

# RAPPORT ANNUEL

## D'IDENTITO-VIGILANCE 2020

### ANNUAIRE RÉFÉRENTIEL D'IDENTIFICATION DES PATIENTS

---

15/04/2021 – V1.0

## Table des matières

1.	Introduction.....	4
1.1.	Annuaire référentiel d'identification des patients.....	4
1.2.	Chiffres clés 2020 de l'annuaire référentiel d'identification des patients.....	8
1.3.	Événements majeurs en 2020.....	9
1.4.	Prochaines étapes.....	10
2.	Indicateurs d'identito-Vigilance de L'annuaire référentiel d'identification des patients.....	11
2.1.	Contrôles préalables.....	11
2.2.	Contrôles des processus.....	11
2.2.1.	Peuplement des identités.....	12
2.2.2.	Rejets d'identités pour non-conformité.....	14
2.2.3.	Nombre d'identités par état.....	15
2.2.4.	Nombre de rapprochements par état.....	18
2.2.5.	Nombre de rapprochements par domaine d'identification.....	21
2.2.6.	Anomalies sur les identités.....	23
2.3.	Indicateurs d'identito-vigilance secondaires.....	25
3.	Indicateurs d'identito-vigilance des établissements.....	25
3.1.	Rappel des indicateurs de qualité au niveau des établissements de santé.....	25
3.2.	Indicateurs d'anomalies sur les identités au niveau des établissements de santé.....	26

## Liste des figures

Figure 1 - Parcours de soins de Madeleine dit « Maddy » entre divers Professionnels de Santé .....	4
Figure 2 - Exemple de doublon.....	
Figure 3 - Exemple de collision .....	
Figure 4 - Schéma d'un serveur d'identité et de rapprochement de Domaines d'Identification .....	
Figure 5 - Nombre d'injections journalières réelles (MM sur 14j) et moyennes annuelles.....	12
Figure 6 - Ratio nombre d'injections journalières / moyennes journalières sur un an.....	13
Figure 7 - Nombre d'identité rejetée, moyenne et médiane annuelles des identités rejetées.....	14
Figure 8 - Ratio nombre de rejets au quotidien / moyennes journalières des rejets sur un an .....	15
Figure 9 - Évolution nombre d'identités valides dans l'annuaire référentiel d'identification des patients.....	15
Figure 10 - Ratio entre identités valides créées quotidiennement et moyenne sur un an.....	16
Figure 11 - Évolution du nombre d'identités désactivées dans l'annuaire référentiel d'identification des patients.....	17
Figure 12 - Évolution cumulée du nombre d'identités provisoires dans l'annuaire référentiel d'identification des patients.....	17
Figure 13 - Évolution du nombre de rapprochements valides dans l'annuaire référentiel d'identification des patients .....	19
Figure 14 - Évolution du nombre de rapprochements à confirmer dans l'annuaire référentiel d'identification des patients.....	20
Figure 15 - Évolution du nombre de rapprochements provisoires modifiés dans l'annuaire référentiel d'identification des patients.....	20
Figure 16 - Évolution du nombre de rapprochements désactivés dans l'annuaire référentiel d'identification des patients .....	21
Figure 17 - Répartition des rapprochements (effectifs et pourcentages) par Domaine d'Identification ..	21
Figure 18 - Évolution du nombre de doublons potentiels sur le domaine d'identification de l'annuaire référentiel d'identification des patients .....	23

## Liste des tableaux

Tableau 1 - Récapitulatif des contrôles préalables .....	11
Tableau 2 - Répartition des rapprochements par domaine d'identification (vue annuelle cumulée).....	22

## 1. INTRODUCTION

### 1.1. Annuaire référentiel d'identification des patients

Pour assurer une gestion sécurisée et qualitative de l'identification des patients, l'Agence eSanté met à disposition des prestataires de soins de santé un annuaire référentiel d'identification des patients, (Règlement Grand-Ducal du 6 décembre 2019<sup>1</sup>), aussi appelé MPI pour Master Patient Index.

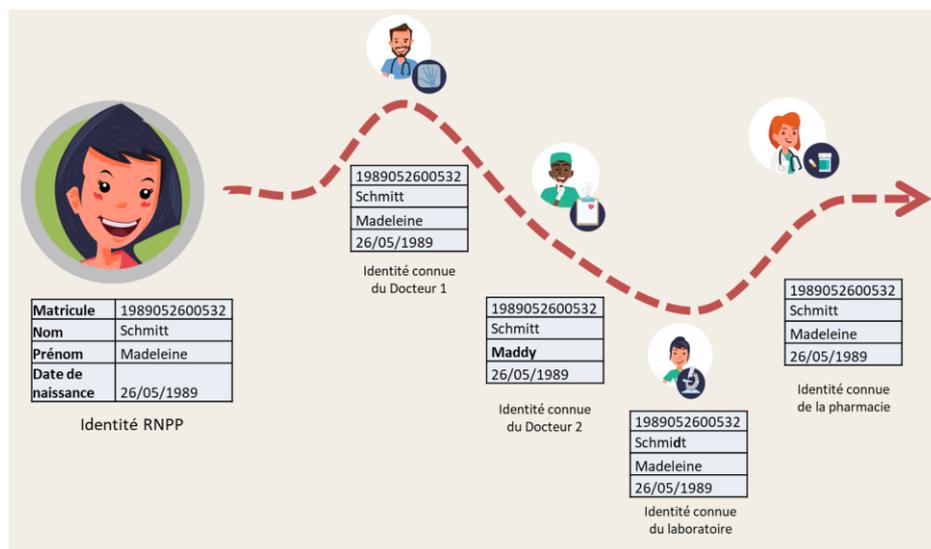


Figure 1 - Parcours de soins de Madeleine dit « Maddy » entre divers Professionnels de Santé

L'annuaire fait partie des services socles de la plateforme eSanté. Son rôle est de lier toutes les identités d'un même patient connues par les prestataires de soins de santé, afin que ces derniers puissent échanger et partager de la documentation médicale sans difficulté.

Les principaux problèmes d'identification sont les doublons et les collisions.

**Cas des doublons :** Les prestataires de soins de santé connaissent une même personne avec des identités différentes. Cela peut conduire à une dispersion de l'information médicale entre les différentes identités, une vue partielle de l'état de santé du patient et de mauvaises décisions médicales car une information ne serait pas disponible. Lorsque ce cas est détecté, une fusion des identités permet de résoudre cette situation.

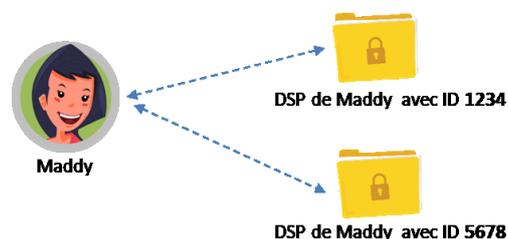


Figure 2 - Exemple de doublon

<sup>1</sup> Règlement grand-ducal du 6 décembre 2019 précisant les modalités de gestion de l'identification des personnes et les catégories de données contenues dans les annuaires référentiels d'identification des patients et des prestataires.

**Cas des collisions :** Deux personnes distinctes sont identifiées avec la même identité. Le risque est alors de collecter sous cette même identité l'information de deux patients. Cela peut conduire à une vue erronée de l'état de santé du patient et de mauvaises décisions médicales. Lorsque ce cas est détecté, il est aussi très difficile de réassigner le bon document au bon patient.

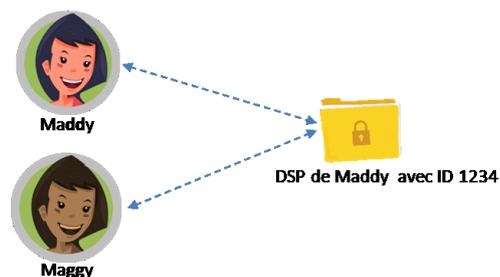


Figure 3 - Exemple de collision

Lorsque l'on partage des informations, les erreurs d'identification et leurs conséquences se propagent aux autres prestataires de soins de santé. Il est donc important d'avoir :

- des règles communes d'identification des personnes dans les échanges électroniques,
- un système de surveillance et de gestion des risques et erreurs liés à l'identification des personnes : l'identito-vigilance.

L'identifiant unique connu de tous est théoriquement la manière la plus simple et la plus sécurisée d'identifier un patient. À cette fin, l'identifiant national (matricule) pourrait être un candidat théoriquement idéal. Mais dans la réalité, une erreur de saisie au niveau du Registre National des Personnes Physiques va parfois générer l'attribution de nouveaux matricules. Une personne peut donc avoir différents matricules tout au long de sa vie. Une erreur de saisie du matricule, au niveau des prestataires de soins de santé peut aussi créer une identité en doublon, ou pire la saisie d'un matricule d'un autre patient va créer une collision d'identité. Tant que la médecine est pratiquée par des hommes, il faut accepter que des erreurs humaines puissent perturber le processus d'identification et donc mettre en place des contre-mesures.

L'annuaire référentiel d'identification des patients n'a donc pas sélectionné un modèle d'identification reposant sur un unique identifiant, mais a retenu une identification du patient qui se base sur un ensemble d'informations composé de :

- 4 traits stricts obligatoires : *Nom de naissance, Prénoms, Date de naissance et Sexe,*
- 2 traits stricts optionnels : *Nom d'usage/marital et Matricule,*
- Et de *l'Identifiant Permanent du Patient local* provenant du *Domaine d'Identification* du prestataire de soins de santé.

Cette méthode a l'avantage de se reposer sur une multiplicité de critères qui réduit drastiquement le risque de collision. L'annuaire référentiel d'identification des patients va alors comparer l'ensemble des traits stricts d'un patient, qu'un prestataire de soins de santé lui soumet, pour déterminer si ce patient est déjà référencé ou s'il doit être ajouté à l'annuaire. Cette comparaison va permettre de lier l'identité soumise par le prestataire à une identité fédératrice qui relie toutes les identités d'un même patient. L'annuaire est donc conçu sur un modèle de fédération hiérarchique de rapprochement d'identités entre divers Domaines d'Identification d'établissements de santé du Luxembourg. Techniquement, c'est un Serveur de Rapprochement d'Identité qui évalue en permanence les relations potentielles entre les identités issues de plusieurs Domaines d'Identité. Pour cela, il utilise un algorithme de rapprochement d'identité qui va traiter tous les messages d'identités provenant d'établissements de santé, puis va automatiquement créer, modifier ou rapprocher des identités. Cet algorithme a été configuré de manière très stricte pour rapprocher les cas certains (correspondance à 100%). Pour les cas où les identités sont proches mais pas identiques, cette automatisation est complétée par des outils d'identito-vigilance qui détectent a posteriori des potentielles identités à rapprocher qui devront être traitées manuellement par les diverses cellules d'identito-vigilance.

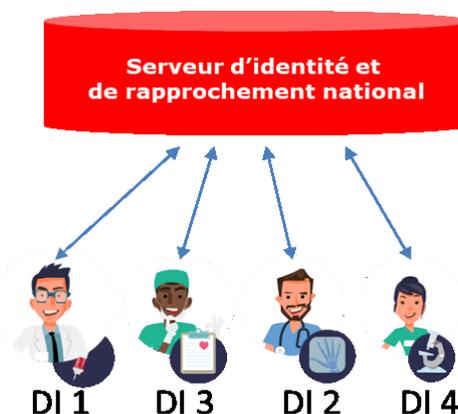


Figure 4 - Schéma d'un serveur d'identité et de rapprochement de Domaines d'Identification

L'annuaire référentiel d'identification des patients est peuplé et mis à jour quotidiennement par des identités officielles provenant du Centre Commun de la Sécurité Sociale et du Registre National des Personnes Physiques. Le périmètre des identités couvertes est donc très large puisqu'il couvre toutes personnes en lien avec une administration ou à la Sécurité Sociale du Luxembourg. Sont alors référencées aussi bien les personnes affiliées et co-affiliées à la Sécurité Sociale luxembourgeoise que des fonctionnaires européens résidant au Luxembourg ou des personnes vivant à l'étranger qui bénéficient de soins de santé planifiés au Luxembourg.

En théorie, les données sont fiables et actuelles si le patient se conforme bien à ses obligations de mise à jour des divers registres alimentant le Registre National des Personnes Physiques. En conséquence, la modification d'identité Fédératrice provenant du Registre National des Personnes Physiques n'est pas autorisée et a été bloquée. Dans les cas où les prestataires de soins de santé rencontreraient des données non actuelles dans l'annuaire, ils seraient donc invités à demander au patient de se rapprocher de leur commune ou du CTIE pour effectuer la mise à jour des données.

Si des établissements de santé souhaitent également gérer des personnes absentes de l'annuaire, comme des touristes qui bénéficieraient pour la première fois du système de santé luxembourgeois, ils peuvent demander à passer d'un profil de simple consommateur de données à un profil de producteur de données d'identification pour pouvoir créer ces identités et les corriger si besoin (tant que la personne n'est pas connue d'un registre officiel).

Les d'établissements de santé utilisent 4 types de messages d'identité pour communiquer avec l'annuaire.

**PDQ:** L'établissement de santé recherche les dernières données démographiques officielles du patient pour vérifier l'identité du patient qu'il prend en charge. Elle soumet une requête à l'annuaire de type Patient Demographics Query (PDQ) avec l'identifiant local du patient, si l'identité du patient est déjà rapprochée, sinon avec son matricule ou/et ses traits stricts. L'annuaire répond avec une liste de patient candidat avec leurs traits démographiques, adresses et les identifiants des divers domaines d'identification référencés dans l'annuaire.

**A04 :** L'établissement de santé informe de l'admission d'un patient dans leur système et souhaite référencer ce patient dans l'annuaire. L'annuaire recherche si une identité Fédératrice existe déjà pour ce patient, si oui il va créer un rapprochement avec l'identité Fédératrice existante, sinon il va créer une nouvelle identité fédératrice et créer un rapprochement avec cette nouvelle identité.

**A08/31 :** L'établissement informe que les données démographiques d'un patient déjà rapproché ont été modifiées dans son système. L'annuaire va évaluer cette modification et décider si elle conserve le rapprochement existant avec l'identité Fédératrice ou si elle le casse.

**A40 :** L'établissement informe qu'il a fusionné deux identités patients et indique quel identifiant conserver. L'annuaire va désactiver le rapprochement avec l'identité non conservée et créer un lien de doublon avéré avec l'identité rapprochée au bon identifiant.

Chaque établissement de santé qui souhaite se connecter à l'annuaire doit passer par un processus de labellisation auquel chacun de ces messages sera scruté pour valider leur syntaxe technique et leur efficience.

Dès lors qu'un établissement devient producteur de données d'identification de l'annuaire, il entre dans une chaîne de coresponsabilité des données d'identification qu'il a renseignées dans l'annuaire. Il doit alors participer à l'effort d'identito-vigilance, coordonné par l'Agence eSanté, en traitant les identités identifiées comme doublons potentiels et les rapprochements d'identité à confirmer avec leur domaine d'identification. À cette fin, un outil de gestion déléguée des identités est mis à disposition des cellules d'identito-vigilance des établissements.

De plus, toute personne qui constate une anomalie dans l'annuaire, comme par exemple, la présence d'identité en doublon, ou une collision de personne sur une même identité, est tenue de la signaler à la cellule d'identito-vigilance nationale de l'Agence eSanté (email : [confidentiel.sante@esante.lu](mailto:confidentiel.sante@esante.lu)). Gardons en tête que l'on travaille en réseau, et que les erreurs d'identification se propagent et peuvent engendrer un peu plus tard des conséquences très graves.

L'identifiant fédérateur, auquel toutes les identités d'un même patient sont rattachées dans l'annuaire, ne doit pas être considéré comme un identifiant unique. Bien que cet identifiant fédérateur soit actuellement disponible dans les réponses des requêtes (PDQ), il ne doit pas être conservé car il est de nature volatile. En effet, il existe un risque que l'identifiant fédérateur soit modifié, notamment après une fusion d'identité. Dès lors, le prestataire qui garderait une copie de l'identifiant disposerait d'une donnée non valide qui pourrait induire un risque de mauvaise identification du patient. Une prochaine évolution devrait d'ailleurs permettre de masquer cet identifiant fédérateur.

Les prestataires et les établissements ne sont pas autorisés à extraire des données de l'annuaire en masse. L'accès légitime se fait au moment d'une interaction avec le patient en question à des fins d'identito-vigilance. D'autres usages tels que des fins commerciales, de marketing ou de relance de factures sont strictement interdits. Par ailleurs, tous les accès à l'annuaire sont tracés et accessibles au citoyen via la plateforme myguichet.lu.

En complément de labellisation technique, les établissements doivent donc respecter un ensemble de règles d'identification des personnes définies dans la charte d'identification et rapprochement d'identité pour la plateforme eSanté. Cette charte est rédigée par la cellule d'identito-vigilance nationale de l'Agence eSanté et validée par le comité national d'identito-vigilance constitué de représentants de la Direction de l'Agence eSanté, du CCSS (Centre Commun de la Sécurité Sociale), de la FHL, du Ministère de la santé et de la sécurité sociale.

## 1.2. Chiffres clés 2020 de l'annuaire référentiel d'identification des patients

L'annuaire référentiel d'identification des patients a été mis en place en juin 2013, nous présentons ci-dessous les principaux indicateurs qui le caractérisent au 31 décembre 2020 (suivi de la variation annuelle):

- 4 hôpitaux généraux connectés avec un profil Producteur d'identité (+1)
  - 2 hôpitaux créent un rapprochement systématique pour chaque patient admis
  - 2 hôpitaux créent un rapprochement seulement en cas de Dossier de Soins Partagé
- 2 470 878 identités fédératrices dont<sup>2</sup>
  - 2 451 211 identités valides (+ 83 933)
  - 495 identités provisoires (+ 225)
  - 13 402 identités désactivées (+ 664)
  - 5 770 identités doublons avérés (+ 1290)
- 3 826 368 rapprochements dont<sup>3</sup>
  - 3 481 160 rapprochements valides (+1 097 722)
  - 2 256 rapprochements à confirmer (+1 135)
  - 327 874 rapprochements provisoires (+15 532)
  - 15078 rapprochements désactivés (+1 008)
- 1,6 rapprochements en moyenne par identité valide et provisoire (+0,2)
- 49 547 liens de doublons potentiels (+ 422) entre deux identités
- 0 identité sans nom de naissance ou sans prénom ou sans sexe ou sans date de naissance
- 2 connexions à l'outil de Gestion Délégué des Identités par les établissements en 2020

---

<sup>2</sup> Voir définitions section 2.2.3 page 16

<sup>3</sup> Voir définitions section 0 page 18

## 1.3. Événements majeurs en 2020

### Réorganisation de la cellule nationale d'identito-vigilance de l'Agence eSanté

Par le passé, la cellule nationale d'identito-vigilance était placée sous la responsabilité d'un médecin au sein du pôle identito-vigilance de l'Agence eSanté. L'expérience nous a appris que les missions du pôle identito-vigilance sont de l'ordre de l'analytique et relèvent que très exceptionnellement de compétences médicales. En effet, le pôle IDV n'accède pas aux données de santé (i.e. contenu document), mais uniquement aux métadonnées du document (i.e. description) et aux données démographiques des patients pour analyser une problématique donnée. Tout accès à chaque dossier ou identité est tracé et doit avoir une finalité. Lors du départ de l'ancien responsable en août 2020, la Direction de l'Agence eSanté a décidé de promouvoir l'ingénieur biomédical, par ailleurs opérant déjà au sein du pôle identito-vigilance, comme responsable du pôle identito-vigilance. Il est supporté par un médecin et un infirmier de l'Agence pour apporter une vue métier et traiter les tâches comme l'ajout de document sur demande patient ou la consultation de document, soit sur demande explicite de patient ou en cas de collision de dossier.

### Généralisation du Dossier de Soins Partagé

L'année 2020 a été marquée par la généralisation du DSP. Un courrier d'activation du DSP a été envoyé à chaque personne titulaire d'un DSP pour les informer de l'ouverture de leur DSP. Cette action a permis de tester les adresses de l'annuaire référentiel d'identification des patients. Au 31 décembre 2020, sur 896 130 DSP créés, 9 149 courriers ont été retournés (1%). Bien que le nombre de retour ait légèrement augmenté en 2021, le taux de retour reste faible (1,5%) et montre une bonne qualité des adresses des patients dans l'annuaire.

Les courriers ont été adressés en juxtaposant nom de naissance et nom d'usage (marital) séparés d'une virgule. Suite à la réception de plaintes de personnes ne comprenant pas l'ajout du nom marital non présent dans le Registre National des Personnes Physiques, une analyse a permis de découvrir que le Centre Commun de la Sécurité Sociale enrichit chaque identité avec le nom du conjoint, s'il est connu dans le nom d'usage. Cette adressage s'est avéré le plus efficace pour le Centre Commun de la Sécurité Sociale notamment pour les envois de courriers à des non-résidents. Le Centre Commun de la Sécurité Sociale n'est pas en mesure de retransmettre les identités, telles que présentes dans le Registre National, mais il garde un indicateur d'adressage « M » pour identifier les personnes qui ne souhaitent pas être adressées avec le nom de leur conjoint. Un correctif a été mis en place le 17 juin 2020 pour que le Centre Commun de la Sécurité Sociale ne transmette plus le nom marital pour les personnes identifiées avec l'indicateur d'adressage « M ». Le premier transfert a permis de corriger 3751 identités.

### Correction d'incohérences

Il a été demandé aux Hôpitaux Robert Schuman de nous soumettre les messages de fusion (A40) de leurs identités locales rapprochées et de ne plus transmettre les identités temporaires ou avec un matricule incohérent.

## Génération des algorithmes externes (scripts R) pour identifier les problèmes des rapprochements d'identité

Afin de faciliter le travail de la cellule nationale d'identito-vigilance, des algorithmes externes (en langage R) à l'annuaire ont été créés pour supporter l'analyse des doublons potentiels détectés et extraits de l'annuaire. En fonction des actions à mener, un fichier Excel est transmis à l'établissement concerné par messagerie sécurisée pour information ou actions. Le script met en avant les différences entre ces doublons et aide à identifier les diverses causes potentielles d'anomalies : usage d'un ancien nom, l'inversion de nom / prénom, fautes d'orthographe, doublon intra-établissement, erreurs dans un des traits...

### 1.4. Prochaines étapes

En 2021, de nouveaux établissements devraient se connecter à l'annuaire référentiel d'identification des patients comme le laboratoire BioneXt ou le CHNP qui réalisent actuellement des tests dans l'environnement d'intégration.

L'usage de l'application de gestion déléguée d'identités devra se généraliser dans les cellules d'identito-vigilance des établissements connectées pour le traitement des anomalies d'identités détectées par la cellule d'identito-vigilance nationale.

Les liens de filiation parent-enfant et le statut de l'affiliation à l'assurance maladie/maternité au Luxembourg devraient pouvoir être référencés dans l'annuaire en 2021.

Il sera étudié la possibilité de revoir le peuplement des genres du Centre Commun de la Sécurité Sociale pour recoller aux valeurs du Registre national (Masculin, Féminin ou Indéterminé).

Les identités sans prénom sont retraitées automatiquement en mettant un « --- » dans le champ prénom pour pouvoir être intégrées dans l'annuaire. Cette règle actuellement est appliquée manuellement par la cellule nationale d'identito-vigilance.

Enfin dans une démarche qualité continue, la cellule d'identito-vigilance nationale ira à la rencontre des cellules d'identito-vigilance des établissements connectées. L'objectif de ces rencontres est d'échanger sur différents sujets, en lien avec l'identito-vigilance, dont entre autres :

- Prendre connaissance des problématiques métiers liées à la gestion des identités
- Présentation d'un processus d'admission (recherche/création d'un patient)
  - service normal
  - service des urgences
- Présentation d'un processus de modification d'identité notamment en lien vers l'annuaire référentiel d'identification des patients
- Présentation d'un processus de gestion des doublons, notamment en lien vers l'annuaire référentiel d'identification des patients
  - fusion d'identité
  - défusion
- Identifier les usages, besoins, attentes et feedbacks des établissements vis-à-vis des services liés à l'annuaire référentiel des patients fournis par l'Agence eSanté

- accès à l'annuaire
- communication des doublons
  - par mail sécurisé
  - par l'application de gestion déléguée d'identités
- Discuter les indicateurs d'identito-vigilance de l'établissement à remonter au rapport d'activité national
- Connaître les actions d'amélioration planifiées, le cas échéant

Suite à ces rencontres, la cellule d'identito-vigilance nationale proposera une nouvelle charte d'identification et de rapprochement de la plateforme eSanté et un référentiel de bonnes pratiques en matière d'identito-vigilance.

## 2. INDICATEURS D'IDENTITO-VIGILANCE DE L'ANNUAIRE RÉFÉRENTIEL D'IDENTIFICATION DES PATIENTS

### 2.1. Contrôles préalables

L'intégration d'identités dans l'annuaire référentiel d'identification des patients ne peut s'effectuer sans contrôle préalable, garant de la qualité des données présentes dans cet annuaire patients, et touchant aussi bien les données proprement dites que les processus visant à leur importation dans l'annuaire. Le Tableau 1 ci-dessous donne un récapitulatif de ces contrôles :

	2020
Existence format d'import des données d'identité (oui/non)	Oui
Existence de procédures de peuplement annuaire (oui/non)	Oui
Existence de procédures de traitements d'anomalies potentielles dans l'annuaire (oui/non)	Oui
Existence de règles de gestion de rapprochements automatiques évolutifs (oui/non)	oui
Complétude des traits stricts (oui/non)	oui
Existence de l'identifiant national (oui/non)	oui
Existence du nom de naissance (oui/non)	oui
Existence du prénom (oui/non)	oui
Renseignement de la date de naissance (oui/non)	oui

Tableau 1 - Récapitulatif des contrôles préalables

### 2.2. Contrôles des processus

Cette section présente les principaux indicateurs d'identito-vigilance relatifs à l'annuaire référentiel d'identification des patients et aborde tout d'abord les flux d'identités (sous-sections 2.2.1 et 2.2.2) qui y sont intégrés, puis le « stock » d'identités qu'il contient, respectivement : statistiques sur les identités (2.2.3), sur les rapprochements (2.2.4 et 2.2.5) et enfin sur les anomalies d'identités (2.2.6).

Les données brutes des différentes figures présentées dans ce rapport sont disponibles sur demande au pôle identito-vigilance de l'Agence eSanté.

## 2.2.1. Peuplement des identités

L'annuaire référentiel d'identification des patients est alimenté par une injection de données quotidienne qui résulte des nouvelles créations et modifications des identités provenant du Centre Commun de la Sécurité Sociale et du Registre National des Personnes Physiques. Tous les jours, l'annuaire reçoit l'injection d'un fichier delta qui met à jour toutes les modifications constatées la veille. Chaque injection contient un certain nombre d'identités qui varie d'un jour à l'autre.

Nous analysons dans cette sous-section la constance de ces injections à savoir le nombre d'identités journalier réellement injecté, leur moyenne journalière de l'année et le ratio de ces deux indicateurs. Les lignes refusées par l'annuaire sont analysées dans la sous-section suivante : les rejets.

La Figure 5 présente ci-dessous le nombre d'identités journalier réellement injecté et leur moyenne journalière de l'année depuis 2015.

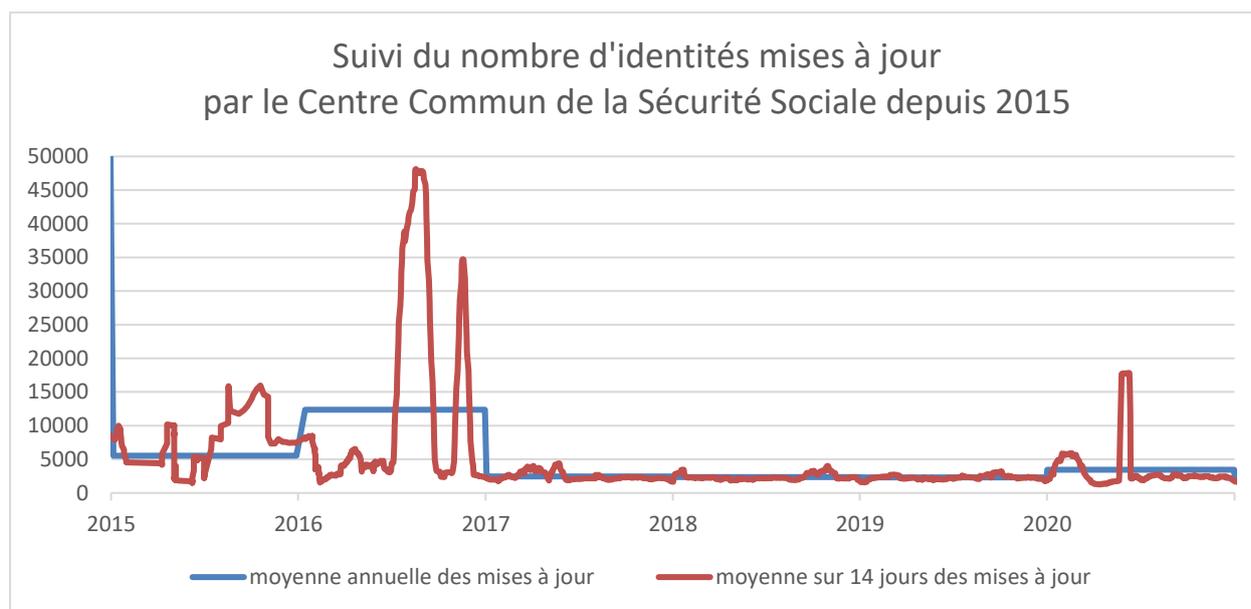


Figure 5 - Nombre d'injections journalières réelles (MM sur 14j) et moyennes annuelles

**Note :** Pour des raisons de lisibilité, l'année 2014 a été exclue du graphique, car l'import initial de 776 378 identités provoquait un aplatissement des courbes non pertinent.

On observe qu'après une première période présentant de très fortes variations, dues essentiellement à divers imports initiaux (du Centre Commun de la Sécurité Sociale, puis de différents établissements), le volume des injections se stabilise sensiblement début 2017 pour s'établir autour de 2770 identités injectées par jour en moyenne. On aperçoit clairement deux pics début 2020 qui correspondent à des mises à jour de nationalité et à un nouveau transfert de certaines identités dont les personnes avaient déjà informé le Centre Commun de la Sécurité Sociale de ne pas utiliser leur nom marital. Ces 2 pics sont également visibles dans la Figure 6 qui présente le ratio entre le nombre d'identités journalier réellement

injecté, et la moyenne journalière de l'année. À l'exception de légers pics ponctuels et des variations début 2020, le ratio s'établit autour de 1, ce qui traduit une relative constance du nombre d'identités injectées dans le temps. Le 25/05/2020, il y a eu un pic à 40 qui correspond à une injection de 140 799 identités mises à jour.

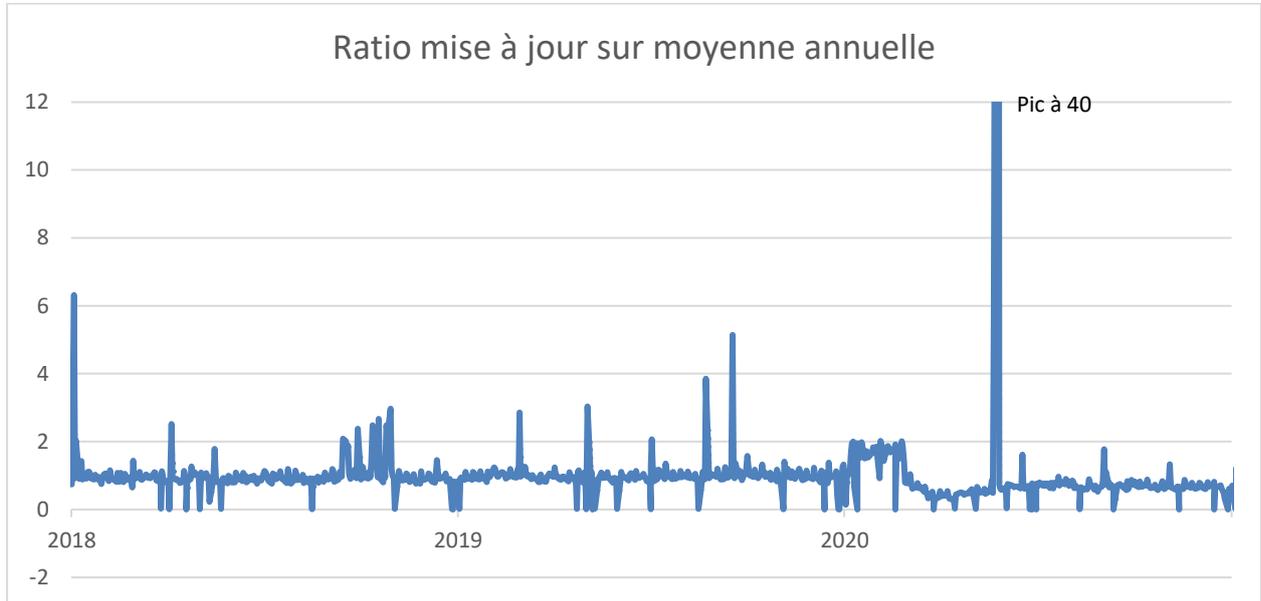


Figure 6 - Ratio nombre d'injections journalières / moyennes journalières sur un an

## 2.2.2. Rejets d'identités pour non-conformité

Chaque identité à injecter doit respecter précisément le format d'import des données d'identité. L'injection des identités rejette systématiquement celles qui ne sont pas conformes. Depuis la mise en place de ce mécanisme, les différentes causes de non-conformité ont été progressivement éliminées, il ne reste plus que l'absence de prénom.

Cette non-conformité provient du fait que dans certaines cultures (notamment en Asie et en Afrique) les personnes n'ont pas de prénom sur leur pièces d'identités ce qui conduit au non-remplissage de ce champ par le Registre National des Personnes Physiques.

Pour résoudre ce problème, la cellule d'identité-vigilance nationale réinjecte manuellement les identités rejetées en ajoutant « --- » dans le champ prénom. Une mise à jour du système est attendue pour automatiser cette tâche.

La Figure 7 ci-dessous présente le nombre de rejets d'identités au quotidien pour non-conformité et la moyenne annuelle de rejets d'identités au quotidien des trois dernières années. En moyenne, 2 identités sont rejetées par jour.

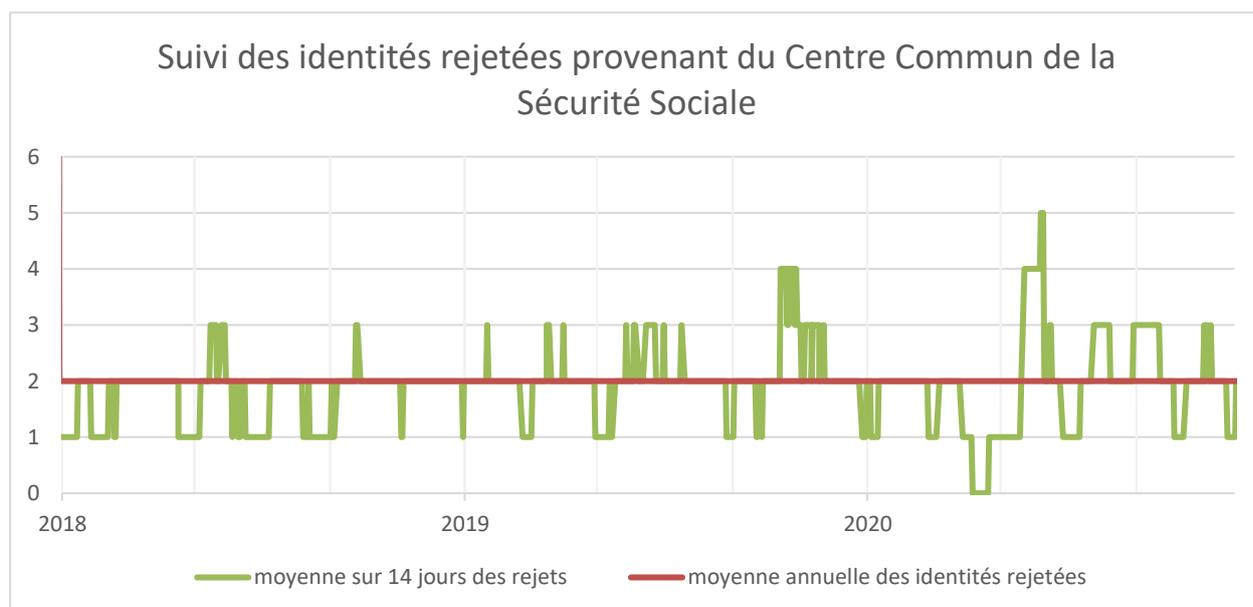


Figure 7 - Nombre d'identité rejetée, moyenne et médiane annuelles des identités rejetées

La Figure 8 présente le ratio entre le nombre d'identités rejetées au quotidien et la moyenne journalière des rejets sur un an. On observe une relative stabilité du ratio depuis 2018, celui-ci s'établit aux alentours de 1. Il y a eu un seul pic qui pourrait paraître alarmant le 25/05/2020 avec un ratio de 25 qui correspond en réalité à 38 rejets sur une injection de 140 799. L'anomalie relève plus de la quantité de données mises à jour ce jour-là que du nombre de rejets.

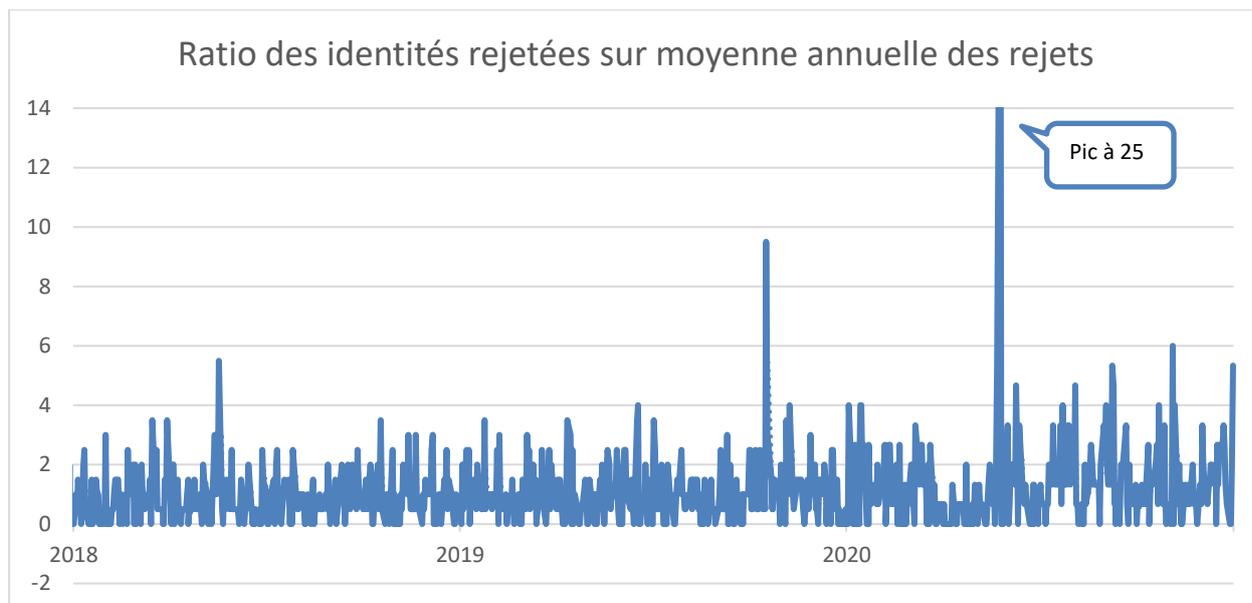


Figure 8 - Ratio nombre de rejets au quotidien / moyennes journalières des rejets sur un an

### 2.2.3. Nombre d'identités par état

Les identités présentes dans l'annuaire peuvent avoir plusieurs états : provisoires, valides, désactivées ou doublons. Ce dernier cas de figure étant abordé plus en détail dans la section 2.2.6, il est donc ici volontairement laissé de côté.

On présente ci-dessous successivement l'évolution du nombre d'identités pour chaque état.

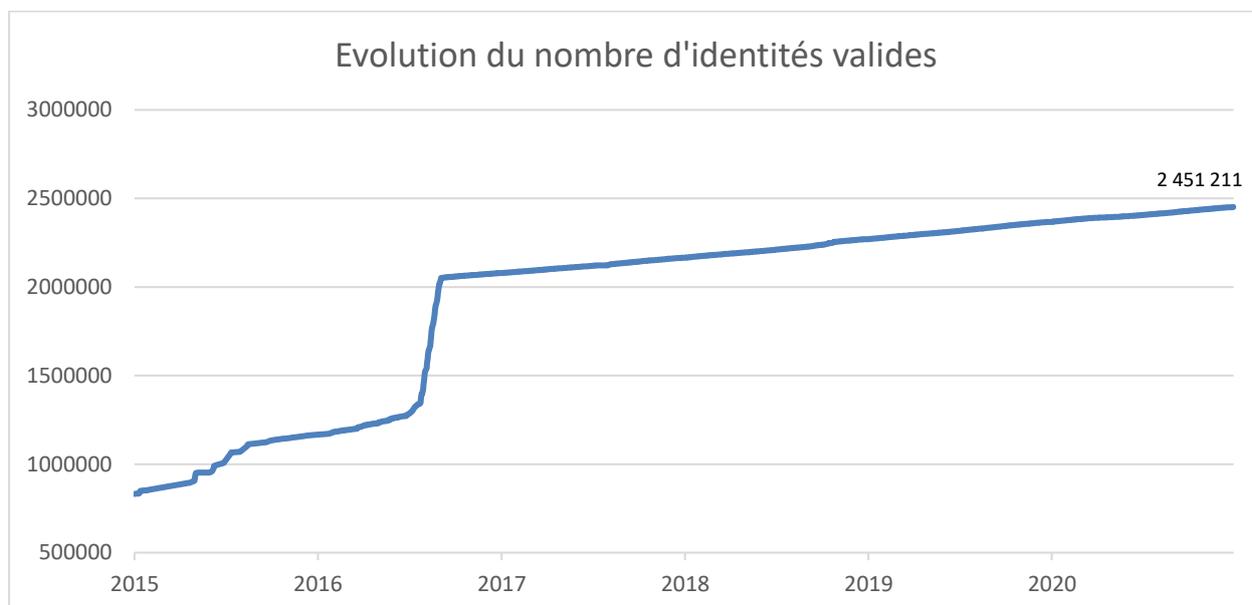


Figure 9 - Évolution nombre d'identités valides dans l'annuaire référentiel d'identification des patients

On remarque un saut dans l'évolution du nombre d'identités valides au cours du 3<sup>ème</sup> trimestre 2016 qui résulte de l'automatisation des imports CCSS puis de la resynchronisation des données d'identités opérée entre le Centre Commun de la Sécurité Sociale (CCSS) et le registre national des personnes physiques

(RNPP). Par la suite, le nombre d'identités valides croît de manière constante. Au 31 décembre 2020, le nombre d'identités valides présentes dans l'annuaire est de 2 451 211.

En 2020, on compte une augmentation de 83 933 identités valides, ce chiffre devrait correspondre au nombre de nouvelles naissances (enfants de résidents (6 459 en 2020<sup>4</sup>) ou de travailleurs frontaliers), plus le flux migratoire des nouveaux arrivants (résidents (22 490 en 2020) ou travailleurs frontaliers), plus d'éventuels doublons non détectés. Par ailleurs, le CCSS nous a également indiqué avoir transmis à l'annuaire 82 718 nouvelles identités en 2020. La différence de 1 215 identités provient certainement des identités créées par les établissements et les éventuels doublons. Une analyse plus fine est en cours pour valider ce postulat.

La constance de la courbe est en réalité assez surprenante, cela indiquerait que l'association de tous ces paramètres est constante. Ce qui est important de remarquer, est qu'il n'y a plus d'envolée du nombre d'identités suite à la connexion progressive de 3 hôpitaux entre 2016 et fin 2020. La courbe indiquerait donc un fonctionnement correct de l'algorithme de rapprochement avec potentiellement une erreur de doublon non détecté fidèle dans le temps.

Par ailleurs, la Figure 10 montre un ratio relativement stable proche de 1 entre le nombre d'identités valides créées quotidiennement et la moyenne annuelle des identités valides créées quotidiennement.

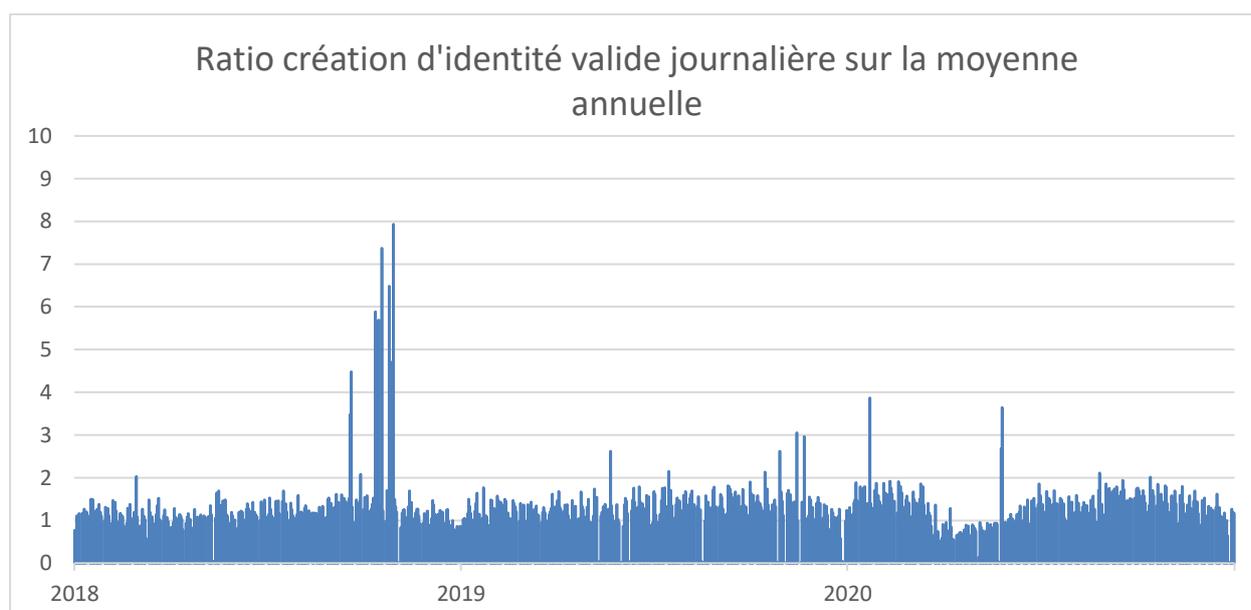


Figure 10 - Ratio entre identités valides créées quotidiennement et moyenne sur un an

Les identités désactivées correspondent soit à des identités traitées comme doublons avérés, ou à des identités temporaires soumis par un établissement. L'évolution cumulée du nombre d'identités désactivées (Figure 11) suit assez logiquement une dynamique similaire à celle du nombre d'identités valides, même si l'ordre de grandeur est bien moindre : ainsi, au 31 décembre 2020, le nombre d'identités désactivées s'établit à 13 402.

<sup>4</sup> <https://statistiques.public.lu/catalogue-publications/en-chiffres/2021/010421-demographie-en-chiffres.pdf>

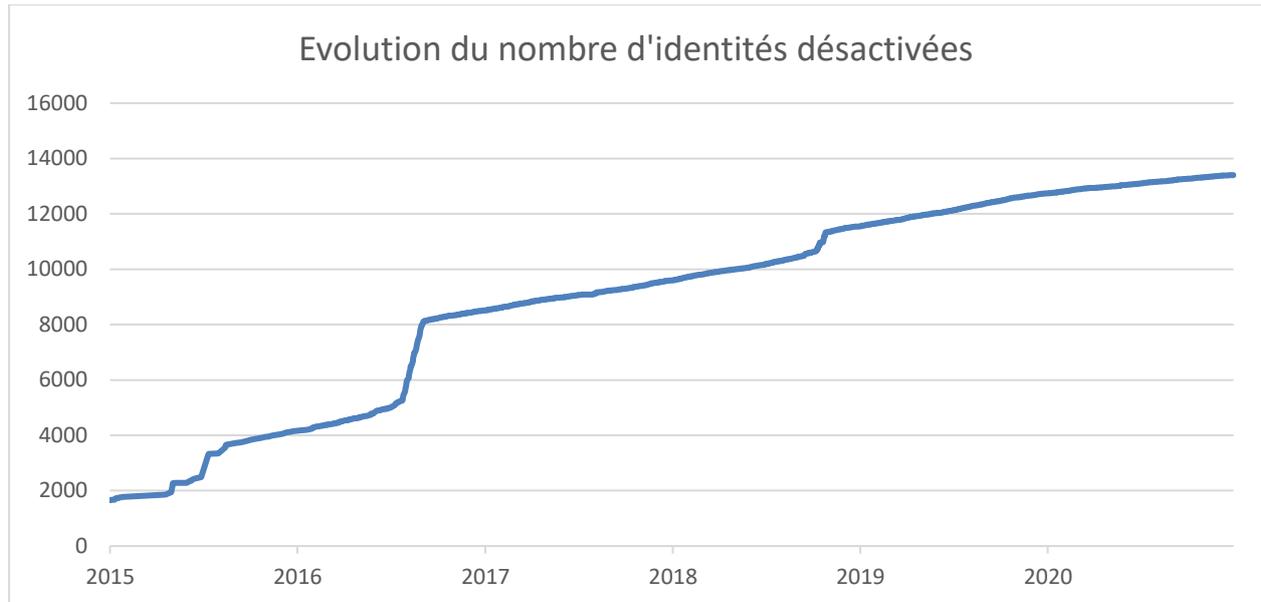


Figure 11 - Évolution du nombre d'identités désactivées dans l'annuaire référentiel d'identification des patients

Les identités provisoires correspondent à des cas assez spécifiques lorsque des informations relatives aux traits stricts ont été modifiées, notamment suite à des cassages de rapprochements ou des fusions. Ce sont des identités à valider. Leur volume est relativement faible en comparaison au volume des identités valides à terme, finissent le plus souvent par passer à l'état valide. C'est la raison pour laquelle on observe ici des effectifs particulièrement faibles (495 identités seulement au 31 décembre 2020).

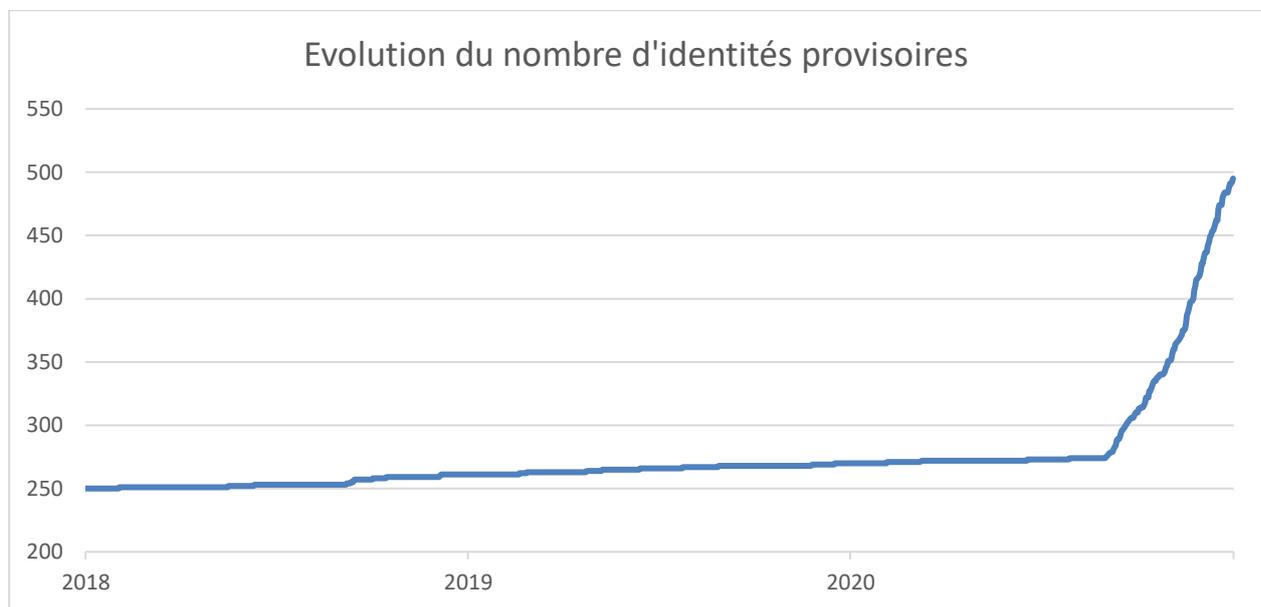


Figure 12 - Évolution cumulée du nombre d'identités provisoires dans l'annuaire référentiel d'identification des patients

## 2.2.4. Nombre de rapprochements par état

L'annuaire référentiel d'identification des patients de la Plateforme eSanté fédère les Domaines d'Identification de plusieurs établissements; on retrouve donc dans l'annuaire de nombreuses identités possédant des rapprochements sur plusieurs de ces DI. A l'instar des identités, les rapprochements peuvent également être qualifiés avec un état : *Valide*, *Provisoire – Modifié*, à *Confirmer* et *Désactivé*.

Lorsqu'une identité locale est rapprochée pour la première fois (taux de correspondance à 100%), le rapprochement de l'identité locale avec l'identité Fédérateur est à l'état *Valide*. Si le taux est inférieur à 100%, une nouvelle identité fédérale est créée et rapprochée à l'identité locale avec un état de rapprochement *Valide*.

Lorsque les traits d'une identité locale déjà rapprochée sont modifiés, l'état du rapprochement peut être également modifié pour signaler aux cellules IDV de contrôler ce rapprochement. Les états de rapprochement sont donc à considérer comme des indicateurs avec différents niveaux d'alertes :

- Si le taux de correspondance reste supérieur à 70%, on ne modifie pas l'état du rapprochement, donc le plus souvent ce rapprochement certain reste dans l'état de rapprochement *Valide*.
- Si le taux de correspondance passe entre 50 et 70%, on modifie l'état du rapprochement en *Provisoire – Modifié* pour indiquer que les traits d'identités ont été modifiés de manière sensible depuis sa création originelle.
- Si le taux de correspondance passe sous les 50%, le rapprochement est *Désactivé*, et une nouvelle identité fédérateur est créée.
- Enfin si suite à la modification des traits, plusieurs identités candidates correspondent à 100%, le rapprochement nécessite une confirmation de la part de la cellule IDV, ce qui est signalé par un état de rapprochement *A confirmer*.

Hormis l'état *Désactivé*, tous les rapprochements sont utilisables par les applications clientes peu importe l'état.

Le total des rapprochements (hors désactivés) est de 3 811 290 au 31 décembre 2020. Les ordres de grandeur des différents états étant très différents, les évolutions sont présentées séparément.

On s'intéresse tout d'abord aux rapprochements valides, qui sont de loin les plus nombreux. Au 31 décembre 2020, l'annuaire compte 3 481 160 rapprochements valides.

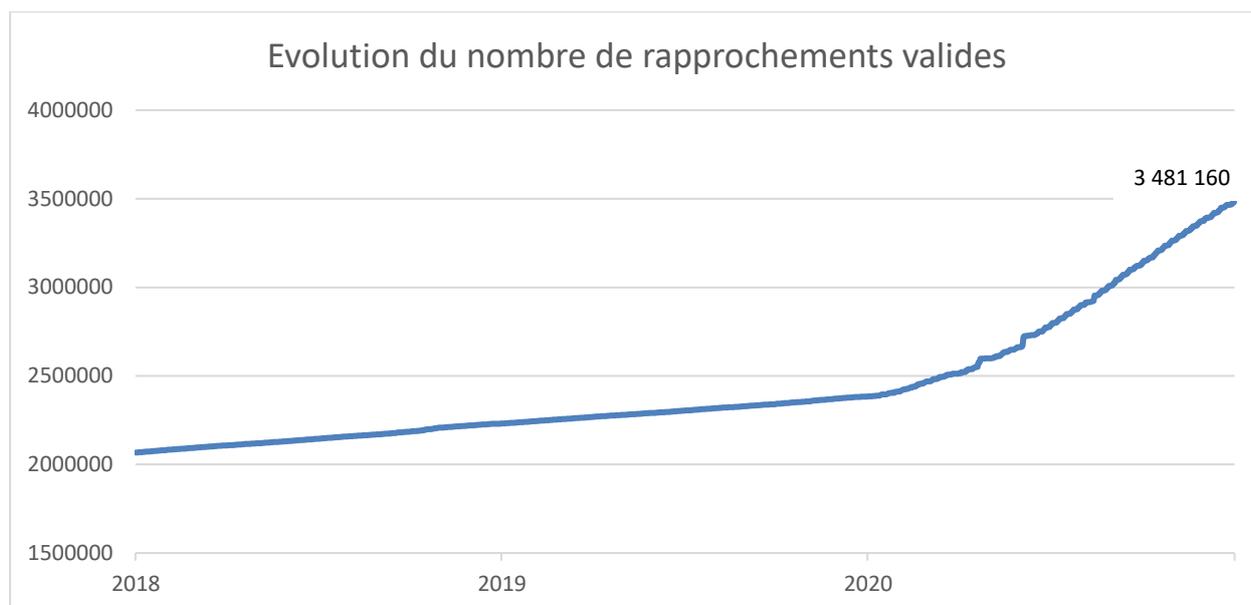


Figure 13 - Évolution du nombre de rapprochements valides dans l'annuaire référentiel d'identification des patients

Avant 2020, nous avons une dynamique stable proche de celle du nombre d'identités en Figure 9, à cette époque les principaux rapprochements provenaient du Centre Commun de la Sécurité Sociale et d'un établissement hospitalier. En 2020, la généralisation du Dossier de Soins Partagés et la connexion de 2 établissements hospitaliers supplémentaires à l'annuaire ont boosté la création de rapprochements.

Il est intéressant de comparer les évolutions du nombre d'identités (Figure 9) et du nombre de rapprochements (Figure 13). La différence de dynamique en 2020 indique que l'annuaire rapproche correctement les identités, puisque le nombre d'identités poursuit une dynamique stable vis-à-vis des années précédentes avec 83 933 identités supplémentaires, tandis que le nombre de rapprochements a augmenté de 1 097 722. Le rapport de ces deux indicateurs est saisissant.

Le nombre de rapprochements à confirmer a également doublé en 2020, mais dans une moindre mesure passant à 2 256 au 31 décembre 2020. Idéalement, cet indicateur devrait être à zéro, la cellule d'identification nationale devrait revoir, et le cas échéant valider l'ensemble de ces rapprochements.

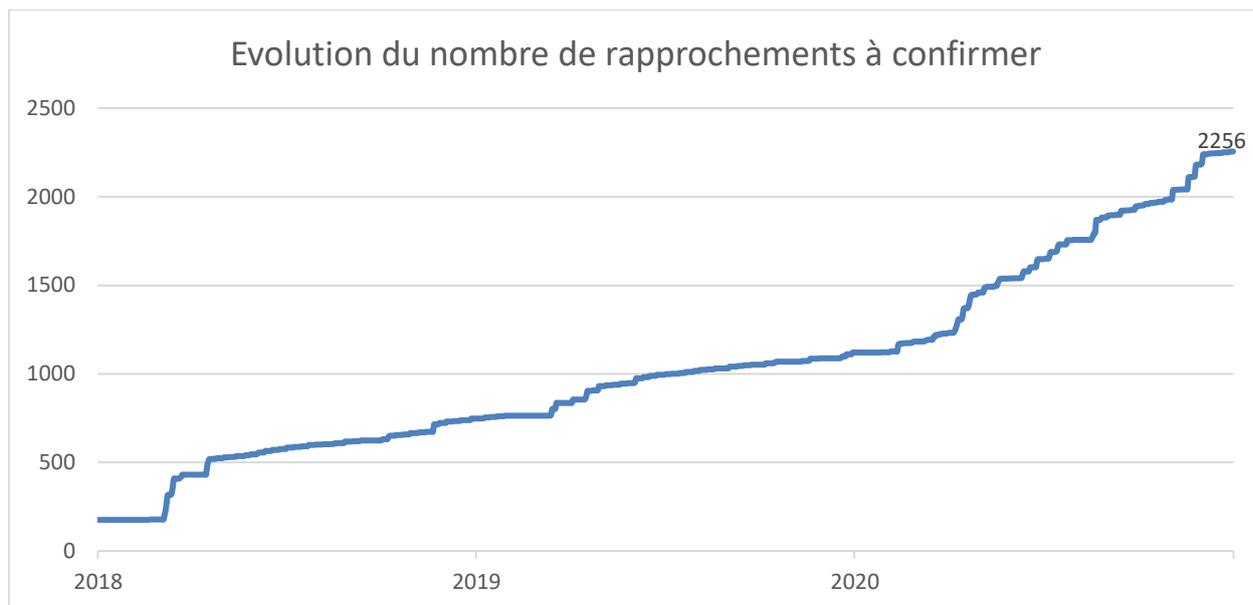


Figure 14 - Évolution du nombre de rapprochements à confirmer dans l'annuaire référentiel d'identification des patients

Les rapprochements *provisoires modifiés* constituent la seconde catégorie la plus importante en termes d'effectifs (327 874 au 31 décembre 2020), mais leur ordre de grandeur reste 10 fois inférieur à celui des rapprochements valides. Leur nombre croît assez régulièrement au cours de la période observée.

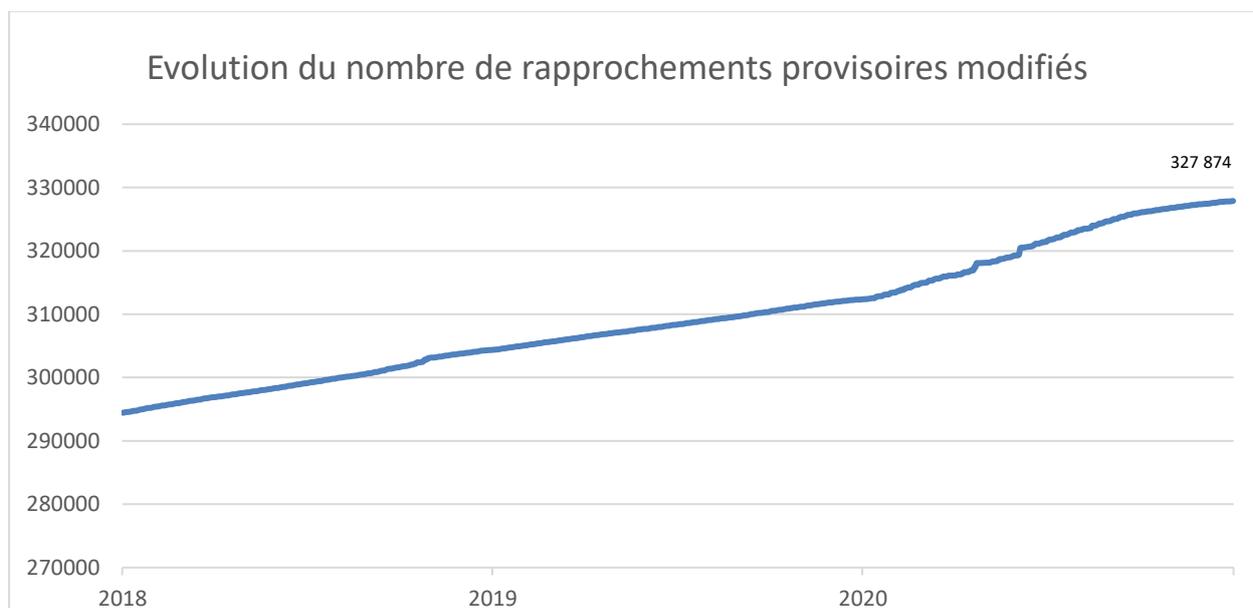


Figure 15 - Évolution du nombre de rapprochements provisoires modifiés dans l'annuaire référentiel d'identification des patients

Les rapprochements à l'état *désactivé* sont relativement peu nombreux (15 078 au 31 décembre 2020) au regard du nombre total de rapprochements. Comme dans le cas des identités, on retrouve ici une dynamique comparable à celle des rapprochements valides, même si l'ordre de grandeur est bien moindre, par un facteur de plus de 200.

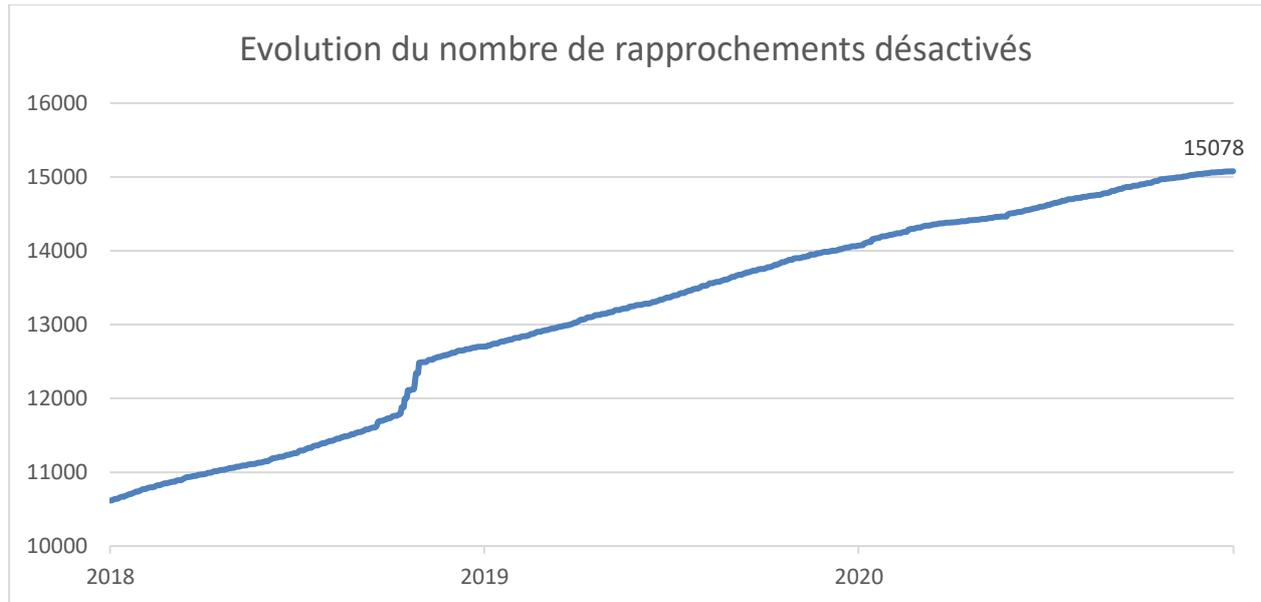


Figure 16 - Évolution du nombre de rapprochements désactivés dans l'annuaire référentiel d'identification des patients

### 2.2.5. Nombre de rapprochements par domaine d'identification

Le nombre de rapprochements d'identités par domaine d'identification est un bon indicateur de l'utilisation de l'annuaire référentiel d'identification des patients par un établissement ou de la couverture de la population des services de la Plateforme eSanté. La Figure 17 présente d'ailleurs la répartition des rapprochements par établissement en date du 31 décembre 2020. Les domaines d'identification transverses tels que le CCSS ou le DSP sont volontairement non représentés car ils écraseraient le graphique, mais ils sont bien visibles dans le Tableau 2 suivant.

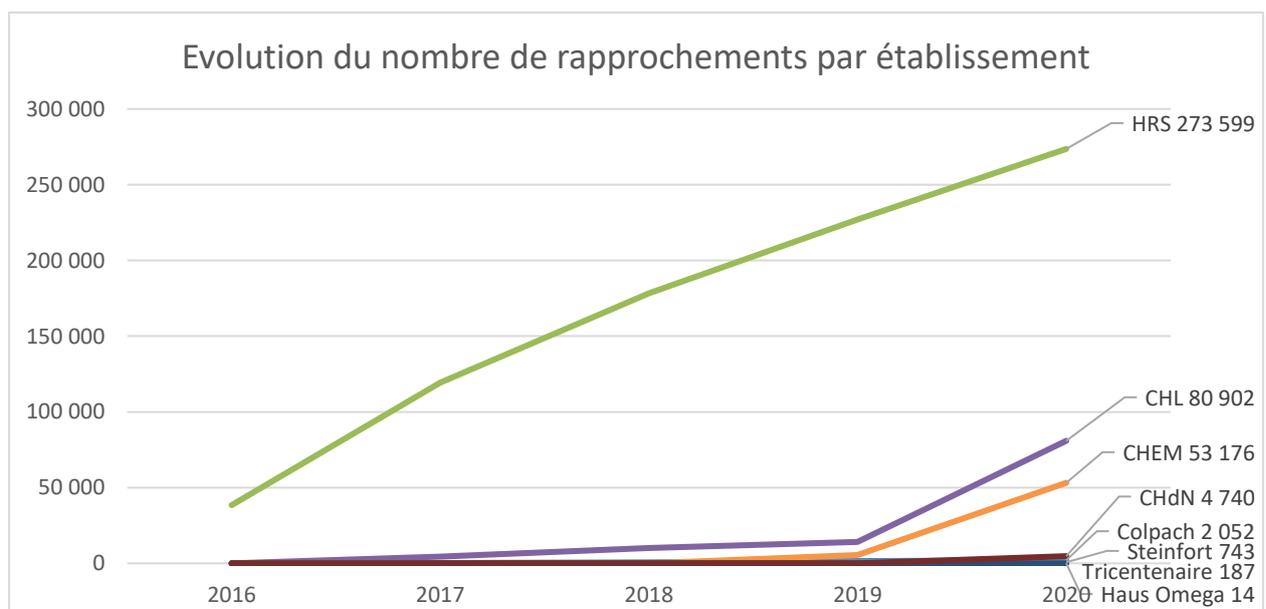


Figure 17 - Répartition des rapprochements (effectifs et pourcentages) par Domaine d'Identification

Le domaine d'identification HRS représente 66% des rapprochements liés aux établissements. Historiquement, c'est le premier établissement à s'être connecté à l'annuaire et chaque patient admis est rapproché avec l'annuaire. Le CHL et le CHEM se sont connectés plus tardivement et rapprochent uniquement les patients disposant d'un DSP. La généralisation de ce dernier a logiquement fait croître le nombre de patients rapprochés. Le CHdN ne s'est connecté à l'annuaire qu'en décembre 2020, ce qui explique le nombre faible de rapprochements.

Cette différence de stratégie de rapprochement n'est pas encore problématique car le DSP est le seul service actuellement en production qui utilise l'annuaire. Un questionnement sera nécessaire lors du lancement de la prescription électronique, du carnet de vaccination ou pour utiliser le service de pseudonymisation.

Le Tableau 2 ci-dessous détaille le cumul des rapprochements par année pour tous les domaines d'identification:

DOMAINE D'IDENTIFICATION	TYPE DE DOMAINE	2016	2017	2018	2019	2020
<b>COLPACH</b>	Producteur	38	157	705	1 437	2 052
<b>TEST POUR AGENCE</b>	Producteur	2	0	14	27	36
<b>TRICENTENAIRE</b>	Producteur	21	66	90	91	187
<b>HRS</b>	Producteur	38 432	119 236	178 375	226 969	273 599
<b>CCSS</b>	Réseau	2 086 970	2 172 880	2 277 915	2 375 167	2 458 940
<b>CHL</b>	Producteur	0	4 441	10 200	14 200	80 902
<b>STEINFORT</b>	Producteur	0	0	80	447	743
<b>CHEM</b>	Producteur	0	0	0	5 483	53 176
<b>HAUS OMEGA</b>	Producteur	0	0	0	1	14
<b>CHDN</b>	Producteur	0	0	0	0	4 740
<b>PSEUDONYMISATION SPS</b>	Réseau	517	0	519	614	629
<b>DSP</b>	Réseau	22 867	51 829	55 836	59 074	896 091
<b>IDEOMED</b>	Réseau	34	156	758	1 772	2 686
<b>MÉDECIN RÉFÉRENT</b>	Réseau	22 104	23 005	24 173	25 613	27 365
<b>IDÉORCP</b>	Réseau	5	0	0	0	0
<b>EDOCUMENTS</b>	Réseau	0	0	0	0	25 111
<b>TOTAL</b>		<b>2 170 990</b>	<b>2 371 770</b>	<b>2 548 665</b>	<b>2 710 895</b>	<b>3 826 271</b>

Tableau 2 - Répartition des rapprochements par domaine d'identification (vue annuelle cumulée)

Nous pouvons calculer un nombre de rapprochements moyen par identité de 1,4 en 2019 et de 1,55 en 2020, en se basant sur le nombre de rapprochements valides (3 481 160), à confirmer (2 256), provisoires modifiés (327 874) sur le nombre d'identités valides (2 451 211) plus les 495 provisoires. Cet indicateur devrait croître avec l'augmentation des usages de la Plateforme eSanté et le nombre d'établissements connectés.

## 2.2.6. Anomalies sur les identités

Le nombre de collisions détectées est pour le moment nul. Précisons que de manière générale en cas de doute, la bonne pratique en matière d'identito-vigilance est toujours de créer une nouvelle identité (et donc éventuellement un doublon) plutôt que de risquer une collision. Ce raisonnement a été retranscrit dans la configuration de l'algorithme de rapprochement d'identité de telle manière qu'un rapprochement est créé seulement si le taux de correspondance des traits stricts est égal à 100%.

Les doublons sont qualifiés de *potentiels*, quand on soupçonne que les identités concernées correspondent à la même personne, ou *d'avérés*, quand le doute ne subsiste plus à ce sujet. Dans ce dernier cas, le traitement des doublons consiste à opérer une *fusion* des identités, l'une acquérant le statut de *maître* et l'autre de *doublon* qui sera alors désactivé. Les outils d'analyse de l'annuaire permettent de détecter bon nombre de ces doublons. Dans cette section, nous observons l'évolution du nombre de doublons potentiels et de fusions réalisées sur le domaine d'identification de l'annuaire.

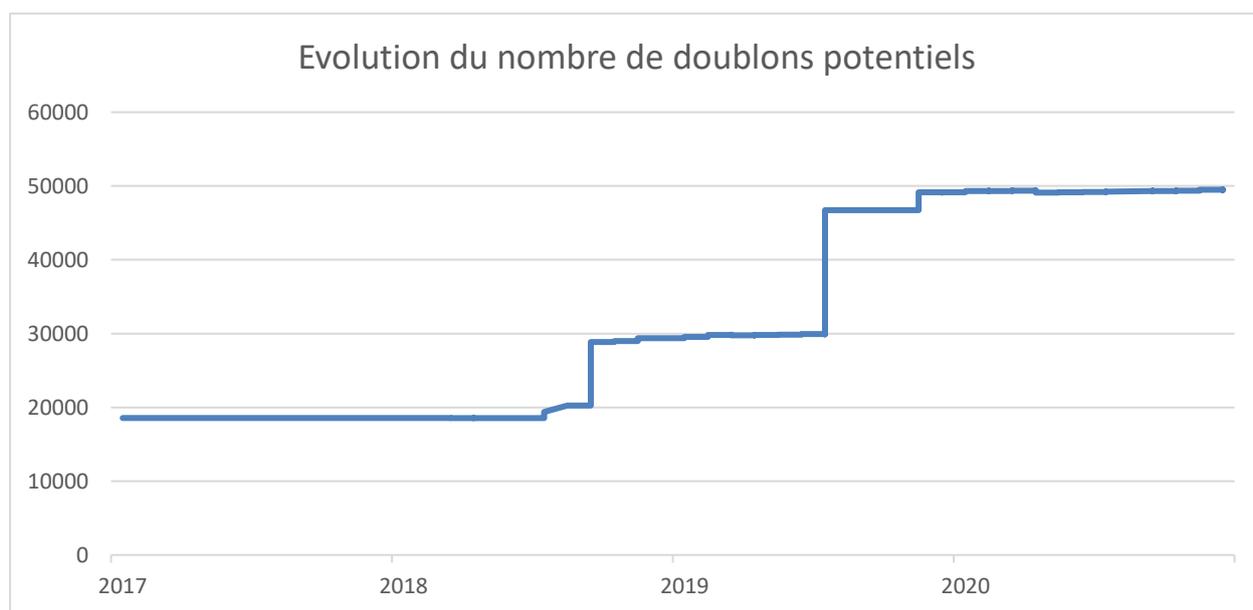


Figure 18 - Évolution du nombre de doublons potentiels sur le domaine d'identification de l'annuaire référentiel d'identification des patients

Le nombre de doublons potentiels s'établit à 49 547 au 31 décembre 2020, la croissance des doublons potentiels s'est limitée à seulement 422 nouveaux cas. Le taux de doublons potentiels sur le nombre total d'identités valides se stabilise à 2%.

Il est à noter qu'il s'agit là de doublons *potentiels* : il se trouve donc parmi eux des cas où l'on a bien affaire à deux personnes différentes que l'on peut qualifier *d'homonyme*. L'évolution du nombre de doublons potentiels n'est pas linéaire. Les sauts s'expliquent par des recherches de doublons plus larges qu'usuellement (par exemple, en considérant des recherches sur la date de naissance uniquement sur le mois et l'année). Plus il y a de domaines d'identification connectés à l'annuaire, plus les chances de créer des doublons augmentent.

Le nombre de doublons potentiels peut néanmoins diminuer : c'est le cas quand ils sont avérés et font donc l'objet d'une fusion. Ces cas sont donc soustraits du nombre des doublons potentiels et viennent

augmenter l'effectif des fusions (doublons avérés). A la date du 31 décembre 2020, 4 810 fusions d'identités fédératrices ont été réalisées. Le traitement des doublons faisant l'objet d'un travail régulier au sein de l'Agence eSanté, on observe logiquement une croissance sensiblement régulière du nombre de fusions d'identités au cours du temps, dans les cas où des doublons avérés sont rencontrés. Ce travail peut être observé en faisant le suivi trimestriel du nombre de fusions (voir Figure 19). On s'aperçoit alors que le nombre de fusions a plus que doublé (1468 en 2020 et 599 en 2019). Cependant, il faudra affiner cette statistique dans les prochaines années, pour différencier les fusions provenant des messages A40 des établissements des fusions manuelles. On estime au 1<sup>er</sup> avril 2021 que nous avons reçu 63 identités par les établissements. La cellule d'identito-vigilance a créé des algorithmes externes (scripts R) pour faciliter son travail et détecter plus facilement les similitudes et différences entre les identités. Ce travail a permis d'augmenter significativement le nombre de cas fusionné.

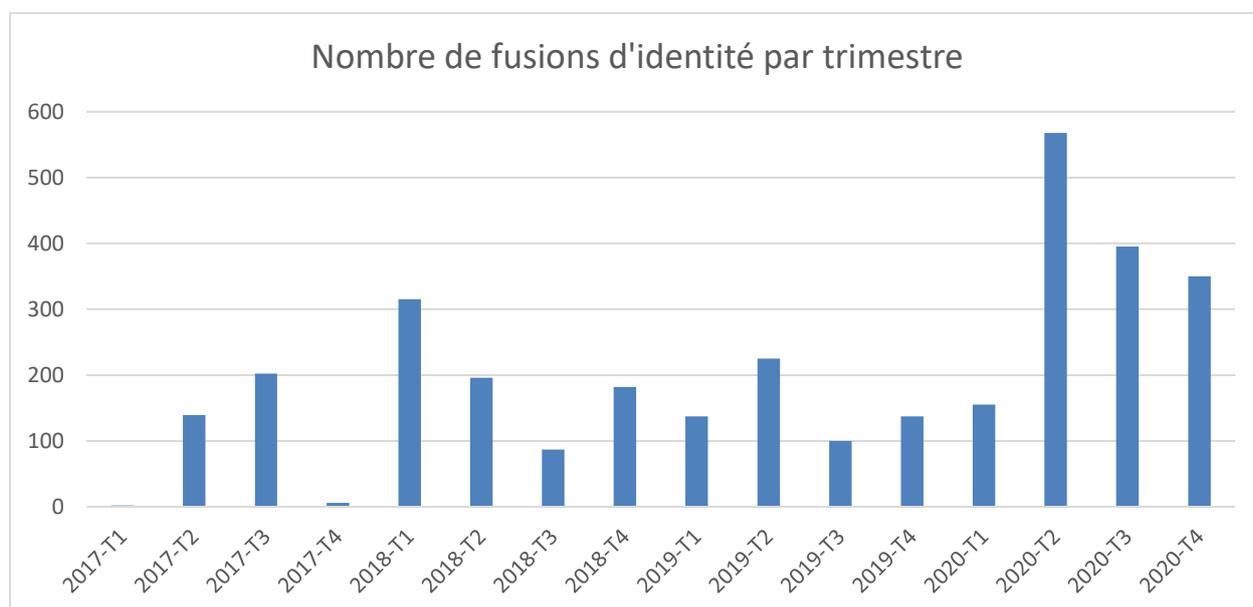


Figure 19 - Nombre de fusions sur le domaine d'identification de l'annuaire référentiel d'identification des patients

Une analyse des causes de doublons sont le plus souvent des anomalies de rapprochements liées la plupart du temps à une confusion entre un nom de naissance et le nom marital, puis des fautes d'orthographe, les caractères diacritiques ü, ö, ß non utilisés, ou bien à un doublon intra-établissement (une autre identité locale était déjà rapprochée). Pour les personnes sans prénom, nous préconisons l'usage de « --- ». Il est également impératif de nommer le patient de son vrai nom (pas de pseudonyme ou nom de religieux).

Lors d'une admission, nous préconisons de faire une recherche sur l'annuaire pour identifier les éventuelles différences et donc, soit corriger l'annuaire de l'établissement, soit demander au patient de signaler l'erreur d'enregistrement dans le Registre National des Personnes Physiques. Ce processus éviterait des erreurs de saisie (nom d'usage, faute d'orthographe et inversion nom / prénom). De même, une recherche systématique de l'existence d'un patient doit être faite avant création d'une nouvelle identité locale, pour ne pas créer de doublon d'établissement.

## 2.3. Indicateurs d'identito-vigilance secondaires

Les indicateurs d'identito-vigilance secondaires correspondent au nombre de signalements des utilisateurs de l'annuaire référentiel d'identification des patients, notamment :

- Le nombre de signalements d'évènements indésirables relatifs à l'identification primaire des patients sur un an
- Le nombre de signalements d'évènements indésirables relatifs à l'identification secondaire des patients sur un an

L'identification primaire désigne l'identification administrative du patient lors de son admission dans l'établissement. Tandis que l'identification secondaire concerne celle du patient par les différents services dudit établissement et inclut l'aspect médical. Ainsi, par exemple, un problème d'identification secondaire peut être le rattachement d'un échantillon biologique au mauvais patient.

Jusqu'à présent aucun signalement n'a été communiqué à l'Agence eSanté par les établissements.

## 3. INDICATEURS D'IDENTITO-VIGILANCE DES ETABLISSEMENTS

Les indicateurs d'identito-vigilance attendus des établissements de santé ont été présentés dans le comité national d'identito-vigilance du 5 décembre 2019, puis rediscutés le 22 septembre 2020. La discussion a mené sur la nécessité de retravailler sur la définition des indicateurs et d'écarter dans un premier temps, les indicateurs « processus » plus difficiles à capter.

Pour le moment aucun établissement n'a communiqué à l'Agence eSanté ses indicateurs. Certains établissements nous ont fait part qu'ils étaient en train de collecter ces informations, tandis que d'autres souhaitaient des informations complémentaires. Dans les sections suivantes nous rappelons les indicateurs demandés.

L'Agence eSanté rencontrera en 2021 les établissements connectés pour redéfinir de manière plus concrète les indicateurs listés dans la charte actuelle de rapprochement de l'annuaire.

### 3.1. Rappel des indicateurs de qualité au niveau des établissements de santé

Les indicateurs relatifs à la qualité des processus, devant être collectés auprès des établissements de santé, sont les suivants :

- Contrôles préalables d'identito-vigilance par l'établissement de santé connecté Plateforme eSanté
  - Existence d'une procédure de recherche et création d'une identité (oui/non)
  - Existence d'une procédure de contrôle et correction des identités (oui/non)

- Indicateurs d'habilitation sur le système d'information de l'établissement
  - Nombre de personnes habilitées à créer des identités dans un établissement
  - Nombre de personnes habilitées à modifier les identités dans un établissement
  - Nombre de établissements ou d'applications habilitées à créer une identité
- Indicateurs de formation
  - Nombre d'agents de l'établissement ayant reçu une formation continue sur l'identification
- Indicateurs de processus
  - Nombre total d'identités actives au sein de l'établissement
  - Nombre total d'identités désactivées/archivées (pour décès ou erreur de saisie d'une identité) au sein de l'établissement
  - Nombre total d'identités dans un état valide (si applicable au sein de l'établissement)
  - Nombre total d'identités dans un état provisoire (si applicable au sein de l'établissement de santé)
  - Nombre total d'identités dans un état imprécis (si applicable au sein de l'établissement)
  - Nombre total d'identités dans un état douteux (si applicable au sein de l'établissement)

## 3.2. Indicateurs d'anomalies sur les identités au niveau des établissements de santé

Les indicateurs relatifs aux anomalies sur les identités devant être collectés auprès des établissements de santé, sont les suivants :

- Nombre de doublons avérés par l'établissement de santé
- Moyenne du nombre de doublons avérés par établissement de santé sur 5 ans
- Ratio du nombre de doublons avérés par l'établissement de santé/moyenne du nombre de doublons avérés par l'établissement de santé sur 5 ans
- Nombre de collisions avérées par l'établissement de santé
- Moyenne du nombre de collisions avérées par établissement de santé sur 5 ans
- Ratio du nombre de collisions avérées par l'établissement de santé/moyenne du nombre de collisions avérées par l'établissement de santé sur 5 ans