

CUB Trajectory : Un projet multicentrique pour analyser et prédire l'évolution des maladies bronchiques chroniques grâce aux données en vie réelle

INTRODUCTION

L'asthme et la bronchopneumopathie chronique obstructive (BPCO) sont des maladies bronchiques chroniques majeures dont la prise en charge optimale est essentielle pour améliorer la qualité de vie des patients et prévenir les exacerbations. L'exploitation des données en vie réelle, en particulier celles issues des milieux hospitaliers, joue un rôle clé dans la compréhension de l'évolution de ces pathologies et l'optimisation des stratégies thérapeutiques.

Cependant, la collecte de ces données représente un véritable défi, notamment dans un contexte multicentrique, où l'hétérogénéité des dossiers médicaux et leur structuration complexe compliquent leur harmonisation et leur analyse. De plus, les données déjà existantes comme celles du système national des données de santé (SNDS) ne permettent pas d'avoir ces informations cliniques.

C'est dans cette perspective qu'a été initié le projet « CUB Trajectory », visant à analyser et prédire l'évolution de ces maladies grâce à une approche innovante et intégrée des données de santé.

MÉTHODE

Cette étude repose sur l'analyse des données d'une cohorte multicentrique ambispective. Elle inclut des patients adultes diagnostiqués avec de l'asthme, une BPCO, une bronchectasie ou une maladie respiratoire obstructive rare. La sélection des patients se fait sur la base d'un suivi régulier supérieur à un an et d'une obstruction respiratoire détectée par mesure de la fonction pulmonaire entre le 1er janvier 2023 et le 1er janvier 2030.

Les données, issues des comptes rendus médicaux, sont collectées et structurées à l'inclusion et lors du suivi via une combinaison d'IA générative, qui identifie et organise les variables, en utilisant un modèle de langage étendu (LLM). Elles englobent des informations socio-démographiques, cliniques, biologiques et respiratoires, ainsi que le parcours de soins hospitalier des patients. En parallèle, les données de mortalité ont été extraites automatiquement du registre public français, l'INSEE, pour une intégration précise dans l'analyse.

Critères d'inclusion dans la cohorte CUB

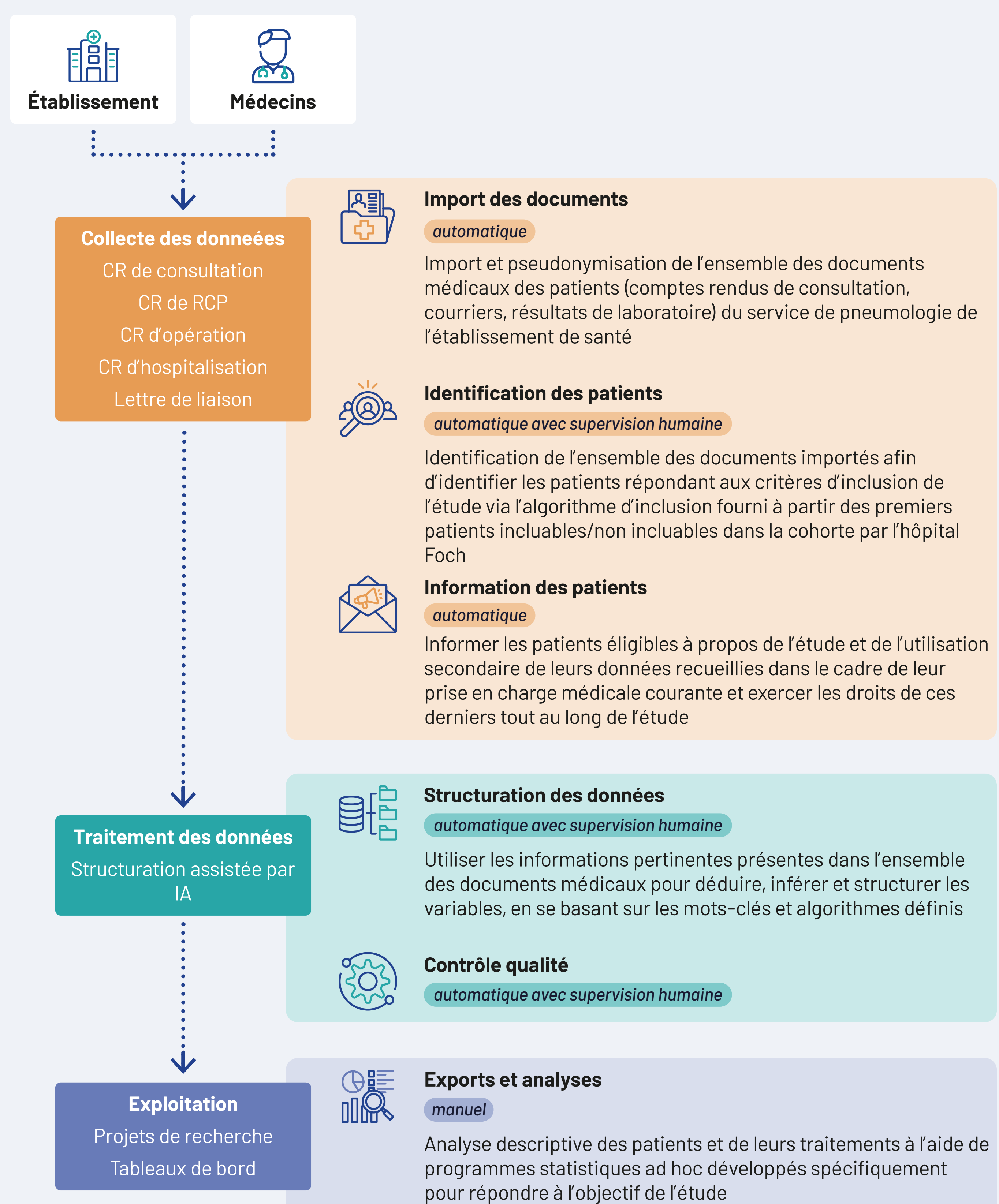
Patient suivi depuis au moins un an dans le même établissement, recruté entre le 01/01/2023 et le 01/01/2030,

- ET un antécédent de maladie chronique des voies aériennes (CAD) : diagnostic d'asthme, de BPCO, de bronchectasie seule ou associée, ou d'autres maladies chroniques des voies aériennes plus rares.
- OU au moins deux mesures de la fonction respiratoire montrant : $VEMS \leq 80\%$ de la valeur théorique ET rapport $VEMS/CVF \leq 0,8$.

Constitution de la base

La spécificité de ce projet réside dans son approche innovante pour identifier les patients et constituer la cohorte.

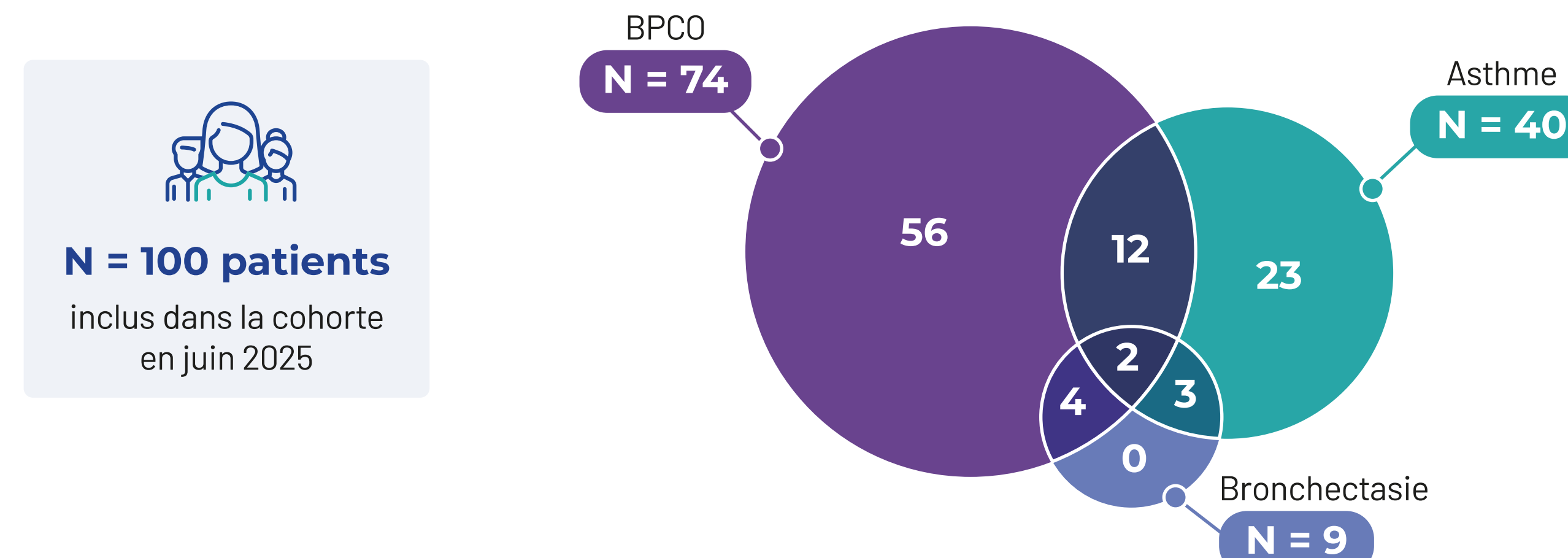
La cohorte est fondée sur une structuration automatisée des données utilisant le traitement du langage naturel par intelligence artificielle (IA-NLP) pour structurer les données.



RÉSULTATS

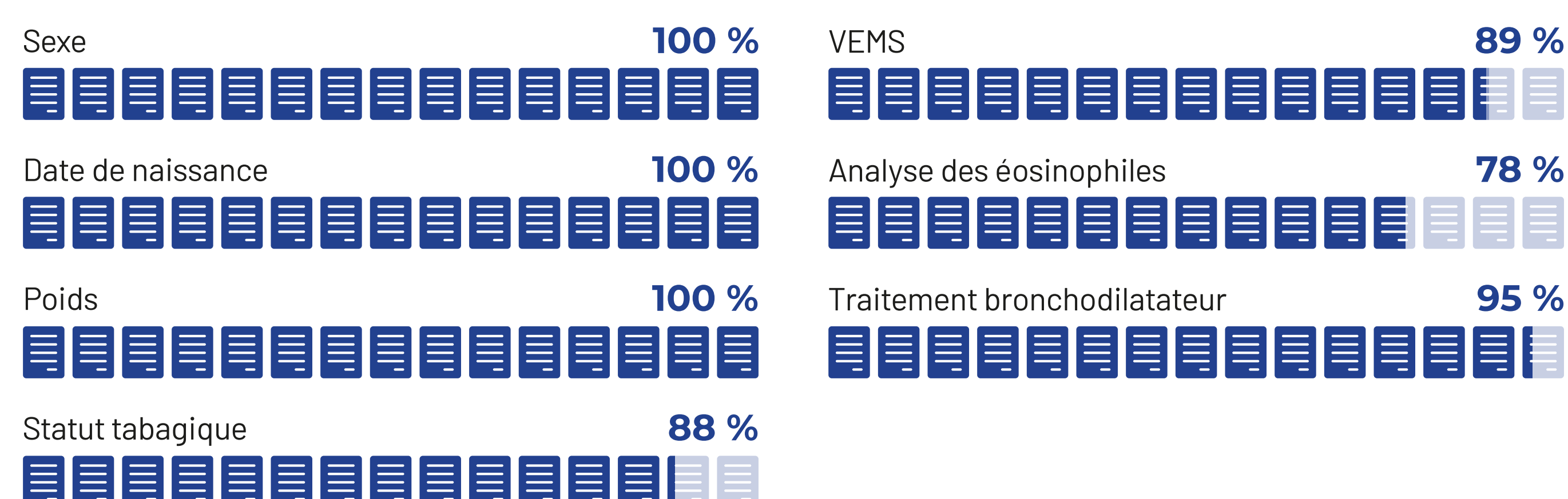
Inclusion et répartition des patients par pathologie

Sur les 100 patients inclus en juin 2025, 74 étaient atteints de BPCO et 40 d'asthme.



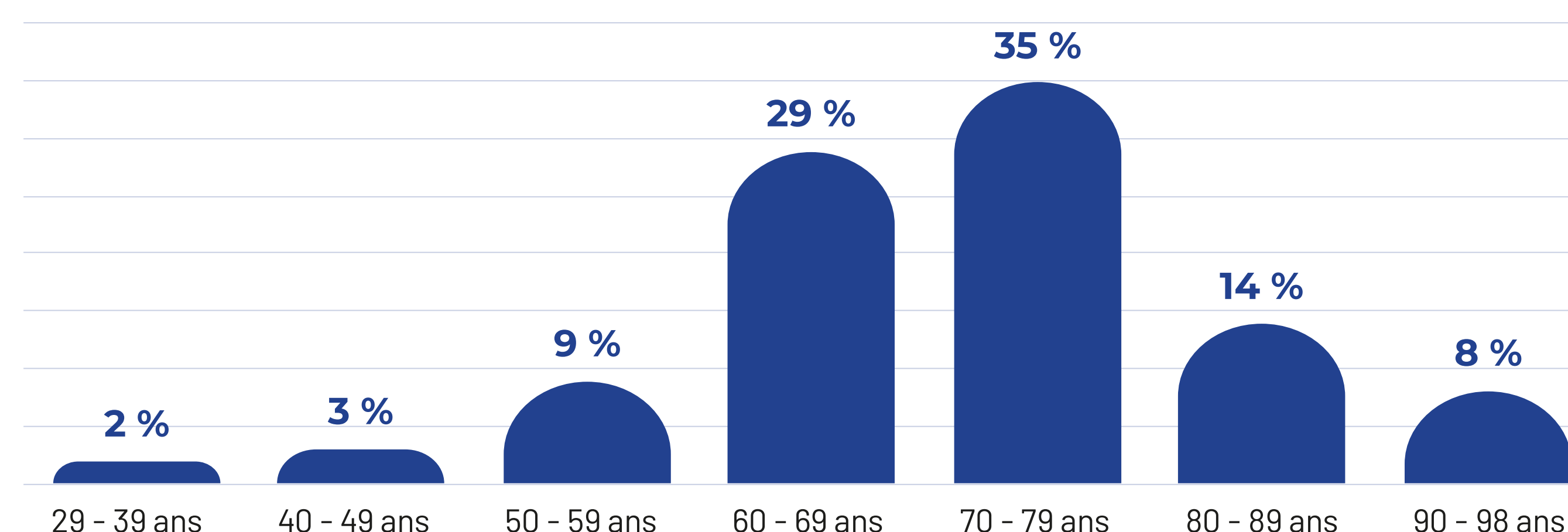
Exemple de variables clefs

La complétion des variables était globalement élevée. Parmi les 89 disponibles, le statut tabagique était renseigné pour 88 % des patients, le poids pour 100 %, et les paramètres respiratoires, dont le volume maximal expiratoire en une seconde (VEMS), étaient disponibles pour 89 % des patients.



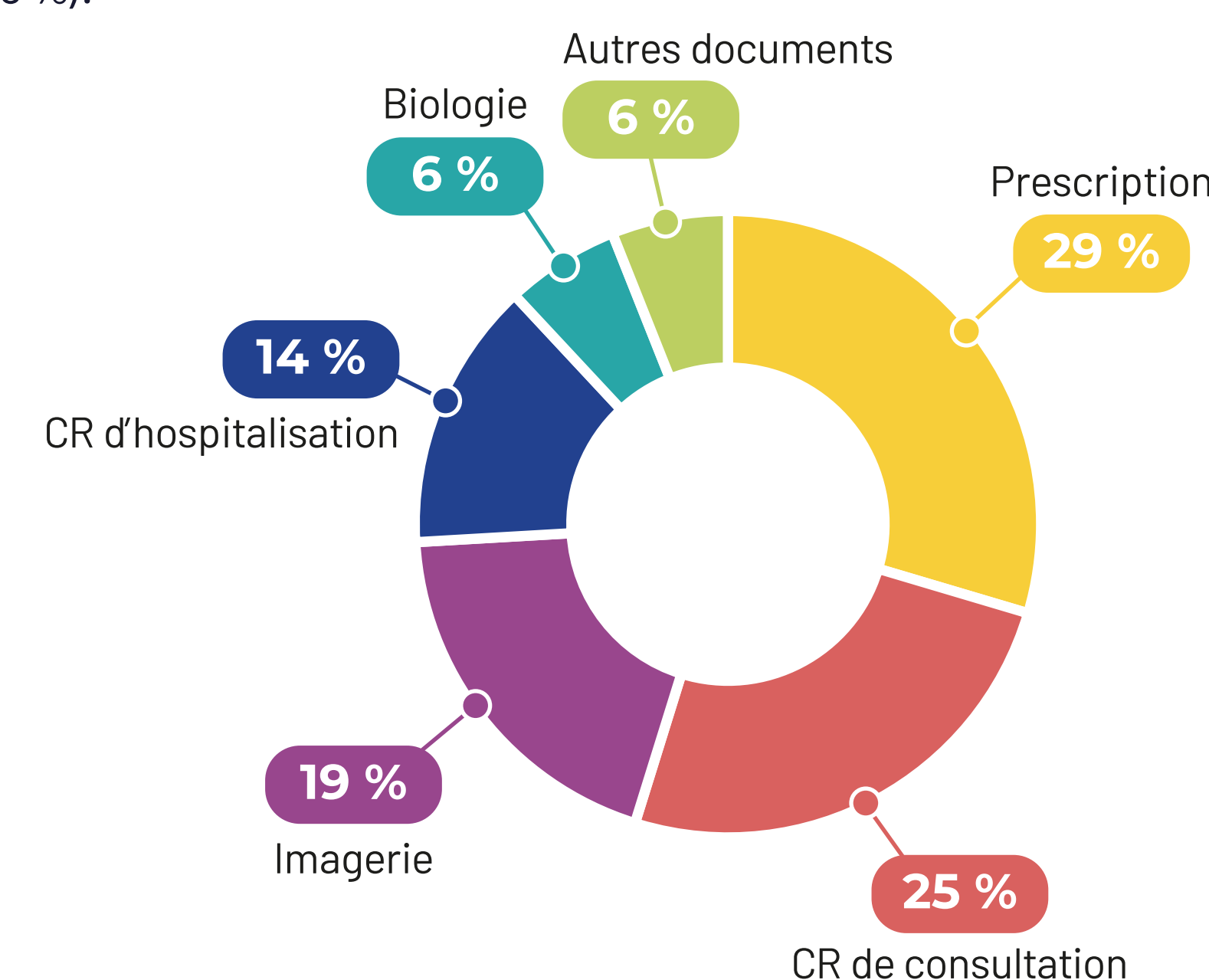
Âge des patients à l'inclusion

L'âge des patients à l'inclusion dans la cohorte était compris entre 29 et 98 ans. La majorité des patients avaient entre 60 et 80 ans au moment de l'inclusion.



Répartition du type de documents collectés

Les types de documents collectés étaient majoritairement des prescriptions (29 %) et des CR de consultations (25 %).



CONCLUSION

La cohorte CUB Trajectory illustre le potentiel des données en vie réelle pour mieux comprendre et décrire l'évolution des maladies bronchiques chroniques. Les premières données collectées permettent d'obtenir des résultats sur les caractéristiques des patients, leur suivi clinique et leurs prises en charge hospitalières qui permettront une analyse plus fine des trajectoires de soins et à l'identification de facteurs prédictifs d'évolution.