

RWE et données de santé : l'apport du SNDS

Dans le domaine des soins de santé, les **études en vie réelle** (Real-World Evidence, RWE) correspondent aux connaissances cliniques tirées de données du monde réel pour évaluer l'usage, l'efficacité et la sécurité des traitements en pratique courante.

Les analyses de ces études s'appuient sur des **données de santé** variées qui, d'après la Commission Nationale de l'informatique et des Libertés (CNIL), sont relatives à la santé physique ou mentale, passée, présente ou future, d'une personne, et révèlent des informations sur l'état de santé de cette personne.

Plusieurs sources de données de santé coexistent et peuvent se compléter, telles que : les cohortes de patients, les registres de pathologies, ou les entrepôts de données de santé.

En France, le Système National des Données de Santé (SNDS) représente la plus large ressource mobilisable pour réunir ce type de données. **Base médico-administrative publique** alimentée par les soins présentés au remboursement, il regroupe des **données pseudonymisées** (une donnée dont l'identifiant a été remplacé par un code, réidentifiable avec une clé séparée donc toujours soumise au RGPD) issues de la ville, de l'hôpital et des causes médicales de décès.

Le SNDS permet d'obtenir une vision exhaustive du parcours de soins de l'ensemble de la population française, avec une profondeur d'historique pouvant aller jusqu'à 20 ans. Au sein de cette base, il est notamment possible d'analyser :

- **Les soins en ville** : les données incluent les dates et lieux des consultations, les éventuels actes techniques réalisés, ainsi que les noms et quantités de médicaments dispensés en officine.
- **Les hospitalisations** : les données disponibles intègrent les dates de séjour, les services d'hospitalisation, les diagnostics et les actes médicaux réalisés. Il est également possible d'identifier la réalisation d'examens, bien que leurs résultats numériques ou textuels ne soient pas renseignés.
- **Les décès** : les dates ainsi que les causes médicales de décès sont enregistrées.

En RWE dans l'industrie pharmaceutique, l'ensemble de ces informations permettent, par exemple, de rapporter des données épidémiologiques sur une population cible dans l'optique du développement d'un médicament.

En contrepartie de sa richesse, l'accès au SNDS reste réglementairement exigeant et long. Pour l'industrie pharmaceutique, le processus requiert généralement un délai minimal d'un an et nécessite obligatoirement de recourir à un laboratoire de recherche

ou un bureau d'études indépendant. La CNIL instruit le dossier industriel et autorise l'accès aux données du SNDS, puis la Caisse Nationale de l'Assurance Maladie (CNAM) établit une convention et met à disposition les données.

Dès lors, **l'identification des données** pertinentes comporte des difficultés liées au data management :

- Extraction des variables : il est nécessaire d'isoler les variables d'intérêt via des nomenclatures spécifiques (CIM, CCAM, ATC...)
- Jointures de tables : L'analyse impose d'effectuer des jointures complexes entre plusieurs tables distinctes (séjours hospitaliers, actes, prescriptions) pour reconstituer l'ensemble du parcours de soins au niveau d'un même patient.
- Evolution du codage : Il est à noter que les pratiques de codage des données évoluent au fil des années et diffèrent entre établissements, augmentant le risque de mauvaise classification dans le SNDS. Ces biais engendrés doivent ainsi être soulignés dans les études RWE et, éventuellement, pris en compte par des analyses supplémentaires.

Par ailleurs, certaines **dimensions cliniques** sont peu renseignées ou indisponibles dans la base (résultats d'examens biologiques, état clinique...), obligeant à recourir à des proxys. En oncologie, par exemple, l'information de stade métastatique n'est pas disponible de manière précise ; on recourt dès lors à l'utilisation de médicaments caractéristiques de ce stade comme proxy. Il convient ainsi d'accepter une incertitude résiduelle lors de l'interprétation des résultats.

Une fois ces contraintes intégrées, le SNDS apporte l'échelle (au sens quantité de données), la représentativité et la continuité nécessaires pour conduire des études de vie réelle robustes et généralisables.