

## [ Chaînage de données : pourquoi et comment faire ? ]

### ① Chaînage avec des données SNDS ?

**On peut chaîner des données collectées (bases de données cliniques, données recueillies dans le cadre de cohortes épidémiologiques, ...) avec les données du SNDS** : cela permet de compléter les données cliniques dont on dispose par les données du SNDS, et donc d'être sûr des patients étudiés et d'obtenir le parcours complet des patients (en termes

de consultations, d'hospitalisations, de consommations médicamenteuses... d'estimer le coût de la prise en charge...). Dans certains cas, cette démarche permet de limiter le recueil de données auprès des patients ou de leurs professionnels de santé, en sachant qu'on va récupérer certaines données dans le SNDS.

### ② Comment chaîner des données au SNDS ?

Deux méthodes sont possibles :

- **Appariement direct avec l'utilisation du NIR (numéro de sécurité sociale)**

Une base de données contenant le NIR des patients peut désormais être appariée directement au SNDS (*Article 193 de la loi de modernisation de notre système de santé, décret n° 2019-341 du 19 avril 2019*)<sup>(1)</sup>. Après avis favorable du CESREES et autorisation de la CNIL, l'investigateur transmet directement la table des NIR à la CNAM selon le process mis en place par la CNAM, pour procéder à l'extraction des données du SNDS correspondant aux NIR de la table transmise.

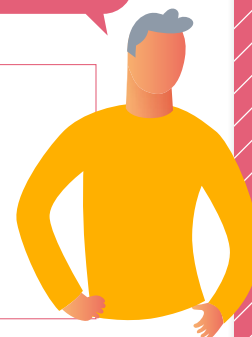
Plus d'info et de nombreuses questions/réponses : <https://www.cnil.fr/fr/tout-savoir-sur-le-decret-cadre-nir-dans-le-champ-de-la-sante>

- **Appariement indirect à partir de données indirectement identifiantes (date de naissance, sexe, ...)**

Si l'on ne dispose pas du NIR des patients, on peut réaliser un appariement indirect (également appelé « probabiliste »). Dans ce cas, on transmet à la CNAM une base de données contenant *a minima* le sexe, le mois et année de naissance, le lieu de résidence (département), la date de décès le cas échéant, après avis favorable du CESREES et autorisation de la CNIL. L'appariement est réalisé par la CNAM à partir des variables communes aux deux bases.

Par exemple :

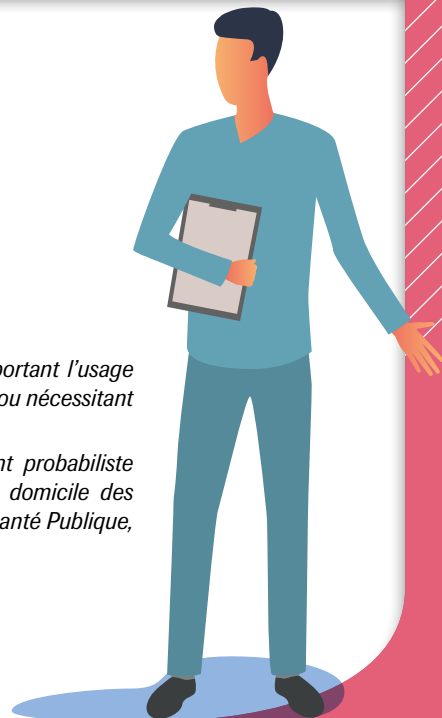
L'appariement indirect d'un programme de télésurveillance à domicile des patients atteints d'insuffisance cardiaque en Normandie au Système National des Données de Santé a permis d'identifier 87,4% des patients de manière unique<sup>(2)</sup>. L'objectif était d'étudier l'impact de la télésurveillance sur les réhospitalisations et la mortalité des patients.



**Important**



**La présence d'un acte traceur  
(date d'une hospitalisation  
dans un établissement donné,  
délivrance d'un médicament par exemple)  
permet d'augmenter les probabilités d'appariement.**



*(1) Décret n° 2019-341 du 19 avril 2019 relatif à la mise en œuvre de traitements comportant l'usage du numéro d'inscription au répertoire national d'identification des personnes physiques ou nécessitant la consultation de ce répertoire.*

*(2) L. Courouve, M. Jodar, M. Gabbas, D. Legallois, V. Donio, R. Sabatier, Appariement probabiliste au Système national des données de santé d'un programme de télésurveillance à domicile des patients atteints d'insuffisance cardiaque en Normandie, Revue d'Épidémiologie et de Santé Publique, Volume 68, Supplement 1, 2020.*

