WP2 : ENTREPÔT DE DONNÉES INTÉGRÉ

Leader : Dr Anne-Laure Roux - Co-leader : Pr Martin Rottman

Partenaires : UVSQ/AP-HP I&D

Qu’est qu’un Data Warehouse ?

Les Data Warehouse ou entrepôts de données sont des bases de données permettant de stocker des données historiques structurées non volatiles orientées sujet afin de les analyser. Les données stockées dans une Data Warehouse présentent plusieurs spécificités. Elles sont séparées des systèmes opérationnels, mais aussi accessibles et disponibles pour les requêtes. Ces données sont orientées sujet et intégrées. Elles sont aussi nommées et définies de façon consistante. Par ailleurs, ces données sont

associées à des périodes de temps définies. Enfin, elles sont statiques (non volatiles), ce qui signifie qu'aucune mise à jour n'est effectuée sur ces données.

Pourquoi un Data Warehouse dans notre FHU ?

Les entrepôts de données revêtent actuellement une importance majeure dans différents domaines et notamment en médecine en permettant l'intégration d'énormes quantités de données provenant de sources différentes et dans des formats très hétérogènes. Dans le domaine du sepsis, les grands entrepôts de données font encore défaut. Cette plateforme innovante de gestion des données sera essentielle pour assurer la qualité et la valorisation des données et des produits biologiques de notre étude sur le sepsis.

Nous allons concevoir une Data Warehouse SEPSIS avec notre partenaire AP-HP (I&D - EdS) qui permettra la valorisation des données SEPSIS. Cet entrepôt de données sécurisé sera capable d'intégrer les données cliniques, biologiques et radiologiques des patients à partir de plusieurs sources de données.

La base de données sera d'abord initiée à partir des bases de données de cohortes existantes. Les données existantes seront complétées soit par une collecte de données prospectives, soit par une extraction automatisée et sécurisée à partir des logiciels de soins. L'URC HUPIFO et le département I&D du Département des systèmes d'information de l'AP-HP (DSI) assureront la consolidation, la structuration, la normalisation et la documentation (dictionnaire de données, mesures de qualité) des données. La politique de sécurité de l'AP-HP sera appliquée à l'entrepôt de données et à chaque flux de données. Le partenaire de l'AP-HP (I&D) fournira au partenaire WP8 des données intégrées pour le développement des signatures de réponse SEPSIS.

Cet entrepôt de données sera associé à un centre de ressources biologiques qui centralisera toutes les collections biologiques issues des différents partenaires.

Qu'est qu'un Centre de Ressources Biologique (CRB) ?

Un Centre de Ressources Biologiques (CRB) est une plateforme de gestion d'échantillons biologiques associés à des données qui peuvent être d'ordre clinique, pathologique, biologique et/ou moléculaire. Il a pour mission de préparer, de conserver et de mettre à disposition des échantillons biologiques de qualité au service des projets scientifiques des équipes de recherche et des éventuels partenaires académiques ou industriels. Les CRB doivent satisfaire aux critères élevés de qualité et d'expertise exigés par la communauté internationale des chercheurs, et par les industriels concernant la

diffusion d’informations et de matériels biologiques.

Les biobanques assurent la gestion complète des échantillons et des données associées, en encadrant les pratiques de collecte et de conservation avec des procédures standardisées. Pour valoriser leurs collections et ainsi multiplier les utilisations présentes et futures des ressources biologiques, les biobanques sont organisées en réseaux nationaux et internationaux.
Les échantillons biologiques conservés sont de tout type : tumeurs, blocs paraffine, sang, sérum, plasma, cellules ainsi que leurs produits dérivés -ADN, ARN.

Les biobanques sont aujourd’hui reconnues comme des catalyseurs de la recherche biomédicale. Le développement rapide de ces structures est la conséquence d’un besoin croissant d’échantillons biologiques de qualité. À l’origine "entrepôts" de matériels biologiques, elles sont devenues des infrastructures de recherche à part entière. Les biobanques représentent également un enjeu capital pour le développement de la médecine personnalisée. Elles sont devenues des supports essentiels contribuant à améliorer le diagnostic : l’accès à des ressources biologiques en grande quantité permet de cibler des mécanismes biologiques impliqués dans l’apparition de maladies telles que le cancer, les maladies inflammatoires et nous l’espérons le sepsis.

Cette nouvelle approche thérapeutique est fondée sur l’identification de biomarqueurs, permettant d’établir une prédiction de la réponse du patient à un traitement proposé. La découverte de nouveaux biomarqueurs candidats entraîne la mise en place d’études rétrospectives, permises par de grandes collections d’échantillons, accompagnées de données épidémiologiques, cliniques, et génomiques précises comme dans le cadre de notre FHU.

Contact

Dr Anne-Laure Roux
Pr Martin Rottman