

Défis et opportunités de l'appariement d'une base d'accès précoce à la base du Système National des Données de Santé (SNDS) : l'expérience de l'étude SHELDON-PRO

Lauren Majed¹, Aleksandra Anchim¹, Arguichti Set Aghayan¹, Baptiste Jouaneton², Diane-Charlotte Baillet³, Cécile Artaud¹

¹AstraZeneca, 31 Place des Corolles, 92400 Courbovoie, France ; ²Heva, 186 Avenue Thiers, 69006 Lyon, France ; ³Docaposte, 45/47 boulevard Paul Vaillant Couturier, 94200 Ivry-sur-Seine, France

Objectifs de la présentation

L'objectif de ce retour d'expérience est de partager à la communauté médicale et scientifique la méthode, les challenges rencontrés et les premiers résultats d'appariement entre une base d'accès précoce de plus de 27 000 patients et le SNDS.



Éléments clés

Une méthodologie d'appariement robuste et systématique permet de pallier l'absence du NIR, dans le cadre d'un chaînage de données d'AP aux données du SNDS.

Introduction

AZD7442 est une combinaison d'anticorps monoclonaux pour la prévention de la COVID-19 chez les personnes immunodéprimées.

Le 9 décembre 2021, la Haute Autorité de Santé (HAS) a autorisé un Accès Précoce (AP)¹ pour AZD7442 dans l'indication

« en prophylaxie pré-exposition de la COVID-19 chez les patients adultes de 18 ans et plus : Insuffisamment ou non répondeurs après un schéma vaccinal complet conformément aux recommandations en vigueur et appartenant à l'un des sous-groupes à très haut risque de forme sévère de COVID-19 tels que définis par l'ANRS-MIE (patients ayant reçu une greffe d'organe solide, patients atteints d'hémopathie lymphoïde, patients recevant un traitement par anticorps anti-CD20...) OU non éligibles à la vaccination et qui sont à haut risque de forme sévère de COVID-19 ».

Une étude impliquant un recueil des données d'AP chaînées au Système National des Données de Santé (SNDS) a été anticipée conformément à la doctrine de la HAS sur les AP².

Cette étude, intitulée SHELDON-PRO³ visera à évaluer l'efficacité en vie réelle de AZD7442 sur les hospitalisations pour COVID-19 et la mortalité toutes causes confondues évalués 6 mois après l'administration initiale.



Le recueil des données de l'AP a été réalisé pendant la vague épidémique COVID-19 dominée par le variant OMICRON.



AZD7442 était mis à disposition des établissements de santé dans le cadre de stock d'état, ne nécessitant pas de codage obligatoire par les hôpitaux dans le SNDS.



Les dynamiques de remontée des dates d'administrations varient grandement selon la source de données.

Méthode



Source de données

Figure 1 : Description des sources de données

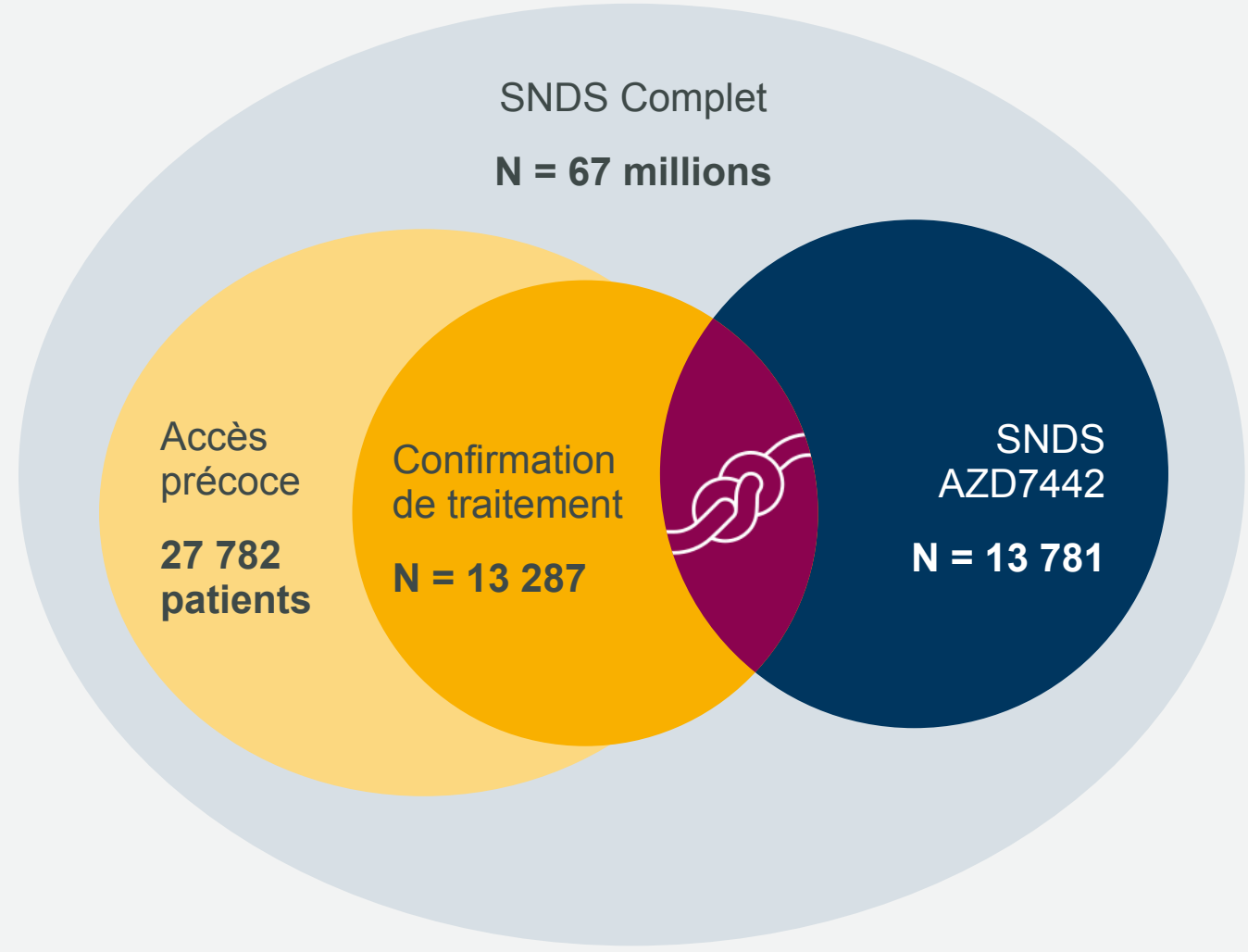
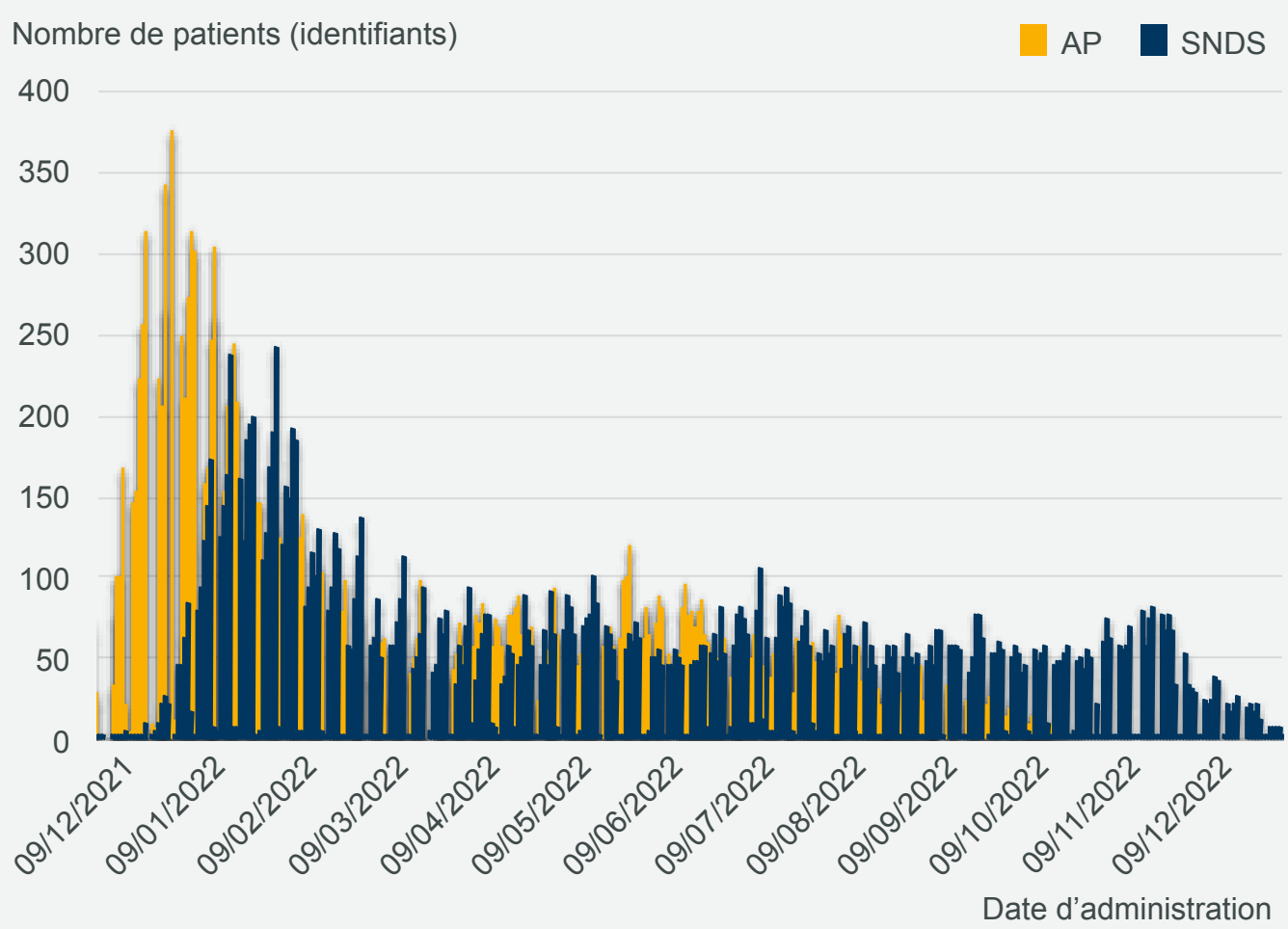


Figure 1 : Nombre d'identifiants par date d'administration d'AZD7442



Période d'étude

Inclusion des patients entre le 09 décembre 2021 et le 30 novembre 2022 (période de l'AP).

Appariement

La collecte du numéro d'inscription au répertoire de la sécurité sociale (NIR) n'étant pas autorisée dans l'AP, l'appariement entre AP et SNDS a été réalisée selon une approche indirecte, en prenant en compte le fait qu'aucune des deux sources de données n'est exhaustive.

Afin de maximiser la qualité de cet appariement, une stratégie complexe en plusieurs étapes a été mise en place, selon la méthode de Fellegi & Sunter⁴. Des taux de correspondances minimales ont été utilisés afin de s'assurer de la qualité de l'appariement.

- Sélection des critères à rapprocher dans les deux bases : mois et année de naissance, sexe, date admin, établissement de prise en charge, ...
- Ordonnancement de ces critères
- Définition des modalités qu'ils peuvent prendre (0, 1, 2, ...)
- Calcul du score pour chacun des couples existants dans les deux bases
- Élection du couple qui obtient le score le plus élevé

AP x SNDS		Vecteurs poids					
		C1	C2	C3	C4	...	Cn
a ₁	b ₁	1	1	1	5	...	1
a ₁	b ₂	1	1	0	5	...	1
a ₂	b ₁	0	1	1	4	...	0
...	...	...	...	...	...	...	...
a _n	b _{n+1}	0	0	0	1	...	1
a _n	b _n	0	0	0	0	...	0

→ Le couple a1b1 est retenu car 11 151 > 11 051

Résultats



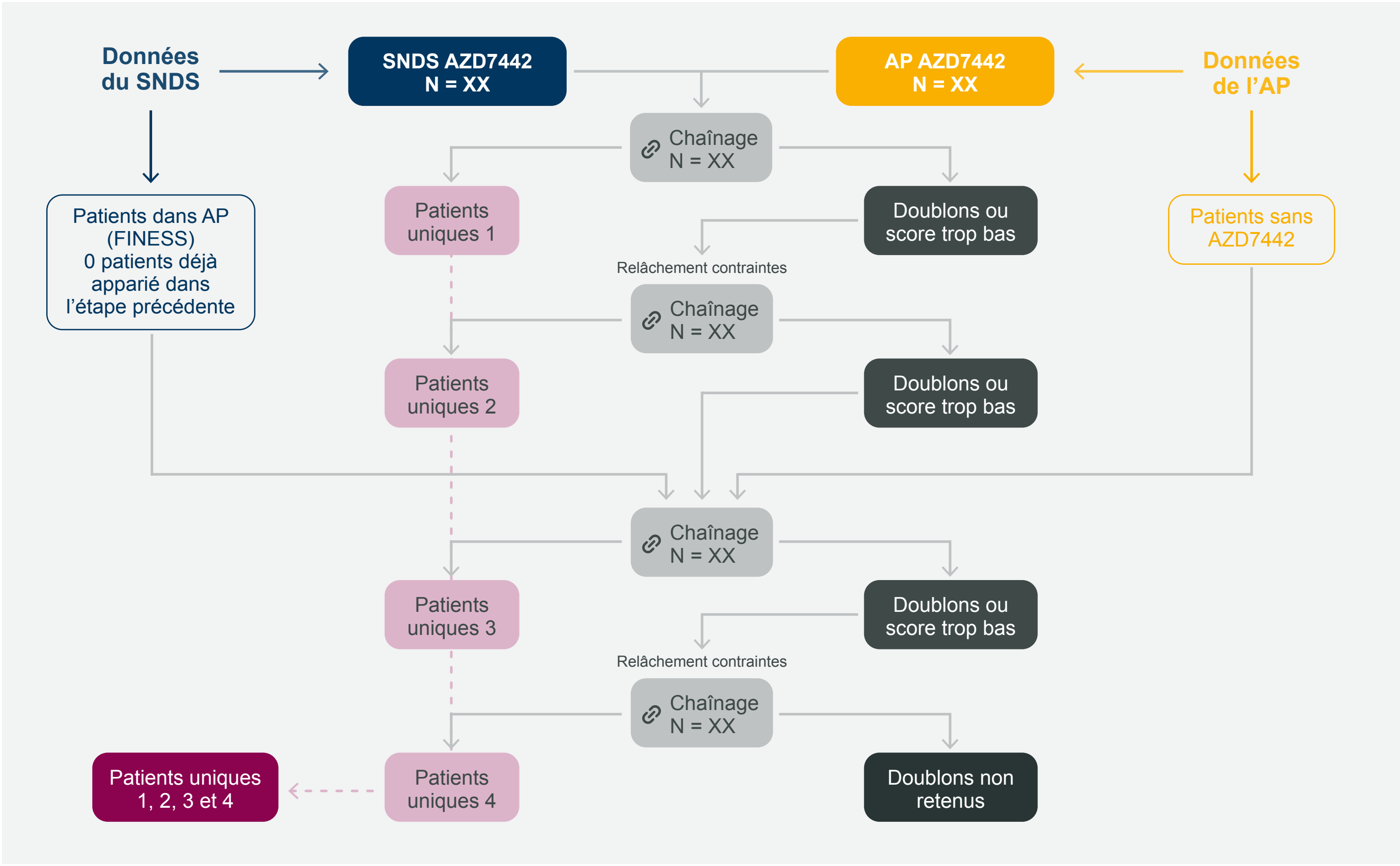
Variables d'appariement

- Année de naissance : 1 ou 0
- Sexe : 1 ou 0
- Date de décès (ou nulle dans l'AP) : 1 ou 0
- Mois de naissance : 1 ou 0
- Couple date AZD7442 / établissement : 5 à 0
- Date casirivimab / imdévimab : 1 ou 0
- Autre date d'hospitalisation (avec durée) : 3 à 0
- Date test PCR : 2 ou 0
- Date de vaccination COVID-19 : 2 à 0
- Nombre de vaccination COVID-19 : 1 ou 0

Sélection du couple d'identifiants AP / SNDS avec le score le plus élevé

- Score maximal = 1 111 513 221
- Score minimal avec 5 premiers critères obligatoires ≥ 1 111 400 000

Flow chart d'appariement



Conclusion



Deux facteurs combinés ont complexifié la méthodologie d'appariement :

- Collecte limitée de données au sein des AP afin de faciliter leur recueil, entraînant une restriction dans leur ré-utilisation.
- Disponibilité de AZD7442 sous forme de stock d'état, entraînant un codage non systématique par les établissements.

De ce fait, il a été mis en œuvre une méthodologie robuste permettant d'optimiser le succès du chaînage des données, et in fine assurer la qualité des analyses de l'étude SHELDON-PRO. L'utilisation d'une méthodologie probabiliste permettrait d'améliorer encore la qualité d'appariement.

Dans le cas de la présente étude, combinant les données du plus grand accès précoce (27 000 patients) et du SNDS, le chaînage va permettre d'évaluer l'efficacité et la tolérance en conditions réelles d'utilisation, avec un suivi à moyen terme.

Ces résultats sont d'autant plus attendus dans le cas de la COVID-19, avec une épidémiologie des variants changeante et un besoin d'actualiser les estimations d'efficacité de produit innovant.

Références

¹HAS. Décision d'accès précoce. Disponible en ligne : https://www.has-sante.fr/jcms/p_3304034/fr/evusheld-tixagevimab/cilgavimab-covid-19-adultes

²HAS. Autorisation d'accès précoce aux médicaments : doctrine d'évaluation de la HAS. Disponible en ligne : Autorisation d'accès précoce aux médicaments : doctrine d'évaluation de la HAS ([has-sante.fr](https://www.has-sante.fr))

³Nguyen et al. Opportunities and challenges in linkage of early access databases with the French National claims database: a RWE effectiveness study of AZD7442 for Pre-Exposure prophylaxis against COVID-19 in Immunocompromised patients. ISPOR 2023.

⁴CépiDC. Note méthodologique sur les appariements probabilistes. Disponible en ligne : https://cepidc.inserm.fr/sites/default/files/2019-10/Note%20sur%20les%20appariements%20probabilistes_0.pdf

⁵<https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000045923753/>

Source des données

Les données proviennent de l'AP chaînées au SNDS, et seront traitées sur la plateforme AGORIA Santé, qui dispose d'une autorisation de la CNIL⁵.

Abréviations

AP : Accès Précoce
CNAM : Caisse National d'Assurance Maladie
HAS : Haute Autorité de Santé
NIR : Numéro d'Inscription au Répertoire de la sécurité sociale
SNDS : Système National des Données de Santé

Financement

Cette étude a été financée par AstraZeneca.

Remerciements

Les auteurs tiennent à remercier les équipes de la Direction de la stratégie, des études et des statistiques, du Département Accès, Traitement et Analyse de la Donnée et des Demandes Externes de la CNAM, en particulier Abdelhakim Boughida pour l'extraction des données et la réalisation de l'appariement.

Conflits d'intérêt

L.M., A.A., A.S.A et C.A. sont des employés d'AstraZeneca.