

# Les cohortes en bref

par Ryma Hachi, cheffe de projet E-city



# qu'est-ce qu'une cohorte?

c'est "la réunion d'un groupe de sujets partageant ensemble un certain nombre de caractéristiques, suivis longitudinalement à l'échelle individuelle, selon un protocole pré-établi" (Spira, 2013)

"l'étude de cohorte est un terme qui a été largement utilisé à partir de la fin des années 1950 pour désigner des études portant sur des groupes d'individus "exposés" ou "non exposés" à des facteurs de risque, suivis, pour une période plus ou moins longue, de la mortalité ou de l'incidence des maladies" (Singh-Manoux, 2012)



# quel statut? quelle inscription dans les politiques de la recherche ?

Les cohortes sont des "**infrastructures de recherche**", ce qui reconnaît leur utilité au service d'une communauté de chercheurs et de la société. Les grandes cohortes sont des "très grandes infrastructures de recherche" (TGIR). (Spira, 2013)

Parmi les 5 priorités d'organisation proposées par l'**Institut de Santé Publique** : compléter le dispositif de GIR au service de la santé publique, celle-ci requiert souvent **la collecte et l'analyse de données concernant de grands échantillons de population ainsi qu'un suivi longitudinal sur une large période de cette collecte.** (Aviesan, 2013)



# les types de cohorte

- cohortes de malades souffrant d'une pathologie particulière (59%\*)
- cohorte en population générale (non malades)

(Goldberg, Zins, 2013)

Les cohortes en population générale peuvent être:

- des cohortes généralistes (couverture large de problèmes de santé et de déterminants, c'est le cas de la cohorte Constances) (10%\*)
- des cohortes thématiques, portant sur une ou plusieurs pathologies (31%\*)

\*Proportion de ce type de cohorte, en France en 2009, selon la base de données Epigramme (cité dans Spira, 2013)



# quelle finalité ?

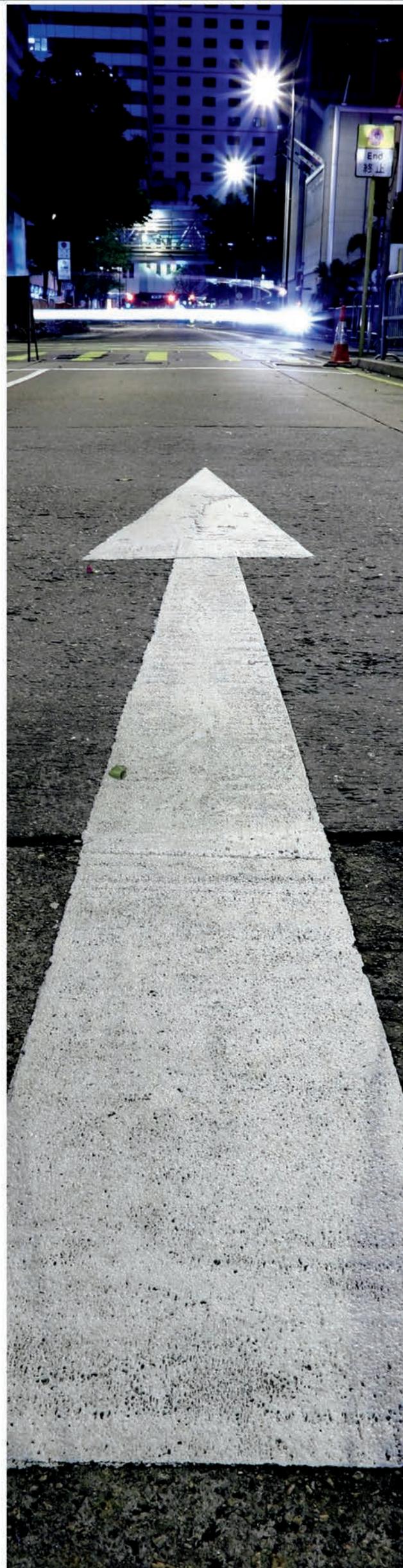
suivi et évaluation des actions dans le domaine de la santé (Spira, 2013)

elles sont "indispensables pour mieux comprendre les conditions d'émergence et d'évolution des maladies et produire des connaissances nouvelles indispensables à leur prévention et à leur prise en charge "

en santé publique, "contribuer à ce que les politiques de santé soient à l'avenir mieux fondées sur l'évidence scientifique qu'elles ne le sont actuellement, et ce faisant introduire plus de rationalité dans les débats sociaux sur ces sujets"

(Spira, 2013)

actuellement, apporter des réponses sur les causes des maladies (de nature environnementales, professionnelles, sociales, ...) en lien avec des risques de faible ampleur donc difficiles à mettre en évidence. (Singh-Manoux, 2012)



# quelle finalité ?

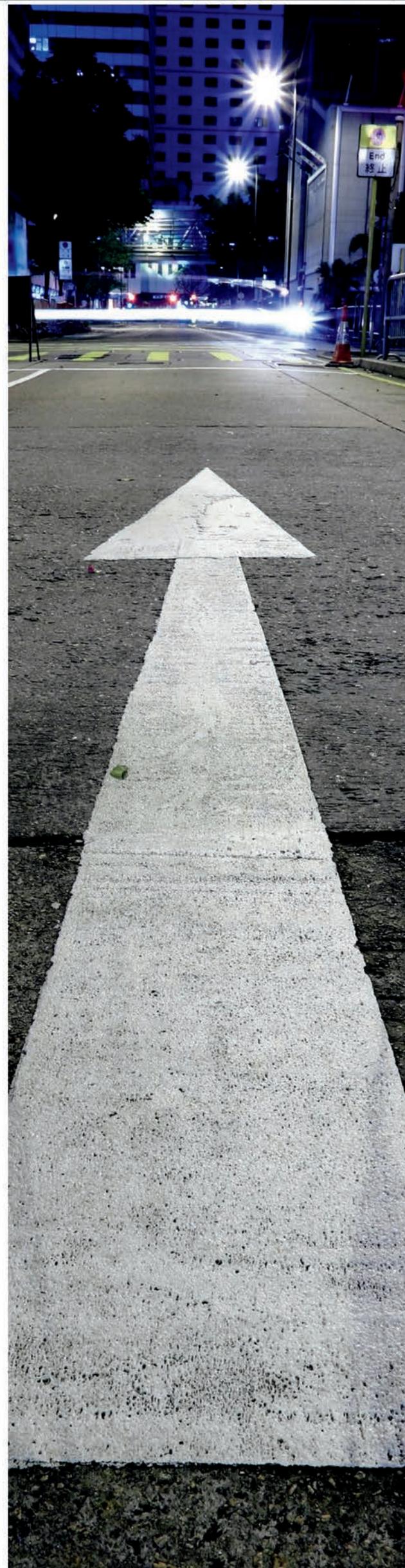
- cohortes de malades : étudier l'évolution d'une maladie. connaître "l'histoire naturelle des maladies"
- cohortes en population générale : causes des maladies, en particulier les plurifactorielles, aux déterminants environnementaux et génétiques multiples

(Goldberg, Zins, 2013)

- tenir compte de phénomènes liés au temps, comme la séquence temporelle exposition-effet
- modéliser l'enchaînement et les interactions des différents facteurs relatifs aux conditions de vie, à l'environnement et à l'état de santé

les cohortes proposent les meilleures conditions pour juger en termes de causalité du **rôle sur la santé des facteurs de risque**

(Goldberg, Zins, 2013)



# combien d'individus dans une cohorte ?

- pour les cohortes de malades : quelques milliers maximum
- pour les cohortes en population générale : parfois beaucoup plus

(Goldberg, Zins, 2013)

les cohortes de très grande envergure comprennent des dizaines de milliers de cas

(Singh-Manoux, 2012)

"plus les phénomènes d'intérêt (maladies, exposition) sont rares, plus les associations facteurs de risque / maladie sont faibles, et plus l'effectif doit être important pour une précision et une puissance données "  
c'est la raison pour laquelle les cohortes ne peuvent répondre à certaines questions

(Goldberg, Zins, 2013)



# comment sont recrutés les participants ?

- cohortes de malades : recrutés en milieu médical
- cohorte en population générale : on choisit un groupe de sujets indemnes de la (des) maladie(s) au début de la période d'observation

(Goldberg, Zins, 2013)

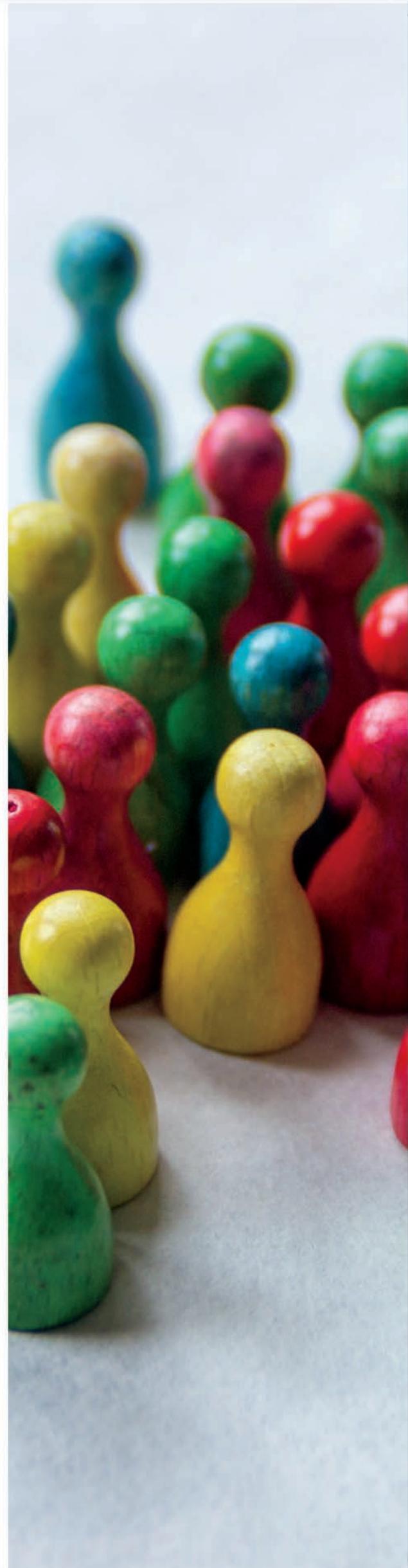
pour les cohortes en population générale, la participation repose sur le volontariat. Ex: pour la cohorte Constances, la Caisse nationale d'assurance vieillesse tire au sort dans ses bases de données un échantillon de personnes correspondant aux critères de sexe, d'âge et de catégorie sociale définis. Les personnes reçoivent un courrier avec un coupon-réponse (Goldberg, Zins, 2013)



# quelle représentativité?

- A priori, le volontariat n'introduit **pas de biais** car la relation exposition-maladie n'est a priori pas différente entre les sujets volontaires et ceux qui ne le sont pas.
- Cela ne signifie **aucunement que ces cohortes ont la prétention d'être représentative** de la population. Par exemple, dans l'étude de Framingham, la population étudiée n'était pas représentative de la population étasunienne (il s'agissait d'une ville riche); pourtant, les résultats de cette étude ont une valeur universelle.

(Goldberg & Zins, 2013)



# quelles données sont collectées ?

- des questionnaires socio-démographiques, comportementaux et cliniques,
- des mesures de caractéristiques physiques, bio-médicales, psychométriques ou/et d'attitudes
- examens cliniques et paracliniques
- prélèvements biologiques

(Spira, 2013)

- cohorte de malades : données très détaillées comme des investigations biocliniques approfondies
- cohorte en population générale : données personnelles, de mode de vie, sociales, professionnelles et environnementales. Biobanques. Données sur l'exposition à des facteurs de risques et incidence des maladies.

## Exemple Cohorte Constances:

- données de santé (antécédents, pathologies, comportements de santé ...)
  - recours au soin et prise en charge
  - caractéristiques socio-démographiques
    - facteurs professionnels
  - satisfaction d'un examen de santé
  - biobanque (sang, urine)
- (Goldberg, Zins, 2013)



# quelles données sont collectées ?

ex Constances: possibilité d'adapter le recueil de données à l'évolution des connaissances. Le dispositif doit permettre d'évoluer au cours du temps et d'intégrer des maladies au fur et à mesure de leur occurrence.

(Goldberg, Zins, 2013)

il faut "s'entendre précisément, dans tous les services où sont suivis ces patients, sur les indicateurs qui permettront de les décrire de manière identique. Puis mettre en place, malgré les différentes pratiques de travail, une même collecte d'information systématique et régulière."  
(source: cohorte sur les maladies inflammatoires du CHU et de l'université de Lille)

Dans la cohorte E-cohortE, qui porte sur la génétique, ils ont besoin d'extraire de l'ADN donc de faire des prélèvements, '<https://e-recherche.fr/wp-content/uploads/2019/09/18b-e-CohortE-French-version.pdf>)



# comment sont collectées les données ?

Ex cohorte Constances :

- suivi "actif" : auto-questionnaire annuel (suivi) + invitation au CES tous les 5 ans, où ils remplissent un questionnaire et bénéficient d'un examen de santé.
- suivi "passif" : grâce à l'appariement régulier de la cohorte avec les bases médico-administratives nationales, les principaux événements socio-professionnels sont extraits régulièrement.

(Goldberg, Zins, 2013)

Les méthodes classiques (interviews, entretiens téléphoniques, questionnaires papier) sont très **couteuses, contraignantes** pour les participants, et la qualité des données obtenues est **controversée** (Spira, 2013)



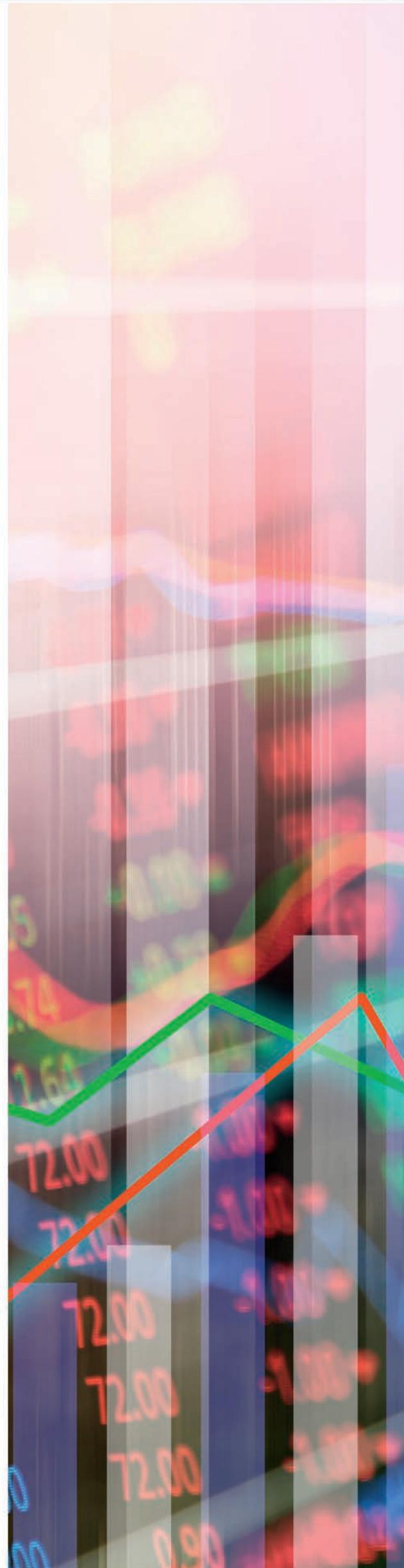
# quelles analyses sont faites sur les données collectées ?

- Méthodes multivariées.
- Utilisation de **marqueurs de progression** de la maladie mesurés au cours du temps, impliquant l'analyse de données longitudinales.

(Singh-Manoux, 2012)

Dans les cohortes en population générale, on dispose de toutes les données utiles pour calculer les **risques associés aux expositions**

Goldberg & Zins 2013



# quelle réglementation?

"toute collecte de données médicales est fortement encadrée : il faut obtenir l'autorisation de nombreuses instances - **comités de protection des personnes (CPP), CNIL, etc.**"

(source: cohorte sur les maladies inflammatoires du CHU et de l'université de Lille)

SI RIPH 2G ([sante.gouv.fr](http://sante.gouv.fr))

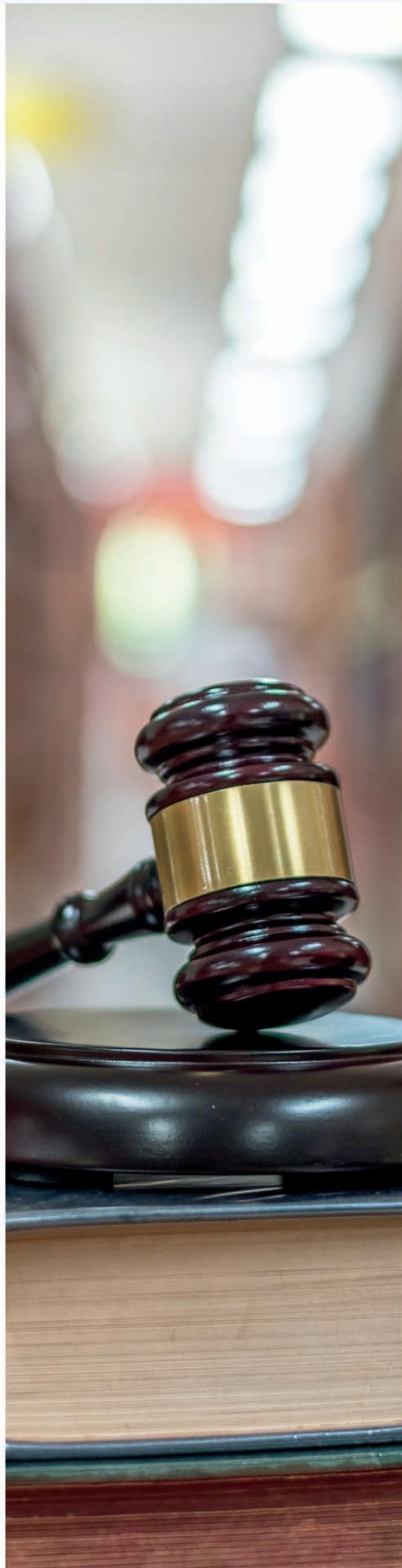
La loi dispose (article L. 1121-4 du CSP) que tous les projets de recherche impliquant la personne humaine (RIPH) ne peuvent être mis en œuvre qu'après un avis favorable (avis éthique) d'un comité de protection des personnes (CPP).

Le CPP compétent pour examiner chaque projet est désigné aléatoirement (article L. 1123-6 du CSP). Il existe 39 Comités de Protection des Personnes (CPP) répartis sur 7 interrégions.

Le système d'information des recherches impliquant la personne humaine (SI RIPH 2G) permet aux promoteurs d'un projet de RIPH de déposer leur dossier de demande d'avis pour leur projet de recherche et de les soumettre à un CPP désigné aléatoirement. Les échanges entre les promoteurs et les CPP sont ensuite gérés par ce même système d'information.

Ce système d'information permet la simplification et la fluidité de la gestion des dossiers de recherches impliquant la personne humaine.

e-CohortE est une RIPH, de catégorie 2 (Loi Jardé)



# quelle contrepartie pour les participants?

Le projet Plastico vise à faciliter l'accès, pour les sujets inclus dans la cohorte Constances, aux informations les concernant contenues dans les bases médico-administratives de l'Assurance maladie (Spira, 2013)

cohorte Constances : les participants bénéficient d'un examen de santé tous les 5 ans (Goldberg, Zins, 2013)

Dans la cohorte E3N, ils préconisent de donner des informations aux participants :

- infos individuelles (résultat de ses analyses de sang, suivi vaccins, données d'activité physique...)
- infos collectives : infos sur d'autres patients, comparaison.

Ils font cela sur le modèle de "Patients like me", un réseau social sur Internet où des personnes atteintes de certaines maladies spécifiques peuvent échanger des informations sur leur maladie, leur traitement et leur expérience.

(Actes du colloque "e-tools and social networks for epidemiology", 2013)



## *Cohorte et nouvelles technologies : adaptation des outils de recrutement, de suivi et de collecte des données au mode de vie de la population actuelle*

- **"Quantified self"**: faire des mesures et analyses de données sur soi-même
- Les **réseaux sociaux de patients**
- **Internet, les smartphones et les objets connectés** : fort potentiel pour la collecte des données

(Spira, 2013)

## *intérêt de ces alternatives*

- recruter et impliquer dans une cohorte un maximum de volontaires
- assurer l'adhésion des volontaires via les réseaux sociaux
- collecter des données (mode de vie, environnement, santé) de façon peu contraignante et ludique

PS. ce chantier est porté par la cohorte E4N (Spira, 2013)



- e-CE permettra alors la création d'une **interface informatique en ligne** (site web) permettant de « connecter » participants de la recherche, médecins, chercheurs investigateurs et coordonnateurs de l'étude **autour du stockage et de l'usage de données** (réponses aux questions, résultats des tests ; base de données) et **d'échantillons biologiques**. Chaque question aura sa traçabilité : provenance (laboratoire public, privé...), son état d'avancement (date de l'étude) et sa bibliographie (état de l'art de la question, légitimité de l'étude). Suite à ces informations, **chaque fois que les participants répondront à une question, ils verront apparaître les statistiques des réponses** (pourcentage de réponse de la cohorte) et plus tard accéderont au **résultat global de l'étude** (publications, séminaires, conférences et vulgarisation de la recherche). Ainsi, pour chaque réponse fournie, ils seront **informés dans le temps (via leur adresse email ou « push » via une application smartphone) des avancées de cette recherche ou si des questions complémentaires sont disponibles dans l'affinement des recherches.**"

source : <https://e-recherche.fr/wp-content/uploads/2019/09/18b-e-CohortE-French-version.pdf>



## quelles bonnes pratiques ?

- Propositions de la table ronde pour la mise en place de règles de bonnes pratiques visant à renforcer l'optimisation de la réalisation des cohortes épidémiologiques :

### 1. Établir l'inventaire des études existantes : un préalable indispensable

- Le portail « Épidémiologie France »[10] permet de réaliser un état des lieux de l'existant dans le domaine concerné, afin de juger si la constitution d'une nouvelle cohorte est nécessaire.
- De plus, parmi les questions de recherche formulées, toutes ne peuvent conduire à la mise en place de nouvelles études et a fortiori la constitution de cohortes ; le choix doit être fait selon des priorités nationales de santé publique

### 2. La cohorte doit s'inscrire dans un projet collectif selon des règles bien établies

- pour cela : "lister tous les acteurs (industriels, chercheurs et régulateurs) opérant dans le domaine cible du consortium. Après avoir informé les potentiels partenaires identifiés sur la nature du projet, l'étape suivante consistera à réaliser une enquête auprès d'eux afin de mieux connaître leurs priorités et attentes. Il conviendra par la suite d'organiser des réunions entre les représentants des différentes structures qui restent intéressées par le projet, ceci dans le but de clarifier leurs objectifs et/ou intérêts. Ces réunions sont aussi le moyen de réfléchir sur les aspects légaux et sur les contributions respectives

source : La mise en place des cohortes en France : pourquoi, pour qui, comment et avec quels moyens ? (2012)



## quelles bonnes pratiques ?

- élaboration d'une charte sur les modalités d'**ouverture des cohortes** à l'ensemble de la communauté scientifique
- réflexion sur la mise en place de **procédures mutualisées pour la validation** des données
- étude de faisabilité d'une plateforme pour l'**appariement des sujets** des cohortes avec les bases de données médico-administratives
- mise en réseau et **rassemblement** de cohortes autour de problématiques communes

la Cellule de coordination nationale des cohortes peut nous aider sur ces points.

(Spira, 2013)

- **mutualisation du personnel** entre les cohortes (mutualiser les moyens et les compétences)

(Spira, 2013)

- mettre des **outils de mesure** de domaines spécifiques (de nutrition, d'environnement, de déterminants sociaux de la santé) **à la disponibilité de l'ensemble des cohortes** susceptibles de les utiliser
- **standardisation de la mesure des diagnostics** de pathologies, pour faciliter les regroupements de données provenant de cohortes différentes

(Spira, 2013)



## quelles bonnes pratiques ?

- **mise en commun de données** de plusieurs grandes cohortes (dans le cadre de consortium de recherche?). Cela permet une puissance statistique.
- **harmonisation** : données ayant la même définition, variables enregistrées identiques (ou avec possibilité de correspondance).  
**Pourquoi ?** permet des études internationales comparatives réunissant des données de cohortes de plusieurs pays, ce qui permet d'aller au-delà des ensembles de données restreints et spécifiques d'une population

--> Le Consortium P3G (public population project in Genomics) "promeut l'utilisation d'outils (questionnaires, échelles diverses, etc.) qu'il met à disposition de la communauté scientifique".

- **ouverture** des bases de données des cohortes à l'ensemble de la communauté de recherche, car les investissements pour la construction et la maintenance des cohortes est tel qu'il n'est plus envisageable d'en limiter l'accès.  
**Pourquoi ?** les grandes cohortes sont conçues pour répondre à de multiples questions provenant d'équipes diversifiées

ne pas négliger :

- les aspects éthiques, juridiques et scientifiques
- la valorisation des équipes qui ont mis en place et gèrent la cohorte

(Singh-Manoux, 2012)



## quelles bonnes pratiques ?

- mise en place de cohortes nationales labellisées pour l'ensemble des grandes pathologies, et de centres de ressources **mutualisés pour le suivi** des grandes cohortes
- **appareiller** les données médico-administratives avec les données des GIR
- renforcer la valorisation économique

(Aviesan, 2013)

- "Des efforts importants ont été déployés pour **harmoniser la méthodologie des cohortes** au niveau européen afin d'ouvrir la possibilité d'agréger et de croiser les données recueillies, comme l'a indiqué le professeur Isabelle Momas lors de son audition. Les rapporteurs considèrent que ces démarches de convergence sont à poursuivre pour les projets de cohorte ayant vocation à pouvoir s'intégrer à un suivi épidémiologique de plus grande ampleur." (Jomier, Lassarade, 2021)



# quel financement des cohortes ? quel accompagnement ?

La "**Cellule de coordination nationale des cohortes**" analyse les besoins et propose "l'octroi de cofinancements pour des projets évalués en fonction de critères explicites" (Spira, 2013). L'auteur détaille ce que fait cette cellule p. 317.

ex Constances : plusieurs sources de financement : CNAMTS, Direction générale de la santé, Programme TGIR (Goldberg & Zins, 2013)

ex cohorte NutriNet-Santé : financée par Le Ministère des Affaires sociales et de la Santé, Santé Publique France, l'Université Paris 13, l'Inserm, l'Inra, le Cnam et le Centre de Recherche en Epidémiologie et Statistiques Sorbonne Paris Cité. (Etude NutriNet-Santé : la nutrition et la santé (etude-nutrinet-sante.fr).

Les Investissements d'Avenir financent des cohortes (Aviesan, 2013)

"Idéalement, les études de cohortes nécessitent 3 types de financement :

- un financement de lancement nécessaire à la mise en place de l'infrastructure de recherche ;
- un financement stabilisé de maintien de la cohorte sur une durée d'au moins 3 à 5 ans ;
- des financements sur projets pour toutes les études qui viennent se greffer au fil du temps sur les données de la cohorte initiale".

(Goehrs et al., 2012)



# quelle gouvernance dans les projets de cohorte ?

- ex Constances : "une **charte** a été établie en précisant les conditions d'utilisation de la cohorte, et un **appel d'offre** en direction des équipes scientifiques françaises et internationales permet aux chercheurs qui souhaitent bénéficier de la cohorte pour leurs propres travaux de **proposer des projets**, qui seront **examinés** par un comité scientifique international"

(Goldberg & Zins, 2013)



# quelques exemples de cohorte

- Cohorte Confins : sur le covid19
- Cohorte ESPERES | APHP : cohorte des personnels des établissements de santé

Etude longitudinale sur la santé des étudiants / Portail Epidemiologie - France | Health Databases (aviesan.fr) : cohorte "Étude longitudinale sur la santé des étudiants" (ELSE).  
Résumé de l'étude : Évolution des comportements à risque des étudiants pendant leur cursus : la cohorte ELSE - EM consulte (em-consulte.com)

- cohorte i-share :
  - Etude de cohorte sur la santé des étudiants | ANR, budget initial demandé de 17millions, ils ont finalement obtenu un financement PIA de près de 6 millions (en 2011). en 2021 : plus de 20 000 jeunes ont rejoint la cohorte.
  - i-Share : le projet « santé » qui a le vent en poupe - Université de Bordeaux (u-bordeaux.fr)
  - M-Soigner - Santé des étudiants : une étude unique
  - La santé de 30.000 étudiants suivis pendant 10 ans | Salle de presse | Inserm
- Cohorte E4N : La plus grande étude sur la santé au sein de la famille.
- Etude NutriNet-Santé : la nutrition et la santé (etude-nutrinet-sante.fr)



# quelques exemples de cohorte

- Cohorte CEMPUS - Le programme de doctorat en médecine - Université de Sherbrooke (usherbrooke.ca) : Cohorte initiée par l'Université de Sherbrooke, elle propose aux citoyens de prendre part à la recherche en matière de santé, en participant à des enquêtes. Il leur sera possible de répondre à des questions sur **leur mode de vie et leurs besoins en matière de santé** par une entrevue téléphonique. Les étudiantes et étudiants de 3e année du programme d'études médicales prédoctorales seront responsables du contact de ces participants. Ils compléteront les différents questionnaires téléphoniques avec eux et seront en mesure d'accroître les informations recueillies dans la banque de données ainsi créée.
- Lancement d'une cohorte exceptionnelle sur les maladies inflammatoires - Université de Lille (univ-lille.fr): le projet fédératif hospitalo-universitaire (FHU) Imminent (maladies inflammatoires et thérapies ciblées), financé par le CHU et l'Université de Lille, a pour but de suivre ensemble des patients atteints de ces différentes pathologies, pendant de nombreuses années. Cette « cohorte » comme l'appellent les médecins, vient d'inclure fin juillet (2020) son premier patient.



# quelques exemples de cohorte

- Home | HKUSPH FAMILY Cohort The FAMILY Cohort is a territory-wide cohort study in Hong Kong to understand the health, happiness and harmony - encompassing physical, mental and social wellbeing - of individuals and the households and neighbourhoods in which they live, in a society undergoing rapid social and economic change.
- E-cohortE <https://e-recherche.fr/wp-content/uploads/2019/09/18b-e-CohortE-French-version.pdf> : Projet de cohorte en génétique humaine, qui utilise les e-technologies et qui est multithématique
- Une cohorte portée par une chercheuse de l'UPEC : ffm - psy-cohorte - Programme dédié à la création d'une cohorte nationale sur les troubles bipolaires et la schizophrénie.



## Lexique

- cohortes rétrospectives = "elles analysent après coup les situations des patients de telle ou telle pathologie, en récupérant les informations dans les courriers des patients. Les données collectées de cette manière sont en général moins fiables et incomplètes". (source cohorte lancée en 2020 par l'université de Lille)



## *Bibliographie et Sites utiles*

- Spira A., Les grandes cohortes en santé en France, Bull. Acad. Natle Méd., 2013, 197, no 2, 315-328, séance du 19 février 2013.
- Singh-Manoux A., Les cohortes au niveau international, histoire et perspective, ADSP n° 78 mars 2012.
- Goldberg M., Zins M, Les cohortes généralistes en population. L'exemple des cohortes Gazel et Constances, Bull. Acad. Natle Méd., 2013, 197, no 2, 299-313, séance du 19 février 2013.
- Résumé de la contribution de l'ITMO Santé Publique d'Aviesan à la Stratégie Nationale de Recherche, mai 2013,
- Actes du colloque "e-tools and social networks for epidemiology", E4N, mai 2013,
- Jomier B., Lassarade F., Les orientations et la gouvernance de la politique de santé environnementale, Rapport d'information n° 479 (2020-2021) fait au nom de la commission des affaires sociales, Sénat, mars 2021.



## *Bibliographie et Sites utiles*

- Goehrs J.-M., Borel T., Costagliola D., et les participants à la table ronde N° 6 de Giens XXVI, La mise en place des cohortes en France: pourquoi, pour qui, comment, et avec quels moyens? *Thérapie* 2012 Juillet-Août; 67 (4).
- Cohorte sur les maladies inflammatoires, lancée en 2020 par le CHU et l'Université de Lille.
- Base de données Epigramme : recensement des cohortes en 2009
- Bienvenue / Portail Epidemiologie - France | Health Databases ([aviesan.fr](http://aviesan.fr))
- Cours sur Les études de cohorte, Université de Lorraine : Recherche clinique et épidémiologique - Les études de cohorte ([univ-lorraine.fr](http://univ-lorraine.fr))
- Dernières descriptions de bases de données publiées / Portail Epidemiologie - France | Health Databases ([aviesan.fr](http://aviesan.fr))

