



HEALTH
DATA HUB

A photograph showing several people's hands and arms around a table. One person is pointing at a tablet displaying a map. There are notebooks and pens on the table. The image is partially overlaid with a purple and blue geometric design.

SNDS Pourquoi un algorithme de sélection des cas ? Pourquoi le valider ?

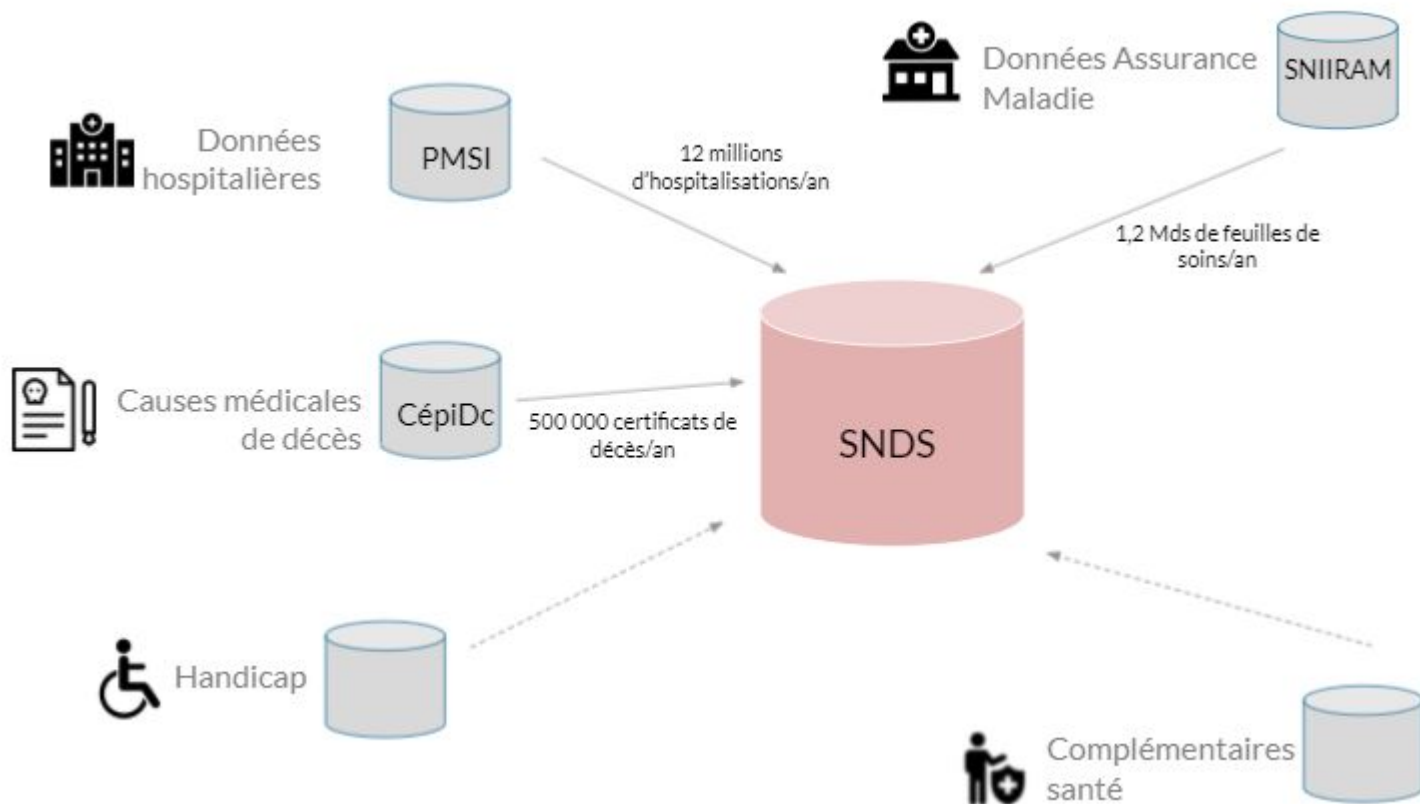
Health Data Hub / Direction scientifique
Septembre 2021

Sommaire

1. **Contexte**
2. **Objectif**
3. **Cas d'usage**
 - 3.1. **Objectifs de l'étude**
 - 3.2. **Développement de l'algorithme de sélection des cas**
 - 3.2.1. **Méthode**
 - 3.2.2. **Résultats**
 - 3.3. **Validation de l'algorithme de sélection des cas**
 - 3.3.1. **Méthode**
 - 3.3.2. **Résultats**
 - 3.4. **Discussion**
4. **Discussion / Conclusion**

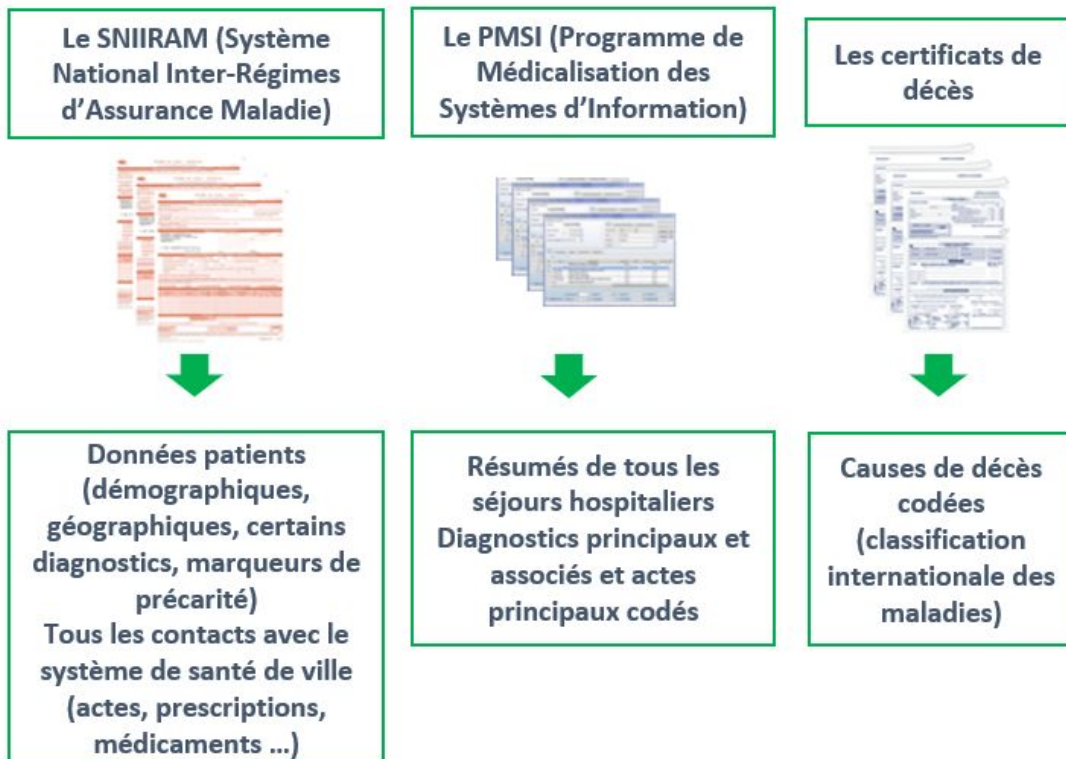
Contexte

Ce que contient le SNDS



Contexte

Ce que contient le SNDS



Ce que contient le SNDS

Soins en ville

- ✓ Prestations remboursées avec le codage détaillé (actes médicaux, biologie, dispositifs médicaux, médicaments)
- ✓ Montants (présentés au remboursement, base de remboursement, remboursé)
- ✓ Détail par date de soins et date de remboursement

Soins en établissement

- ✓ Les séjours facturés directement à l'Assurance Maladie : cliniques privées, une partie du secteur médico-social handicap
- ✓ Les séjours du PMSI pour tous les établissements publics ou privés, pour le MCO, les SSR, la HAD et la PSY

Offre de soins

- ✓ Spécialité du prescripteur et de l'exécutant
- ✓ Lieu d'exécution
- ✓ Département et commune d'implantation du cabinet
- ✓ Statut conventionnel (libéral) et statut juridique (établissement)

Informations sur les bénéficiaires

- ✓ Âge,, régime d'affiliation, sex, département et commune et de résidence
- ✓ Bénéfice de la CMU-C, de l'ACS, et de l'AME
- ✓ Diagnostic de l'ALD (pour une prise en charge à 100%)
- ✓ Date de décès

Ce que NE contient PAS le SNDS :

- Les diagnostics lors des consultations médicales ou paramédicales (en ville)
- Les résultats des examens cliniques (tension artérielle, indice de masse corporelle), de laboratoire, de l'histologie, de l'imagerie médicale...
- Certaines données sociodémographiques : le niveau d'études, les revenus ou la catégorie socioprofessionnelle des assurés, le statut marital...
- Certains facteurs de risque : tabagisme, alcool, activité physique, nutrition, antécédents familiaux...
- Les médicaments délivrés dans les services d'hospitalisation ou de long séjour, autres que ceux figurant sur la liste en sus ou dans le cadre de forfaits (cliniques, établissements de santé, etc.)
- Le dosage prescrit des médicaments (information limitée à la quantité délivrée, nécessitant une estimation de la dose quotidienne) Les médicaments non remboursés ou actes non facturés

Répondre à deux questions :

- Pourquoi faut-il dans certains cas mettre en place des algorithmes de sélection de cas dans le SNDS ?
- Pourquoi une validation de ces algorithmes peut s'avérer nécessaire ?

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS



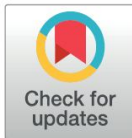
RESEARCH ARTICLE

Identifying chronic thromboembolic pulmonary hypertension through the French national hospital discharge database

V. Cottin¹, D. Avot², L. Lévy-Bachelot², C. A. Baxter³, D. R. Ramey⁴, L. Catella⁵, S. Bénard^{5*}, O. Sitbon⁶, S. Teal⁷

1 National Reference Centre for rare pulmonary diseases, Competence centre for pulmonary arterial hypertension, Louis Pradel hospital, Claude Bernard University Lyon 1, UMR 754, Lyon, France, **2** MSD France, Courbevoie, France, **3** MSD Ltd, Hoddesdon, United Kingdom, **4** Merck & Co., Inc., Kenilworth, NJ, United States of America, **5** stève consultants, Oullins, France, **6** Université Paris-Sud, CHU de Bicêtre, Le Kremlin-Bicêtre, France, **7** Bayer AG, Berlin, Germany

* sbenard@steve-consultants.com



Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

En l'absence de code diagnostic spécifique à l'HPTC:

Objectifs de l'étude :

- **Développer** un algorithme de sélection des cas pour identifier les patients atteints de CTEPH dans la base de données du PMSI.
- **Valider** un algorithme de sélection des cas pour identifier les patients atteints de HPTC dans la base de données du PMSI.
- Estimer le nombre de patients atteints de HPTC hospitalisés en France en 2015 .
-> *Cette partie ne sera pas présentée*

Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Développement de l'algorithme de sélection des cas

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Méthode : Développement de l'algorithme de sélection des cas

En l'absence de code diagnostique spécifique à la HPTC :

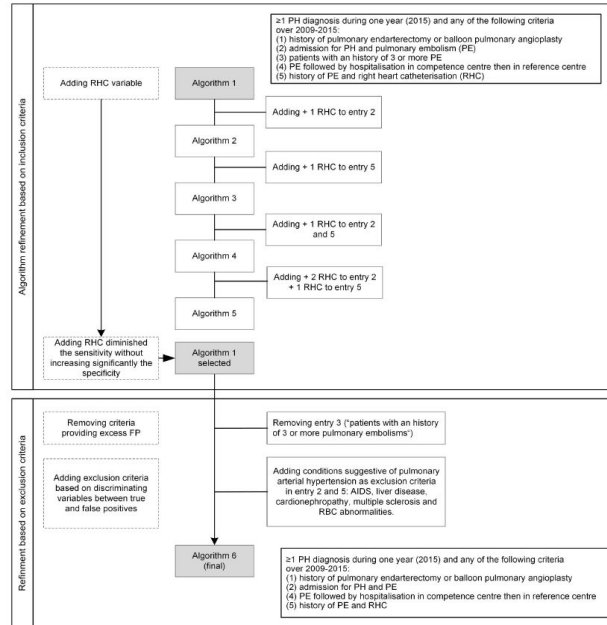
- Caractérisation du parcours de soins pour la prise en charge des patients atteints de HPTC :
 - Un recours à une expertise sur l'HPTC (pneumologues et experts en maladies rare)
 - Revue de littérature sur la pathologie
 - Revue de littérature sur les algorithmes mis en place dans d'autres études

Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Méthode : Développement de l'algorithme de sélection des cas

Au total 5 algorithmes ont été testés pour arriver à l'algorithme final :



Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Résultats : Développement de l'algorithme de sélection des cas

Algorithme final

- Au moins une hospitalisation pour hypertension pulmonaire au cours de l'année 2015
ET
 - Un des critères suivants sur la période 2009-2015 :
 - (i) acte médical (code CCAM de angioplastie par ballonnet ou endartériectomie pulmonaire) ;
 - (ii) admission pour hypertension pulmonaire **ET** embolie pulmonaire ;
 - (iii) embolie pulmonaire suivie d'une hospitalisation en centre de compétence puis en centre de référence ;
 - (iv) antécédents d'embolie pulmonaire et de cathétérisme cardiaque droit.
- Les patients présentant des conditions évocatrices d'hypertension artérielle pulmonaire ont été

Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Validation de l'algorithme de sélection des cas

Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Méthode : Validation de l'algorithme de sélection des cas

Première étape :

Mise en place d'une base de données de cas **confirmé** de HPTC

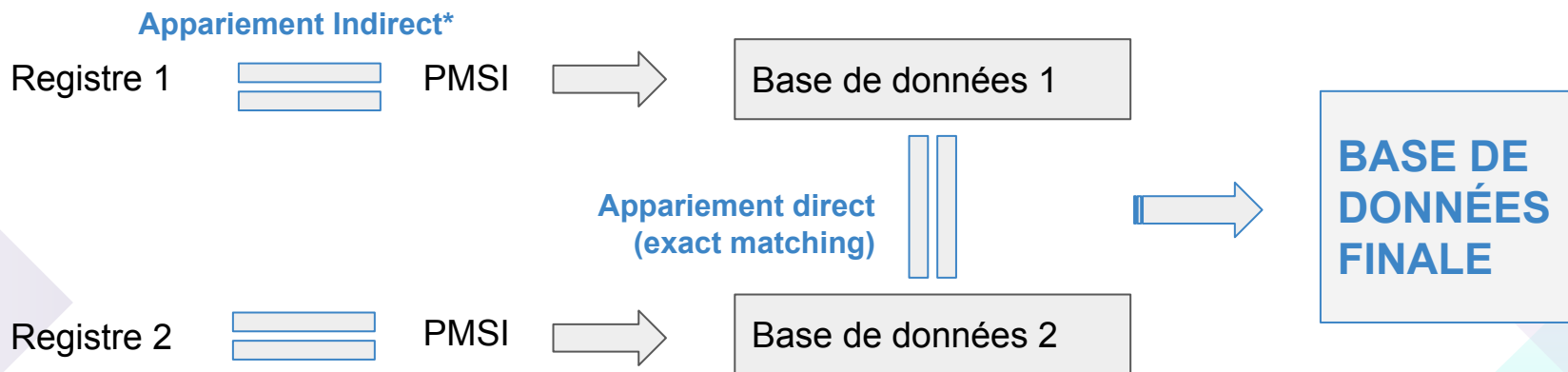
→ à partir des dossiers médicaux de deux hôpitaux en France qui gèrent tous deux un **registre prospectif et exhaustif**.

Cas d'usage

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Méthode : Validation de l'algorithme de sélection des cas

Deuxième étape :



Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Méthode : Validation de l'algorithme de sélection des cas

Troisième étape :

- Calcul de la performance des 6 algorithmes en prenant l'un des registres comme référence :
 - Vraies positifs : correctement identifiées dans le PMSI comme **cas** HPTC
 - Vraies négatifs : correctement identifiées dans le PMSI comme **non-cas** HPTC
 - Faux positifs : identifiés comme cas dans le PMSI mais qui ne sont pas des cas dans les registres
 - Faux négatifs : non-identifiés comme cas dans le PMSI mais qui sont des cas dans les registres
- La sensibilité : nombre de personnes atteint de HPTC qui sont identifiés comme cas
- La spécificité : nombre de personnes non atteint de HPTC qui sont identifiés comme non-cas
- La valeur prédictive positive : le pourcentage de cas positifs prévus et qui sont en réalité positifs.
- La valeur prédictive négative : le pourcentage des cas négatifs prévus et qui sont en réalité négatifs.

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Résultats : Validation de l'algorithme de sélection des cas

Table 1. Performance of tested case ascertainment algorithms on training and validation sets (medical charts records as the reference standard).

Algorithm definition	True-positive/ false-positive	True-negative/ false-negative	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)
1st algorithm (*)	132/123	17/750	88.6 (83.5–93.7)	85.9 (83.6–88.2)	51.8 (45.6–57.9)	97.8 (96.7–98.8)
2nd algorithm (* +1 RHC in entry 2)	121/102	28/771	81.2 (74.9–87.5)	88.3 (86.2–90.4)	54.3 (47.7–60.8)	96.5 (95.2–97.8)
3rd algorithm (* +1 RHC in entry 5)	128/120	21/753	85.9 (80.3–91.4)	86.3 (83.4–88.5)	51.6 (45.4–57.8)	97.3 (96.1–98.4)
4th algorithm (* +1 RHC in entry 2, +1 RHC in entry 5)	117/99	32/774	78.5 (71.9–85.1)	88.7 (86.6–90.7)	54.2 (47.5–60.8)	96.0 (94.7–97.4)
5th algorithm (* + 2 RHC in entry 2 + 1 RHC in entry 5)	101/78	48/795	67.8 (60.2–75.2)	91.1 (89.1–93.0)	56.4 (49.2–63.7)	94.3 (92.7–95.9)
6th algorithm (* entry 3 removed) on training set	129/75	20/798	86.6 (81.1–92.1)	91.4 (89.6–93.3)	63.2 (56.7–69.9)	97.6 (96.5–98.1)
6th algorithm (final) on validation set	38/55	16/1,051	70.4 (58.2–82.5)	95.0 (93.7–96.3)	40.9 (30.9–50.8)	98.5 (97.8–99.2)

CI, Confidence Interval; NPV, Negative Predictive Value; PPV, Positive Predictive Value; RHC, right heart catheterization

* First/preliminary algorithm: ≥ 1 PH diagnosis during one year (2015) and any of the following criteria over 2009–2015: (i) history of PEA (medical procedures-CCAM code) or BPA (ICD-10 code of PH and a CCAM code of BPA), (ii) hospital stay for PH and PE, (iii) history of ≥ 3 PE, (iv) CTEPH care pathway: hospital stay for PE, then for PH in competence centre, then for PH in a reference centre, (v) history of PE and 1 RHC.

Identification de cas d'hypertension pulmonaire thromboembolique chronique (HPTC) à partir des données du SNDS

Discussion :

- Test de plusieurs algorithmes suite à l'avis des experts et à la revue de littérature
- Pour cette pathologie, les auteurs ont fait le choix de l'algorithme qui avait une spécificité et une VPN importantes alors que la sensibilité et la VPP étaient faibles
- Ces paramètres de performance de l'algorithme ont été pris en compte pour ajuster leur estimation de la prévalence (non exposée dans cette présentation)

- ❑ La mise en place d'un algorithme est dans certains cas incontournable pour la sélection des patients
- ❑ Avant de proposer un algorithme de sélection de cas, s'assurer que :
 - ❑ l'algorithme n'existe pas déjà en France ou à l'international
 - ❑ Bien documenter la pathologie (faire appel à des experts médicaux, revue de littérature...)
- ❑ Il est fortement recommandé, et lorsque c'est possible, d'effectuer une validation de l'algorithme en utilisant des bases de données avec des diagnostics validés par un médecin (registres, cohortes ...).

Comment candidater



Formulaire de Candidature

- Le formulaire est à remplir sur la plateforme Démarches Simplifiées au lien suivant : <https://www.demarches-simplifiees.fr/commencer/boas-ami-2021-2022>



Date limite de candidature

- Pour la première vague : 24 septembre 2021 à 12h00 (heure de Paris)
- Pour l'ensemble de l'AMI : Juin 2022



Contact

- opensource@health-data-hub.fr



Suivez-nous sur les réseaux sociaux !

