



**MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION**

*Liberté
Égalité
Fraternité*

CARTOGRAPHIER ET QUALIFIER LES BASES DE DONNÉES EN SANTÉ

Yohann Foucher – Chargé de mission « données de santé »

22 juin 2022

Pourquoi étendre les déclaration à toutes les études de santé ?

1. Une source d'erreur dans la prise de décision des scientifiques, entrepreneurs, financeurs, médecins ou décideurs de santé publique.

Exemple personnel :

Pretransplant sensitization against angiotensin II type 1 receptor is a risk factor for acute rejection and graft loss.

Giral M, Foucher Y, Dufay A, Duong Van Huyen JP, Renaudin K, Moreau A, Philippe A, Hegner B, Dechend R, Heidecke H, Brouard S, Cesbron A, Castagnet S, Devys A, Soullillou JP, Dragun D.

Am J Transplant. 2013 Oct;13(10):2567-76. doi: 10.1111/ajt.12397. Epub 2013 Aug 6.

PMID: 23919486 **Free article.**

The angiotensin II type 1 receptor (**AT1R**) is an emerging target of functional non-HLA antibodies (Ab). ...The study included 599 kidney recipients between 1998 and 2007. Serum samples were analyzed in a blinded fashion for anti-**AT1R** antibodies (**AT1R**-Abs) usin ...

Is pre-transplant sensitization against angiotensin II type 1 receptor still a risk factor of graft and patient outcome in kidney transplantation in the anti-HLA Luminex era? A retrospective study.

Deltombe C, Gillaizeau F, Anglicheau D, Morelon E, Trébern-Launay K, Le Borgne F, Rimbart M, Guérif P, Malard-Castagnet S, Foucher Y, Giral M.

Transpl Int. 2017 Nov;30(11):1150-1160. doi: 10.1111/tri.13009. Epub 2017 Oct 5.

PMID: 28686316 **Free article.**

Anti-**AT1R** antibody levels were measured by ELISA in pretransplantation sera of 940 kidney recipients from three French centers of the DIVAT cohort. ...Our study did not confirm an association between pretransplant anti-**AT1R** antibody levels and kidney transplant outc ...

Pourquoi étendre les déclaration à toutes les études de santé ?

2. Un frein aux projets collaboratifs et au partage des données.

En l'absence de registre unique et exhaustif, les académiques et les entreprises ne sont pas en capacité d'identifier les études qui pourraient faire l'objet de réutilisations.

Mis en place dans le cadre du Contrat Stratégique de Filière des Industries et Technologies de Santé, le GT« Usage des données de santé en vie réelle dans la prise de décision » recommande:

la création d'un registre déclaratif des bases de données de santé en vie réelle (cohortes, registres, etc.), incluant les indicateurs de qualité précédemment décrits.

Incitative de l'IRESF soutenu par le MESRI pour COVID-19

- Depuis 2011, le Portail Épidémiologie France (PEF) permet le recensement et la description de ces études.
- Le PEF a été développé par l'ITMO Santé publique de l'Alliance Aviesan.
- Il recense 1057 bases de données individuelles décrites par leurs responsables (hors essais thérapeutiques).
- 2000 visites mensuelles (dont 1/3 provenant de l'étranger).

En 6 mois, parmi les 477 études COVID-19 identifiées via différentes sources (CPP, CNIL, financeurs, etc.), seulement 75 (16%) sont déclarées et validées

Objectifs

- Rendre obligatoire la déclaration de toute collecte prospective de données individuelles en biologie-santé jusqu'aux études observationnelles.
- Proposer un portail de déclaration interopérable avec les solutions européennes et internationales.



OUVRIR
LA SCIENCE !





Proposer un portail de déclaration interopérable avec les solutions européennes et internationales

- Un outil bilingue FR / EN et interopérable pour éviter double saisie
- Un module de recueil systématique comprenant les items minimums des standards internationaux (OMS)
- Un module de recueil optionnel de métadonnées du plan de partage.
- Un module de recueil optionnel de métadonnées d'auto-évaluation de la qualité.
- Un module de requête, via une interface et une API, de toutes ces informations en libre accès (open data).

Proposer un portail de déclaration interopérable avec les solutions européennes et internationales

- Un outil bilingue FR / EN et interopérable pour éviter double saisie
- Un module de recueil systématique comprenant les items minimums des standards internationaux (OMS)
- Un module de recueil optionnel de métadonnées du plan de partage.
- **Un module de recueil optionnel de métadonnées d'auto-évaluation de la qualité.**
- Un module de requête, via une interface et une API, de toutes ces informations en libre accès (open data).

Membres du groupe de travail pour l'étiquetage de la qualité des bases de données en santé

Romain Casey (Docteur en Epidémiologie, Observatoire Français de la Sclérose en Plaques, Épidémiologiste)

Marie Zins et Marcel Goldberg (INSERM, Épidémiologistes, cohorte Constances)

Christophe Bonaldi (Santé Publique France, Statisticien)

Emmanuel Bacry (Health Data Hub, Directeur scientifique)

Emilie Lanoy (Institut Gustave Roussy, Statisticienne)

Paulo Carita (Sanofi, Économie de la santé)

Laurie Levy-Bachelot (MSD, Directrice des études en vie réelle)

Philippe-Jean Bousquet (INCA, Directeur de la Direction Observation, Science des données, Évaluation)

Objectif, méthode, et livrable principal

- Proposer un auto-questionnaire pour décrire la qualité de bases de données et éclairer les chercheurs et les industriels qui souhaitent exploiter ces données.
- Une bibliographie existante pour initier les réflexions :
 - « Good pharmaco-epidemiological practices » de la Société Internationale de Pharmaco-Epidémiologie (2016)
 - « Good epidemiological practices » de la Société Allemande d'Epidémiologie (2019)
 - Un travail proposé par Whitney et al. (1998)
 - Les pratiques suivies dans le cadre de la cohorte Constances (2016)
- Livrable: 26 questions fermées réparties en 5 axes

Axe #1: structuration pour l'ouverture à la recherche

1. Existe-t-il une publication définissant la base de données et/ou son protocole dans une revue à comité de lecture (critères d'inclusion, variables collectées, etc.) ?
2. La base de données est-elle partageable avec d'autres utilisateurs que votre équipe ?
3. Existe-t-il une procédure formalisée pour accéder à la base de données ?
4. Existe-il une instance scientifique pour l'évaluation des projets portant sur la base de données ?
5. Une équipe est-elle disponible pour accompagner l'utilisabilité des données (au minimum un data-manager) ?
6. Existe-il un dictionnaire des variables (au minimum leurs définitions, formats, et modes de collecte) ?

Axe #2: procédures de mise en qualité des données

1. Existe-t-il une formation du personnel de collecte ?
2. Existe-t-il un protocole formalisé pour standardiser la collecte ?
3. Existe-il des contrôles de cohérence à la collecte ?
4. Existe-il des audits de qualité internes ?
5. Existe-il des audits de qualité externes ?
6. Des classifications standardisées sont-elles utilisées (par exemple : CIM-10, questionnaires validés, etc.) ?
7. Les instruments de mesure sont-ils régulièrement standardisés, calibrés, vérifiés ?
8. Pour les collections biologiques, existe-t-il des consignes de traitement et de stockage des échantillons ?
9. Pour l'imagerie, existe-t-il des consignes de réalisation des clichés ?

Axe #3: traçabilité du processus qualité

1. Existe-il un cahier des charges de data management décrivant le processus de mise en qualité des données ?
2. Les modifications apportées aux données après la collecte sont-elles traçables ?
3. Peut-on accéder à la fois aux données brutes et retravaillées (par exemple : nettoyage, recodage, etc.) ?
4. Existe-t-il des indicateurs de qualité régulièrement calculés à partir des données collectées ?

Axe #4: normes et réglementation

1. La base de données est-elle engagée dans un processus normatif (par exemple : BPC, BPE, NF, ISO, etc.) ?

2. Le régime juridique de la base de données déclarée à la CNIL est :

- Recherche, étude ou évaluation ponctuelle
- Une demande de changement du statut d'étude à celui d'entrepôt est en cours

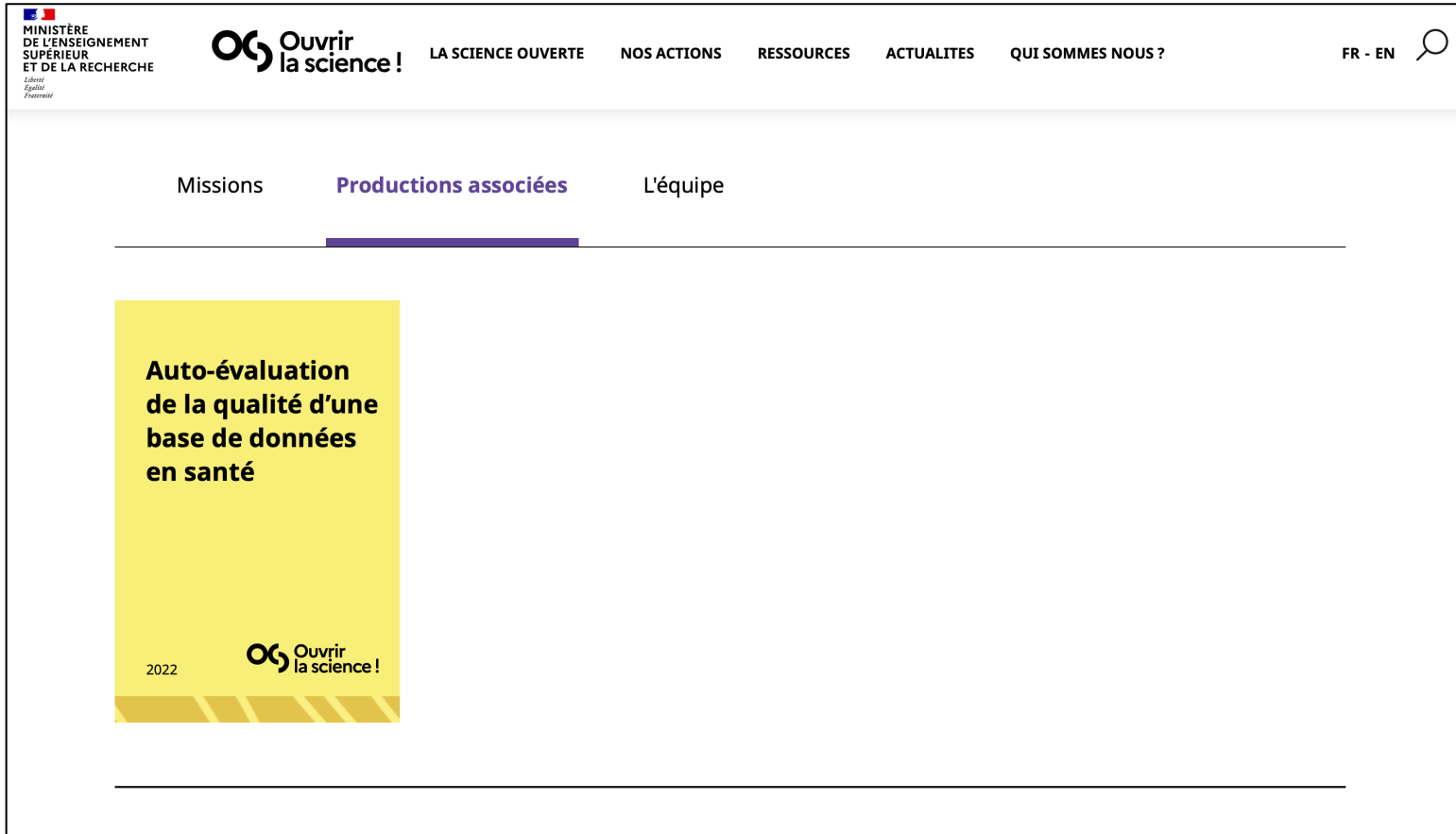
3. Est-ce que l'information à l'inclusion des sujets prévoit la réutilisation secondaire de leurs données ?

Si oui, les canaux d'information des patients sont-ils fonctionnels ?

Axe #5: évaluation externe, publications et animation

1. La base de données a-t-elle fait l'objet d'une évaluation par un organisme indépendant (par exemple : ANR, comité d'évaluation des registres, etc.) ?
2. La liste des projets initiés à partir de la base est-elle publiquement accessible ?
3. La liste des publications scientifiques issues de la base avec les principaux résultats est-elle publiquement accessible ?
4. Existe-il une communauté structurée d'utilisateurs de la base (par exemple : club utilisateurs, consortium) ?


<https://www.ouvrirlascience.fr/qualite-des-bases-de-donnees-en-sante/>



MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
*Liberté
Égalité
Fraternité*

Ouvrir la science !

LA SCIENCE OUVERTE NOS ACTIONS RESSOURCES ACTUALITES QUI SOMMES NOUS ?

FR - EN 

Missions **Productions associées** L'équipe

Auto-évaluation de la qualité d'une base de données en santé

2022 