

SÉMINAIRE MÉTHODOLOGIQUE

proposé par le service méthodes statistiques de l'Ined

24 septembre 2020
13h45-17h30
Ined

LES RENCONTRES DE STATISTIQUE APPLIQUÉE

DONNÉES DE SANTÉ : ENJEUX ET APPLICATIONS

statapp.site.ined.fr/

Les données de santé représentent une source de données majeure dans les études socio-épidémiologiques. Construites initialement dans une optique comptable, leur utilisation dans le domaine de la recherche s'avère complexe. Cependant l'accès des données de santé aux instituts de recherche représente une réelle opportunité pour les analyses, avec une perspective rétrospective de 5 à 20 ans, selon les applicatifs.

Nous nous proposons au cours de cette session, de préciser quelles sont les sources à disposition de la communauté scientifique, et de montrer des applications déjà entreprises dans ce domaine.

Programme et résumés des présentations

13h30	Arnaud Bringé (Ined) et Bénédicte Garnier (Ined) Accueil et informations
13h45	Chantal Cases (Insee) – présidente de séance Introduction de la séance
14h00	Jérôme Brocca (ARS Centre-Val de Loire) L'usage du Système National des Données de Santé (SNDS) pour un meilleur pilotage des politiques La connaissance de l'état de santé de la population est un enjeu majeur pour élaborer une politique de santé publique. Observer pour comprendre, comprendre pour agir ! Depuis quelques années, les systèmes d'information ont considérablement évolué, l'arrivée de données massives « Big data » a changé la donne en matière d'observation, en ouvrant le champ des possibles et en rendant accessible à un plus grand nombre des données souvent peu exploitées. En couvrant l'ensemble du territoire et en facilitant l'analyse individuelle du parcours de soins, le SNDS (système national des données de santé) constitue une avancée considérable pour la connaissance de l'état de santé de la population et la mesure de l'adéquation entre l'offre et le besoin de santé. L'objectif de cette présentation est d'illustrer, au travers d'exemples concrets l'opportunité réelle qu'offre cet outil, mis à disposition des acteurs institutionnels et des chercheurs, aussi bien en matière de pilotage opérationnel de l'action publique que de recherche, études et évaluation pour répondre aux enjeux de demain en matière de santé.
14h30	Romain Girard (Health Data Hub) Point d'étape du Health Data Hub La France dispose d'exceptionnelles bases de données de santé et cohortes, d'un écosystème hospitalier dynamique et de compétences de pointe en recherche médicale et mathématiques appliquées. Pour favoriser le développement des innovations au bénéfice des patients, nous devons décloisonner les sources et permettre un accès fluidifié et plus rapide aux données de santé. C'est pourquoi le Health Data Hub réunit les données de santé jugées prioritaires et les met à disposition sous format déidentifié aux porteurs de projet d'intérêt public dûment habilités par la CNIL. En investissant dans des technologies de pointe et des expertises rares, le ministère ambitionne de réduire les délais d'accès, soutenir les projets innovants souvent jugés infaisables aujourd'hui en France et contribuer

	à diffuser des bonnes pratiques en matière de protection des données personnelles et de cybersécurité. En lien avec la DREES, le Health Data Hub promeut également une dynamique de partage avec tous les acteurs de l'écosystème, afin de mutualiser les efforts et mettre en avant les contributions de chacun au service de tous.
15h00	<p>Louis Falissard (Univ. Paris Saclay / CépiDc INSERM) Apprentissage profond et reconnaissance d'entités médicales CIM10 à partir du langage naturel</p> <p>La reconnaissance d'entités médicales CIM10 à partir de données textuelles est une problématique omniprésente dans le domaine médical, du codage d'acte médical à l'analyse statistique de bases de données médico-administrative à des fins de santé publique. C'est en revanche un exercice complexe qui nécessite la plupart du temps une intervention humaine. Dans cette présentation, nous verrons comment s'appuyer sur les dernières avancées en intelligence artificielle (et tout particulièrement en apprentissage profond) pour développer un nouvel outil de codage automatique à partir des données du CépiDc (Centre d'épidémiologie sur les causes médicales de décès).</p>
15h40	<p>Khaoula Ben Messaoud (Ined /Univ. Paris Saclay, Univ. Paris-Sud, CESP, INSERM) Les barrières d'accès à la Fécondation In Vitro</p> <p>Les coûts des traitements de l'infertilité sont pris en charge à 100 % par l'assurance maladie en France. Cependant, la littérature internationale suggère l'existence possible d'un nonaccès à l'FIV, même en cas de couverture des coûts. Par ailleurs, les études internationales basés sur des données de centres d'AMP, peuvent induire des biais de sélection et limitent l'exploration des facteurs sociaux et territoriaux. Les données du SNDS apparaissent comme une opportunité d'explorer pour la première fois, l'accès à la FIV en population générale. Dans notre étude, nous avons mis en évidence que 70% des femmes en échec d'induction de l'ovulation n'accédaient pas à la FIV. Le désavantage social, la défavorisation de la zone de résidence, les âges jeunes et les âges avancés augmentent le risque de nonaccès à la FIV.</p>
16h10	<p>Emmanuel Oger (EA-7449 REPERES, Univ.Rennes 1) Risque de suicide et initiation d'un traitement pour l'acné : Exemple d'utilisation du SNDS</p> <p>Pour tenter de répondre à la question d'un lien potentiel entre l'initiation d'un traitement par isotrétinoïne pour une acné sévère et une tentative de suicide, il a été mis en œuvre une méthode d'analyse case-time-control à partir des données du SNDS (enrichies des causes de décès issues de la base gérée par le CEPIDC), avec une analyse complémentaire après chaînage entre une extraction physique (système-fils) du SNDS et la base de données OSCOUR gérée par Santé Publique France.</p>
16h40	<p>Pierre-Louis Bithorel (Ined) Panel diagnosis et données administratives de santé : quel intérêt ?</p> <p>L'exhaustivité et la granularité des données de santé leur confèrent un intérêt certain pour la recherche. Mais comment s'assurer de leur qualité à des fins autres qu'administratives ? La présentation se concentrera sur les problématiques rencontrées dans le cadre du projet StimHo et sur les solutions étudiées, en particulier sur le diagnostic de panel. Cette méthode, utilisée le plus souvent en épidémiologie, est utile pour construire un standard de référence en l'absence de <i>gold standard</i>.</p>
17h10	Discussion et conclusion