

Actualités sur les maladies infectieuses: Mégadonnées

Ioannidis JPA, Khoury MJ. **Big data meets public health** Science 28 November 2014: Vol. 346 no. 6213 pp. 1054-1055. <http://www.sciencemag.org/content/346/6213/1054.full?sid=775d7875-0e89-4618-bd28-04fb6f240c6e>

Le terme mégadonnées désigne un ensemble d'information massive, complexe et reliée. Au-delà de la génomique et des autres domaines "omiques", les mégadonnées incluent des informations médicales, environnementales, financières, géographiques et reliées aux médias sociaux. La majorité des informations numériques n'était pas disponible une dizaine d'années passées. Le volume des données va continuer à croître et elles seront stockées par des sources qui sont actuellement inimaginables. Les mégadonnées contribuent à améliorer la santé en fournissant des connaissances sur les causes et les résultats des maladies, à mieux cibler les médicaments pour une médecine plus exacte, et à mieux prédire et prévenir les maladies. De plus, les citoyens scientifiques vont de plus en plus utiliser ces informations pour promouvoir leur propre santé et bien-être. Les mégadonnées peuvent améliorer notre compréhension des comportements liés à la santé... et accélérer le cycle de diffusion des connaissances... Mais la promesse des mégadonnées s'accompagne aussi d'affirmations que "la méthode scientifique en soi devient obsolète" (et) "une grosse erreur" peut tourmenter les mégadonnées.

Pigott DM, Kraemer MU. **Enhancing infectious disease mapping with open access resources** . Euro Surveill. 2014;19(49):pii=20989. <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20989>

Le groupe d'écologie spatiale et d'épidémiologie de l'université d'Oxford a rassemblé une quantité de bases de données à jour et mondialement reconnues à partir de trois sources: (i) recherche exhaustive sur PubMed, (ii) information provenant de sondages liés à la santé non publiés et des études entomologiques rendues disponibles par les collaborateurs, (iii) systèmes de surveillance des maladies sur internet tels HealthMap... Les bases de données sur les maladies et leurs vecteurs sont par la suite disponibles en accès libre à travers un dépositaire en ligne qui traite des maladies mentionnées... Les prochaines démarches incluent le développement de l'Atlas of Baseline Risk Assessment for Infectious Disease (ABRAID), une plateforme cartographiée et automatisée qui intègre le cadre de référence plus haut pour générer des cartes exhaustives, itérativement améliorées, et fondées sur des données probantes sur le risque de maladie à l'échelle mondiale pour des maladies infectieuses ciblées.

Iwasaki Y¹, Abe T, Wada Y, Wada K, Ikemura T. **Novel bioinformatics strategies for prediction of directional sequence changes in influenza virus genomes and for surveillance of potentially hazardous strains**. *BMC Infect Dis*. 2013 Aug 21;13:386. doi: 10.1186/1471-2334-13-386. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3765179/>

Avec la croissance remarquable des données sur la séquence bactérienne et virale obtenue à partir des séquenceurs à haut débit de l'ADN, de nouveaux outils sont requis pour une analyse exhaustive des données et de leur séquence. Nous avons développé "Batch-Learning Self-Organizing Map (BLSOM)" qui peut caractériser..... des millions de séquences génomiques sur un seul plan... Les études bioinformatiques sur les virus de l'influenza comportent des enjeux importants tels la prédiction des changements dans la séquence génomique dans un avenir rapproché et la surveillance des souches potentiellement dangereuses... Des millions de séquences génomiques des bactéries infectieuses et des virus sont devenues disponibles en raison de leur importance médicale et sociale, et BLSOM peut caractériser les mégadonnées et soutenir la découverte efficace de connaissances.