES_DPT='20' AND
        FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
        FLX_TRT_DTD<TO_DATE('2020-03-01','YYYY-MM-DD'))
        THEN '17a'
    WHEN (SUBSTR(ORG_AFF_BEN,1,3)='16A' AND
        FLX_TRT_DTD>TO_DATE('2012-12-31','YYYY-MM-DD') AND
        FLX_TRT_DTD<TO_DATE('2018-01-01','YYYY-MM-DD'))
        THEN '19b'
    WHEN SUBSTR(ORG_AFF_BEN,1,3)='01C' AND SUBSTR(ORG_AFF_BEN,7,3)='895' AND PS='99999999' THEN '15c'
    WHEN (BEN_RES_DPT='98' AND BEN_RES_COM='138')
        THEN '20a'
    WHEN (INDIV_CC5 IN ('06194','05374','94479','42250','06170','05234','52478'))
        THEN '20b1'
    WHEN (INDIV_CC5 IN ('05066','05154','33189','01274','21085','01109','05058','05150',
        '48136','43105','33335','01102','01283','52522','37042','13105'))
        THEN '20b2'
    WHEN SUBSTR(ORG_AFF_BEN,1,6)='01C201' AND BEN_RES_DPT='37' THEN 'anomalie_corseA'
    WHEN SUBSTR(ORG_AFF_BEN,1,6)='01C202' AND BEN_RES_DPT='38' THEN 'anomalie_corseB'
    ELSE "modif"
    END AS modif
FROM ERPRSF_adapted""
    .replace("'modif'", "modif").replace("'prio'", "prio")
    .replace("YYYY-MM-DD", "yyyy-MM-dd")
    .replace("99999999", "A01234567")

```

#les replace finaux permettent de modifier à la marge le code écrit sur le portail de la Cnam, pour faciliter le transfert sur le serveur bigdata (SQL sous-jacent)

```
ERPRSF.createOrReplaceTempView("ERPRSF_adapted")
ERPRSF = spark.sql("""SELECT ORG_AFF_BEN, PS, ID, BEN_NAI_ANN, BEN_SEX_COD, EXE_SOI_DTD, FLX_DIS_DTD,
RGM_GRG_COD,
    FLX_TRT_DTD, PRS_ORD_NUM, ETB_PRE_FIN, BEN_RES_DPT, BEN_RES_COM,
    INDIV_CC5, PS_CC5, CODGEO_2022, CC_INSEE, CORR15B, prio,
    CASE
    WHEN ("modif"='\0' AND BEN_RES_COM='\810'
        AND INDIV_CC5=PS_CC5) THEN '\20c1'
    WHEN ("modif"='\0' AND BEN_RES_COM='\810'
        AND INDIV_CC5 <> PS_CC5 AND SUBSTR(INDIV_CC5,1,2)=SUBSTR(PS_CC5,1,2)) THEN '\20c2'
    WHEN ("modif"='\0' AND BEN_RES_COM='\810'
        AND (SUBSTR(INDIV_CC5,1,2) <> SUBSTR(PS_CC5,1,2))) THEN '\20c3'
    ELSE "modif"
    END AS modif
FROM ERPRSF_adapted""")
    .replace("\modif", "modif").replace("\prio", "prio")
    .replace("YYYY-MM-DD", "yyyy-MM-dd")
    .replace("99999999", "A01234567")
```

priorisation

```
ERPRSF.createOrReplaceTempView("ERPRSF_adapted")
ERPRSF = spark.sql("""SELECT ORG_AFF_BEN, PS, ID, BEN_NAI_ANN, BEN_SEX_COD, EXE_SOI_DTD, FLX_DIS_DTD,
RGM_GRG_COD,
    FLX_TRT_DTD, PRS_ORD_NUM, ETB_PRE_FIN, BEN_RES_DPT, BEN_RES_COM,
    INDIV_CC5, PS_CC5, CODGEO_2022, CC_INSEE, CORR15B, modif,
    CASE
    WHEN ("modif"='\1a') THEN 1
    WHEN ("modif"='\2a') THEN 2
    WHEN ("modif"='\3a1') THEN 2
    WHEN ("modif"='\3a2') THEN 3
    WHEN ("modif"='\3a3') THEN 4
    WHEN ("modif"='\3a4') THEN 5
    WHEN ("modif"='\4a1') THEN 2
    WHEN ("modif"='\4a2') THEN 3
    WHEN ("modif"='\5a') THEN 2
    WHEN ("modif"='\6a') THEN 2
    WHEN ("modif"='\7a') THEN 2
    WHEN ("modif"='\8a') THEN 3
    WHEN ("modif"='\8b') THEN 2
    WHEN ("modif"='\9a1') THEN 2
    WHEN ("modif"='\9a2') THEN 3
    WHEN ("modif"='\9a3') THEN 2
    WHEN ("modif"='\10a1') THEN 2
    WHEN ("modif"='\10a2') THEN 3
    WHEN ("modif"='\10b') THEN 2
    WHEN ("modif"='\10c') THEN 1
    WHEN ("modif" IN ('\11a1', '\11a2')) THEN 2
    WHEN ("modif"='\12a') THEN 3
    WHEN ("modif"='\12b') THEN 2
```

```

WHEN ("modif"='\'13a\'') THEN 2
WHEN ("modif"='\'14a\'') THEN 2
WHEN ("modif"='\'15a\'') THEN 1
WHEN ("modif"='\'15b\'') THEN 2
WHEN ("modif"='\'15c\'') THEN 2
WHEN ("modif"='\'16a\'') THEN 2
WHEN ("modif"='\'17a\'') THEN 2
WHEN ("modif"='\'18a\'') THEN 1
WHEN ("modif"='\'19b\'') THEN 2
WHEN ("modif"='\'20a\'') THEN 1
WHEN ("modif"='\'20b1\'') THEN 1
WHEN ("modif"='\'20b2\'') THEN 2
WHEN ("modif"='\'20c1\'') THEN 1
WHEN ("modif"='\'20c2\'') THEN 2
WHEN ("modif"='\'20c3\'') THEN 3
WHEN ("modif" IN ('\'anomalie_corseA\'','\'anomalie_corseB\'')) THEN 2
ELSE "prio"
END AS prio
FROM ERPRSF_adapted""
.replace("\'modif\'","modif").replace("\'prio\'","prio")
.replace("YYYY-MM-DD","yyyy-MM-dd")
.replace("99999999","A01234567")

```

Modifications :

```

ERPRSF.createOrReplaceTempView("ERPRSF_adapted")
ERPRSF = spark.sql("""SELECT ORG_AFF_BEN, PS, ID, BEN_NAI_ANN, BEN_SEX_COD, EXE_SOI_DTD, FLX_DIS_DTD,
RGM_GRG_COD,
    FLX_TRT_DTD, PRS_ORD_NUM, ETB_PRE_FIN,
    INDIV_CC5, PS_CC5, CODGEO_2022, CC_INSEE, CORR15B, modif, prio,
    CASE
    WHEN ("modif"='\'1a\'') THEN \'98\'
    WHEN ("modif" IN ('\'3a1\'','\'3a2\'','\'3a3\'','\'3a4\'','\'8a\'','\'9a1\'','\'10a1\'','\'12a\'') AND (PS_CC5 IS NOT NULL)) THEN
SUBSTR(PS_CC5,1,2)
    WHEN ("modif" IN
('\'2a\'','\'4a2\'','\'5a\'','\'6a\'','\'7a\'','\'8b\'','\'9a3\'','\'10b\'','\'11a1\'','\'11a2\'','\'12b\'','\'13a\'','\'14a\'','\'15b\'','\'15c\'','\'16a\'','\'17a\'','\'19b\'') AND (CC_INSEE IS
NOT NULL)) THEN SUBSTR(CC_INSEE,1,2)
    WHEN ("modif"='\'20a\'') THEN \'99\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'06194\'') THEN \'04\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'05374\'') THEN \'26\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'94479\'') THEN \'91\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'42250\'') THEN \'03\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'06170\'') THEN \'04\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'05234\'') THEN \'04\'
    WHEN ("modif"='\'20b1\' AND INDIV_CC5='\'52478\'') THEN \'51\'
    WHEN ("modif"='\'anomalie_corseA\'') THEN \'20\'
    WHEN ("modif"='\'anomalie_corseB\'') THEN \'20\'
    ELSE BEN_RES_DPT
    END AS BEN_RES_DPT,

```

```

CASE
WHEN ("modif" IN ('3a1','3a2','3a3','3a4','8a','9a1','12a')) AND (PS_CC5 IS NOT NULL) THEN SUBSTR(PS_CC5,3,3)
WHEN ("modif" IN ('15a')) THEN '6'||SUBSTR(BEN_RES_COM,2,2)
WHEN ("modif" IN
('2a','4a2','5a','6a','7a','8b','9a3','10b','11a1','11a2','12b','13a','14a','15b','15c','16a','17a','19b')) AND (CC_INSEE IS
NOT NULL) THEN SUBSTR(CC_INSEE,3,3)
WHEN ("modif" IN ('10c')) THEN '000'
WHEN ("modif"='20b1') THEN SUBSTR(INDIV_CC5,3,3)
ELSE BEN_RES_COM
END AS BEN_RES_COM
FROM ERPRSF_adapted""
.replace("\modif","modif").replace("\prio","prio").replace("\INDIV_CC5","INDIV_CC5").replace("\CC_INSEE","CC_INSEE")
E")
.replace("YYYY-MM-DD","yyyy-MM-dd")
.replace("99999999","A01234567")

ERPRSF = ERPRSF.withColumn("INDIV_CC5_modif",ps_f.concat(ps_f.col("BEN_RES_DPT"),ps_f.col("BEN_RES_COM")))

ERPRSF.createOrReplaceTempView("ERPRSF_adapted")
ERPRSF = spark.sql("""SELECT ORG_AFF_BEN, PS, ID, BEN_NAI_ANN, BEN_SEX_COD, EXE_SOI_DTD, FLX_DIS_DTD,
RGM_GRG_COD,
FLX_TRT_DTD, PRS_ORD_NUM, ETB_PRE_FIN,
INDIV_CC5, PS_CC5, CODGEO_2022, CC_INSEE, CORR15B, modif, prio,
BEN_RES_DPT,BEN_RES_COM, INDIV_CC5_modif,
CASE
WHEN ("INDIV_CC5" <> "INDIV_CC5_modif") THEN 1
WHEN ("modif" IN ('2a','4a2','5a','6a','7a','8b','9a3','10b',
'11a1','11a2','12b','13a','14a','15b','16a','17a','19b')) AND
"INDIV_CC5"="INDIV_CC5_modif" AND "INDIV_CC5"="CC_INSEE") THEN 1
WHEN ("modif" IN ('4a1','9a1','9a2')) THEN 1
ELSE 0
END AS modified
FROM ERPRSF_adapted""
.replace("\modif","modif").replace("\prio","prio").replace("\INDIV_CC5","INDIV_CC5").replace("\CC_INSEE","CC_INSEE")
E")
.replace("YYYY-MM-DD","yyyy-MM-dd")
.replace("99999999","A01234567")
.replace("\INDIV_CC5_modif","INDIV_CC5_modif")

# Agregation
INDIV_GEO = ERPRSF.groupBy(['ID','RGM_GRG_COD','INDIV_CC5','INDIV_CC5_modif','modif','prio']) \
.agg(ps_f.sum(ps_f.lit(1)).alias('nb_rows'),\
ps_f.min('FLX_TRT_DTD').alias('min_flux'),\
ps_f.max('FLX_TRT_DTD').alias('max_flux'),\
ps_f.min('EXE_SOI_DTD').alias('min_soi'),\
ps_f.max('EXE_SOI_DTD').alias('max_soi')) \
.sort("ID","max_flux",ascending=[True,True])

```


ReDSiam

Réseau pour mieux utiliser les Données
du Système national des données de santé

GT MÉTHODOLOGIE

La géographie dans le SNDS

Novembre 2025

Ce rapport,
fruit de nombreuses contributions,
est structuré en quatre grandes parties.

La première aborde
la géographie dans le SNDS,
la seconde la qualité de la géographie,
la troisième la géographie des
populations spécifiques et la dernière
la problématique de la géographie
pour la déclinaison d'indicateurs
territoriaux.