SNFMI
société nationale française de médecine interne

BIG DATA

& MALADIES RARES :
PARADOXES & OPPORTUNITÉS

Mercredi 12 Décembre 2018

Dans le cadre du 77ème congrès de la SNFMI, Alexpo, Grenoble

Shire

EXA/FR//135 – Décembre 2018

BIG DATA & MALADIES RARES : PARADOXES & OPPORTUNITÉS

Modéré par le **Pr Eric Hachulla,** CHU de Lille

12 DÉCEMBRE 2018 • 12H45 – 14H15

77ème congrès de la SNFMI

Amphithéâtre Pelvoux, Alexpo, Grenoble

L'utilisation des bases médico-administratives : une opportunité pour l'évaluation des maladies rares ?
Mr Alexandre Vainchtock

Que peut-on apprendre de vos patients sur les réseaux sociaux ?
Mr Stéphane Schuck

Dr Warehouse, un accélérateur pour la recherche translationnelle sur les maladies rares ?
Mr Nicolas Garcelon

Vos données personnelles font l'objet d'un traitement par SHIRE. Ces données seront utilisées par Shire France et ses sociétés affiliées pour l'exercice de leurs activités ainsi que pour la mise en œuvre des obligations de transparence. Elles peuvent être également transférées à des prestataires ainsi qu'aux instances ordinales et à toute instance habilitée à les recevoir. Certaines données peuvent être transmises à des destinataires ou des prestataires habilités situés en dehors de l'Union Européenne, Shire mettant alors en œuvre des mesures appropriées pour garantir un niveau de protection adéquat. Vous bénéficiez d'un droit d'accès et de rectification aux informations qui vous concernent, que vous pouvez exercer en envoyant une demande par e-mail à privacyconnect@shire.com. Vous pouvez par ailleurs, pour des motifs légitimes, vous opposer au traitement des données vous concernant – sauf dans le cas où ces données sont utilisées dans le cadre des obligations de SHIRE en matière de transparence tel que prévu à l'article L.1453-1 du Code de la santé publique.

Shire

EXA/FR//1261 – Octobre 2018

L’utilisation des bases médico-administratives : une opportunité pour l’évaluation des maladies rares ?

Alexandre VAINCHTOCK

https://www.hevaweb.com/fr

Sommaire

HEVA

1 LE SYSTÈME NATIONAL DES DONNÉES DE SANTÉ

2 LE POTENTIEL D'UTILISATION DES DONNEES

3 ENRICHISSEMENT DES DONNEES : PROCESS D'APPARIEMENT SNDS/BDD

4 HEALTH DATA HUB

HEVA

1

LE SYSTÈME NATIONAL DES DONNÉES DE SANTÉ

Accès régulé aux données du SNDS

HEVA

Le SNDS doit permettre à compter du 1er avril 2017 :

- L'ouverture des données de santé dans le respect de la vie privée du citoyen
- L'amélioration de la santé des citoyens et l'analyse de la dépense publique

Accès aux données

Toute personne ou structure, publique ou privée, à but lucratif ou non lucratif, peut accéder aux données du SNDS sur **autorisation de la CNIL**, en vue de réaliser une étude, une recherche ou une évaluation présentant un **intérêt public**.

LES FINALITÉS DE TRAITEMENT

HEVA

2 finalités interdites

- La promotion des produits de santé auprès des professionnels de santé et des établissements de santé
- L'exclusion de garanties des contrats d'assurance ou modification de cotisations ou primes d'assurance

6 finalités autorisées

- Information sur la santé, l'offre de soins, la prise en charge médico-sociale, la qualité des données
- Définition, mise en œuvre et évaluation des politiques santé et de protection sociale
- Connaissance des dépenses santé, d'assurance maladie et médico-sociales
- Information des professionnels de santé, des structures et des établissements sanitaires et médico-sociaux sur leur activité
- Surveillance, veille et sécurité sanitaire
- Recherche, études, évaluation, innovation dans les domaines de la santé et de la prise en charge médico-sociale

Composantes du SNDS

HEVA

SNIIRAM
Système national d'information inter-régimes de l'assurance maladie

Remboursement de soins de ville

+

PMSI
Programme de Médicalisation des Systèmes d'Information

Remboursement en établissement hospitalier

+

CépiDC
Centre d'épidémiologie sur les causes de décès

Cause de décès

SIMDPH Système d'Information des Maisons Départementales des Personnes Handicapées

Données relatives au handicap

Échantillon des données de **l'Assurance Maladie Complémentaire**

Données de remboursements de soins des complémentaires santé

N° d'anonymat (N° ANO) permet d'identifier les patients entre ces bases

Nouvelles bases d'ici 2020

SNIIRAM + PMSI + CépiDC = SNDS

HEVA

2

LE POTENTIEL D'UTILISATION DES DONNEES

Potentiels

HEVA

Médico-économie

- Poids économique
- Coûts de prise en charge

Epidémiologie et pharmacovigilance

- Études de cohorte
- Évaluation du risque
- Évaluation des produits de santé

Offre de soins

- Offre sur un territoire
- Comparaisons régionales

Parcours de soins

- Séquences de traitements
- Écarts par rapport aux recommandations

« Appariement » pour enrichir les données

- Direct (NIR du patient)
- Probabiliste

Innovations méthodologiques

- Machine Learning : identifier des <u>profils de patients</u>
- Process Mining : étudier les <u>parcours des patients</u>

HEVA

3

Enrichissement des données : process d'appariement SNDS/BDD

Appariement via le NIR / NIR INS

HEVA

Si la base est une base de données de soins et que le médecin/l'investigateur recueille le NIR du patient

Il doit avoir obtenu au préalable du recueil une autorisation de la CNIL pour recueillir le NIR (cela ne peut rentrer dans le cadre d'une MR car recueillir le NIR est un critère d'exclusion des MR). Ce NIR est appelé NIR-INS (NIR identifiant national de santé)

Le recours à un tiers d'appariement est obligatoire

Attention, ce NIR-INS ne doit en aucun cas figurer directement dans la base de données de soins, mais être répertorié dans une table à part avec la clé de jointure avec la base de données de soins.

Si la base est une base de données à vocation de recherche, et que l'investigateur recueille le NIR du patient,

Il doit avoir obtenu au préalable une autorisation de la CNIL. Ce NIR est appelé simplement NIR et non pas NIR INS.

Le recours à un tiers d'appariement n'est pas obligatoire; mais le recours à un tiers de confiance facilite la démarche

Pour pouvoir faire un appariement avec les données du SNDS, l'investigateur doit déposer un protocole au CEREES et à la CNIL en demandant l'appariement des données sans besoin de passer par un tiers de confiance. Après OK du CEREES et de la CNIL, l'investigateur transmettra directement la table des NIR à la CNAM pour procéder à l'extraction.

*A noter : **NIR-INS et NIR** sont le même identifiant, le nom change pour différencier le type de base dans lequel il est recueilli.*

Appariement probabiliste

HEVA

La procédure d'appariement probabiliste est une méthode d'appariement indirecte. Cette méthode doit être mise en place lorsque la base de données ne contient pas le NIR (ce dernier n'a pas fait l'objet d'une demande auprès de la CNIL lors de la constitution de la base).

La procédure doit être la suivante :

- Identifier les variables communes entre la base de données externe et les données du SNDS. Cette étape nécessite de disposer du cahier des charges de la collecte de la base externe
- Vérifier si les variables communes sont suffisamment discriminantes pour réidentifier un individu.

La procédure décrite à ce jour par la CNAM est la suivante :

- Envoi par la CNAM d'une clef de cryptage afin de procéder au cryptage des données externes
- Envoi par transfert sécurisé des données cryptées
- Décryptage de la base de données par la CNAM
- Appariement des variables par l'équipe DEMEX de la CNAM
- Extraction des données du SNIIRAM
- Cryptage des données extraites par la CNAM
- Envoi des données cryptées à HEVA ainsi que la clef de cryptage (sécurisé)
- HEVA héberge les données au sein de sa bulle sécurisée et procède au décryptage des données via la clef communiquée par la CNAM

HEVA

4

Health Data Hub

Le potentiel français en santé

HEVA

- Une recherche d'excellence en mathématiques et en médecine
- Ressources riches :
  Données de Santé, Hébergeurs,
- Une demande forte des acteurs publics et privés

- Cloisonnement des infrastructures
- Cadre d'échange peu défini nuit au partage de données
- Cloisonnement des savoirs

Le Health Data Hub

Conclusion

HEVA

1/ Une richesse inouïe de datas

- Multiples sources : données médico-administratives, registres, études cliniques
- Vie réelle
- Exhaustives
- Complexes

→ **Une révolution dans l'évaluation des produits de santé, des parcours de soins, de la pertinence des actes, etc.**

→ **Perspective de l'article 42 du PLFSS 2019**

2/ Indissociable d'une expertise collégiale

- Cliniciens
- Codage des données
- Data scientist

Que peut-on apprendre de vos patients sur les réseaux sociaux ?

Dr Stéphane Schück

contact@kapcode.fr
55, rue de La Boétie 75008 PARIS

NOUVEAUX TYPES DE DONNEES AU SERVICE DE LA SANTE PUBLIQUE : LES APPORTS DE L'INFODEMIOLOGIE

Dr Stéphane Schück
CEO Kappa Santé / Kap Code

KAP CODE
CONNECT YOUR HEALTH

L’ESSOR DES RÉSEAUX SOCIAUX

SONDAGE ODOXA

- Réalisé pour le Healthcare Data Institute
- Echantillon représentatif de 987 personnes
  - 29% des Français affirment avoir déjà évoqué un sujet lié à leur santé sur Internet.
    - Les plus jeunes (48% des 18-24 ans et 41% des 25-34 ans) ainsi que les femmes (32%) sont les catégories qui hésitent le moins à évoquer le sujet sur le web.
  - Pour 45% des Français, les témoignages de patients en ligne ont de la valeur pour la recherche
  - Un Français sur deux déclare qu'il serait prêt à faire un don à la recherche de ses témoignages et données de santé.

« SOCIAL MEDIA » DANS PUBMED

GOOGLE FLU TRENDS

nature International weekly journal of science

Access
To read this story in full you will need to login or make a payment (see right).

nature.com > Journal home > Table of Contents

Letter

Nature **457**, 1012-1014 (19 February 2009) | doi:10.1038/nature07634; Received 14 August 2008; Accepted 13 November 2008; Published online 19 November 2008; Corrected 19 February 2009

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg[1], Matthew H. Mohebbi[1], Rajan S. Patel[1], Lynnette Brammer[2], Mark S. Smolinski[1] & Larry Brilliant[1]

1. Google Inc., 1600 Amphitheatre Parkway, Mountain View, California 94043, USA
2. Centers for Disease Control and Prevention, 1600 Clifton Road, NE, Atlanta, Georgia 30333, USA

Correspondence to: Matthew H. Mohebbi[1] Correspondence and requests for materials should be addressed to J.G. or M.H.M. (Email: flutrends-support@google.com).

Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year[1]. In addition to seasonal influenza, a new strain of influenza virus against which no previous immunity exists and that demonstrates human-to-human transmission could result in a pandemic with millions of fatalities[2]. Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza[3...4]. One way to improve early detection is to monitor health-seeking behaviour in the form of queries to online

Top

ARTICLE LINKS
- Figures and tables
- Supplementary info

SEE ALSO
- Editor's Summary

ARTICLE TOOLS
- Send to a friend
- Export citation
- Export references
- Rights and permissions
- Order commercial reprints

SEARCH PUBMED FOR
- Jeremy Ginsberg
- Matthew H. Mohebbi
- Rajan S. Patel
- Lynnette Brammer
- Mark S. Smolinski
- Larry Brilliant
- more authors of this article

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

Large errors in flu prediction were largely avoidable, which offers lessons for the use of big data.

David Lazer,[1,2*] Ryan Kennedy,[1,3,4] Gary King,[3] Alessandro Vespignani[5,6,3]

In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that Google executives or the creators of the flu tracking system would have hoped. *Nature* reported that GFT was predicting more than double the proportion of doctor visits for influenza-like illness (ILI) than the Centers for Disease Control and Prevention (CDC), which bases its estimates on surveillance reports from laboratories across the United States (*1*, *2*). This happened despite the fact that GFT was built to predict CDC reports. Given that GFT is often held up as an exemplary use of big data (*3*, *4*), what lessons can we draw from this error?

The problems we identify are not limited to GFT. Research on whether search or social media can predict *x* has become commonplace (*5*–*7*) and is often put in sharp contrast with traditional methods and hypotheses. Although these studies have shown the value of these data, we are far from a place where they can supplant more traditional methods or theories (*8*). We explore two issues that contributed to GFT's mistakes—big data hubris and algorithm dynamics—and offer lessons for moving forward in the big data age.

The initial version of GFT was a particularly problematic marriage of big and small data. Essentially, the methodology was to find the best matches among 50 million search queries to fit 1152 data points. The ability and dependencies among data (*12*). The core challenge is that most big data that have received popular attention are not the output of instruments designed to produce valid and reliable data amenable for scientific analysis.

run ever since, with a few changes announced in October 2013 (*10*, *15*).

Although not widely reported until 2013, the new GFT has been persistently overestimating flu prevalence for a much longer time. GFT also missed by a very large margin in the 2011–2012 flu season and has missed high for 100 out of 108 weeks starting with August 2011 (see the graph). These errors are not randomly distributed. For example, last week's errors predict this week's errors (temporal autocorrelation), and the direction and magnitude of error varies with the time of year (seasonality). These patterns mean that GFT overlooks considerable information that could be extracted by traditional statistical methods.

Even after GFT was updated in 2009, the comparative value of the algorithm as a stand-alone flu monitor is questionable. A study in 2010 demonstrated that GFT accuracy was not much better than a fairly simple projection forward using already available (typically on a 2-week lag) CDC data (*4*). The comparison has become even worse since that time, with lagged models significantly outperforming GFT (see the graph). Even 3-week-old CDC data

DETEC't

ODE
HEALTH

RESEAUX SOCIAUX ET EPIDEMIES

Using Friends as Sensors to Detect Global-Scale Contagious Outbreaks

Manuel Garcia-Herranz[1*], Esteban Moro[2,3], Manuel Cebrian[4,5,6], Nicholas A. Christakis[7,8,9], James H. Fowler[10,11]

1 Department of Computer Science, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain, 2 Department of Mathematics & GISC, Universidad Carlos III de Madrid, Leganés, Spain, 3 Instituto de Ingeniería del Conocimiento, Universidad Autónoma de Madrid, Madrid, Spain, 4 Computer Science & Engineering Department, University of California San Diego, San Diego, California, United States of America, 5 Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, 6 National Information and Communications Technology Australia, Melbourne, Victoria, Australia, 7 Department of Sociology, Yale University, New Haven, Connecticut, United States of America, 8 Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, United States of America, 9 Department of Medicine, Yale School of Medicine, New Haven, Connecticut, United States of America, 10 Medical Genetics Division, School of Medicine, University of California San Diego, San Diego, California, United States of America, 11 Political Science Department, University of California San Diego, San Diego, California, United States of America

Abstract

Recent research has focused on the monitoring of global–scale online data for improved detection of epidemics, mood patterns, movements in the stock market political revolutions, box-office revenues, consumer behaviour and many other important phenomena. However, privacy considerations and the sheer scale of data available online are quickly making global monitoring infeasible, and existing methods do not take full advantage of local network structure to identify key nodes for monitoring. Here, we develop a model of the contagious spread of information in a global-scale, publicly-articulated social network and show that a simple method can yield not just early detection, but advance warning of contagious outbreaks. In this method, we randomly choose a small fraction of nodes in the network and then we randomly choose a friend of each node to include in a group for local monitoring. Using six months of data from most of the full Twittersphere, we show that this friend group is more central in the network and it helps us to detect viral outbreaks of the use of novel hashtags about 7 days earlier than we could with an equal-sized randomly chosen group. Moreover, the method actually works better than expected due to network structure alone because highly central actors are both more active and exhibit increased diversity in the information they transmit to others. These results suggest that local monitoring is not just more efficient, but also more effective, and it may be applied to monitor contagious processes in global–scale networks.

Citation: Garcia-Herranz M, Moro E, Cebrian M, Christakis NA, Fowler JH (2014) Using Friends as Sensors to Detect Global-Scale Contagious Outbreaks. PLoS ONE 9(4): e92413. doi:10.1371/journal.pone.0092413

Editor: José Javier Ramasco, Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Spain

ARTICLE // Article

TWITTER : UN OUTIL COMPLÉMENTAIRE POUR LA SURVEILLANCE DE L'ÉPIDÉMIE SAISONNIÈRE DE GRIPPE EN FRANCE MÉTROPOLITAINE ET EN RÉGION ?

// TWITTER: A COMPLEMENTARY TOOL FOR MONITORING THE SEASONAL INFLUENZA EPIDEMIC IN METROPOLITAN FRANCE AND THE REGIONS?

Pascal Vilain (pascal.vilain@santepubliquefrance.fr), Laurent Filleul

Santé publique France, Cellule d'intervention en région Océan Indien, Saint-Denis, La Réunion, France

Soumis le 23.03.2018 // *Date of submission: 03.23.2018*

Résumé // Abstract

Introduction – Aujourd'hui, les médias sociaux comme Twitter sont utilisés par les individus pour diffuser de l'information en santé mais également pour partager ou échanger sur leur état de santé. Dans ce contexte, une étude exploratoire a été mise en place afin d'étudier si les données de Twitter peuvent être utilisées comme un proxy pour la surveillance de l'épidémie saisonnière de grippe en France, et plus particulièrement à l'échelle régionale.

Méthodes – Un système automatisé permettant le recueil en temps réel et un prétraitement (géolocalisation et classification) de tweets relatifs à la grippe a été développé. Les données recueillies sur la période du 8 août 2016 au 26 mars 2017 ont ensuite été comparées à celles du système de surveillance syndromique SurSaUD® (réseaux OSCOUR® et SOS Médecins).

Résultats – Au total, sur la période d'étude, le système a permis de recueillir 238 244 tweets relatifs à la grippe, dont 130 559 étaient localisés en France. Après une étape de nettoyage, l'algorithme a permis de classer 22 939 tweets indiquant un état grippal. Une corrélation positive et statistiquement significative a été observé entre les tweets indiquant un état grippal et les consultations pour grippe/syndrome grippal rapportées par les réseaux OSCOUR® et SOS Médecins, à la fois au niveau national et au niveau régional.

Discussion – conclusion – Cette étude exploratoire a permis de montrer que les données de Twitter, en complément des systèmes existants, permettent le suivi de l'épidémie saisonnière de grippe en France et en région. Le système devra être amélioré pour confirmer les tendances observées lors de la prochaine épidémie de grippe.

Introduction – Social media as Twitter are used today by people to disseminate health information, but also to share or exchange on their health condition. In this context, an exploratory study was implemented to investigate whether Twitter data can be used as a proxy for monitoring the seasonal influenza epidemic in France and at the regional level.

Methods – *A real-time automated system allowing the collection, the pre-processing (geolocation and classification) of tweets related to influenza illness was developed. The data collected from 8 August 2016 to 26 March 2017 were compared to those of the French syndromic surveillance system SurSaUD® (OSCOUR® and SOS Médecins networks).*

DETEC't

KAP CODE
CONNECT YOUR HEALTH

RESEAUX SOCIAUX ET MORTALITE

PNAS

Online social integration is associated with reduced mortality risk

William R. Hobbs[a,b,c,1], Moira Burke[d], Nicholas A. Christakis[e,f,g,h], and James H. Fowler[a,i,1]

[a]Division of Social Sciences, University of California, San Diego, CA 92093; [b]Network Science Institute, Northeastern University, Boston, MA 02115; [c]Institute for Quantitative Social Science, Harvard University, Cambridge, MA 02138; [d]Facebook, Menlo Park, CA 94025; [e]Department of Sociology, Yale University, New Haven, CT 06520; [f]Department of Medicine, Yale University, New Haven, CT 06520; [g]Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520; [h]Department of Biomedical Engineering, Yale University, New Haven, CT 06520; and [i]School of Medicine, University of California, San Diego, CA 92093

Edited by Kenneth W. Wachter, University of California, Berkeley, CA, and approved September 15, 2016 (received for review April 6, 2016)

Social interactions increasingly take place online. Friendships and other offline social ties have been repeatedly associated with human longevity, but online interactions might have different properties. Here, we reference 12 million social media profiles against California Department of Public Health vital records and use longitudinal statistical models to assess whether social media use is associated with longer life. The results show that receiving requests to connect as friends online is associated with reduced mortality but initiating friendships is not. Additionally, online behaviors that indicate face-to-face social activity (like posting photos) are associated with reduced mortality, but online-only behaviors (like sending messages) have a nonlinear relationship, where moderate use is associated with the lowest mortality. These results suggest that online social integration is linked to lower risk for a wide variety of critical health problems. Although this is an associational study, it may be an important step in understanding how, on a global scale, online social networks might be adapted to improve modern populations' social and physical health.

social networks | social media | health | longevity | social support

measures were associated with decreased mortality risk in a two year follow-up.

Results

Before analyzing online social connection and social media behavior, we compare mortality rates for the Facebook population vs. the population-at-large. In these analyses, we control for age and gender differences between the two groups, as well as a coarse proxy for race/ethnicity [based on data from the US Census Bureau (11); results are shown in *SI Appendix*] to account for known health disparities and slightly different levels of Facebook use by race/ethnicity. The age- and gender-matched mortality rate for the "full" population of Facebook users (*Materials and Methods*) was 63% of the rate in the California voter record (our data-matching benchmark). This association may result from difficulties in matching Facebook users to vital records. To more confidently evaluate the relative health of Facebook users compared with the general population, we focus our analysis on the "voter" subpopulation, which includes only those Facebook users also present in the California voter record. That is, we compare voters who are on Facebook to those who are not. The age- and gender-matched

Significance

People who have stronger social networks live longer. However, can we say the same about online social networks? Here, we conduct such a study. Using public California vital records, we compare 12 million Facebook users to nonusers. More importantly, we also look within Facebook users to explore how online social interactions—reflecting both online and offline social activity—are associated with longevity. We find that Facebook users who accept more friendships have a lower risk of mortality, but there is no relationship for those who initiate more friendships. Mortality risk is lowest for those with high levels of offline social interaction and moderate levels of online social interaction.

RÉSEAUX SOCIAUX ET SURVEILLANCE SANITAIRE

The use of social media in public health surveillance

Isaac Chun-Hai Fung,[a] Zion Tsz Ho Tse[b] and King-Wa Fu[c]
Correspondence to Isaac Chun-Hai Fung (email: cfung@georgiasouthern.edu).

Social media are broadly understood as a set of online activities that facilitate interpersonal communication, information sharing, collaboration or crowdsourcing among online users. They have become a global phenomenon with over two thirds of worldwide adult Internet users being active on social networking sites in 2014.[1]

Social media are increasingly harnessed for public health and can be used as communication tools to disseminate disease risks and interventions and to promote healthy lifestyles and health policies. There is also the potential use of social media as data sources for public health surveillance. While social media will likely never replace traditional data sources for disease surveillance, they can provide complementary information. However, social media data are, in essence, observational data of online communications and were not designed for public health purposes. Analyses of social media data are subject to limitations that are generally associated with observational studies, i.e. possible confounding factors and no causal conclusion.

Following is a brief overview of some of the uses of social media data for public health surveillance and some of the data's strengths and limitations.

USING SOCIAL MEDIA FOR PUBLIC HEALTH SURVEILLANCE

There are three major applications for social media in public health surveillance: epidemiologic monitoring and surveillance, situational awareness during emergency response and communication surveillance (Table 1).

Epidemiologic monitoring and surveillance

For epidemiologic monitoring and surveillance, social media can be used to perform three specific functions:

Monitoring and retrieving official information

Public health officials use social media to monitor official information released by foreign authorities and to monitor domestic official accounts as these can be more timely, which is important in emergency responses (Table 1).[2,3]

Disease detection

Social media and other population-based digital platforms provide additional data sources for public health surveillance to detect disease outbreaks and estimate disease incidence. Syndromic surveillance can be undertaken by detecting symptoms disclosed by individuals on social media for non-public health purposes by either human readers or computer algorithm,[4] or through participatory epidemiology where applications allow participants to self-report their symptoms to disease-specific digital surveillance systems.[5,6] Diseases may also be detected via event-based surveillance as unofficial information or rumours about a new disease may circulate on social media, as in the case of a medical record of an H7N9 patient uploaded to Weibo in 2013.[7] Print media, radio and television may pick these up and generate news stories that are also circulated on social media and may be detected by event-based surveillance systems that monitor digital news feeds for disease news (e.g. HealthMap).[8] Digital data sources also provide epidemiologists with additional means to detect, investigate and verify outbreaks.

[a] Department of Epidemiology, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, Georgia 30460-8015, United States of America.
[b] College of Engineering, The University of Georgia, Athens, Georgia 30602, United States of America.
[c] Journalism and Media Studies Centre, The University of Hong Kong, Hong Kong Special Administrative Region, China.
Submitted: 3 March 2015; Published: 26 June 2015
doi: 10.5365/wpsar.2015.6.1.019

Table 1. **Summary of the uses of social media in public health surveillance***

Application	Purpose and scenario	Public health information to retrieve, detect or predict	Data targeted	Function of social media	Examples
1. Epidemiologic monitoring and surveillance					
(a) Monitoring official information	To monitor official information	Disease incidence and other case details	Links to original sources of official data	News feed	Retrieval of official information via Weibo during the 2013 H7N9 outbreak[2,3]
(b) Disease detection – syndromic surveillance	To detect outbreaks and to estimate disease incidence	Disease incidence and other case details	Self-reported symptoms	Syndromic surveillance	Twitter tweets of self-disclosed symptoms of influenza infection;[4] self-reports of symptoms via specialized apps[5,6]
(c) Disease detection – event-based surveillance	To detect outbreaks and to estimate disease incidence	Disease incidence and other case details	Media reports, unofficial information or rumours used as proxy measurable outcomes	Event-based surveillance (epidemic intelligence)	Unofficial information released on Weibo about an H7N9 patient;[7] systems that pick up news related to health events (e.g. HealthMap)[8]
(d) Timely estimates and forecasting of disease incidence	To provide timely estimates of current disease incidence or forecast future disease incidence	Disease incidence: start, peak and intensity	Social media text with keywords (diseases or symptoms) that correlate with disease incidence	As data sources for timely estimates or forecasts of disease incidence	US seasonal influenza estimates using Twitter data;[4] disease estimation and forecasts using Wikipedia access log;[9] US seasonal influenza forecast using Google Flu Trends[11]
2. Situational awareness					
(a) Surveillance for situational awareness	Humanitarian crises, usually natural disasters, e.g. typhoons and earthquakes	Reported needs (e.g. water supply and shelter)	Self-reported humanitarian needs	Information feed on humanitarian needs	Earthquake and tsunami in East Japan[12] and earthquake in Haiti[13]
3. Communication surveillance					
(a) Global awareness	To measure social media users' reactions to an outbreak situation	Media news reports, rumours, sentiments, awareness	User-generated data that reflect their knowledge, attitudes and perception of public health events	Monitoring of the general public's awareness and perception	Awareness of Ebola;[14] sentiment towards influenza vaccine[15]
(b) Specific reactions	To measure social media users' reactions to health promotion messages or events	Reception of particular public health messages	User generated data that are reactions to particular public health messages	Monitoring of the general public's reaction to specific public health messages	Breast cancer awareness month[16]

Surveillances des maladies
Détection d'épidémies
Surveillance syndromique
Incidence de maladies
Crises sanitaires
Réactions à des campagnes de santé publique
Pharmacovigilance
Usage des médicaments

METHODOLOGIE

DETEC't

SOURCES

FORUMS

Doctissimo
vivre bien au quotidien

Sujet : après 1 semaine avec Glucor

Posté le 03/03/2018 à 14:15:14

Voilà presque une semaine que je prend Glucor 50 (qui m a été prescrit pr éviter les hypo fonctionnel)
Autant dire que je revis ... Plus d hypo du tout ! Le bonheur ..
Par contre comme l interne de l hôpital m avait dit , effectivement niveau intestinale c est pas le joie ... Mal au ventre , ballonnement et direct aux toilettes après chaque repas .. mais à choisir ... Je préfère ça que les hypo à répétition

J'aime 0

Sujet : Perte de poid avec le glucophage

Posté le 06/01/2010 à 11:42:33

Bonjour a tous !

es ce que qq1 utilise glucophage 500mg ???
Moi sa fait 10 jours que je le prend et g perdu 7 kilo en tous es ce que cela est normal ??? je m'inkiete !!!!

Message cité 1 fois

---

traitement de glucophage depuis le 18 decembre 2009 , ompk .
Rendez vous hez le gygy le 18 mars !!

date de notre mariage le 18 Juin 2011

J'aime 0

FACEBOOK

Dave Gill
January 20, 2015 · Davis, CA, United States

Target and CVS are out of codeine cough medicine. Supplier shortage...I can barely speak my throat is so raw.....

Robert Torres, Nan Becker and 4 others — 20 Comments

Like — Comment

About This Group

Description

This page is for anyone struggling with Codeine Addiction. If you have found this page it means you have been searching for a solution. If you would like to join after reading this please note that it is compulsory to answer... See More

Group Type

Support

Members · 398

Admins

Kristy, Genevieve and 3 other members are admins.

See All Members

TWITTER

GILU DATTEBAYO @ColonelKickassz · 21 mars
G 6 boites de **codeine** jv bibi a des rappeurs new wave pr faire fortune
1 · 1

STRVB2o2s @VicoStrab · 20 mars
fini la **codéine**, toujours sirop rouge mais c'est plus d'la lean mais d'la grenadine
1 · 2

サミュエル @_samuel_bt · 23 mars
la philo sous **codéine**: ptdr
1 · 1

Ms Duckworth @creopatraa · 22 mars
Je peux pas avoir pris un doliprannne **codéiné** et avoir toujours cette putain de migraine !
2 · 2

Afficher cette discussion

kamae amayi @anemaaryy · 21 mars
ptdddr le 3e couplet de kodak dans **codeine** dreaming il mfume ça montre à quele point il était fly
1 · 2

Afficher cette discussion

YOUTUBE

Tramadol is NOT a 'safe alternative' to narcotics.
43 253 vues

susanfrits7 Il y a 2 ans
I've been given tramadol for ms related pain because I am unable to take any other pain meds. It was not first resort. Ive been on the same dosage for 5 years. It's hard to function without it but I can, I just have a better quality with it. I have no side effects. My concern here is that medical professionals get on these kicks to subscribe medicines (ie: adhd meds, codeine , neurontin, lyrica antibiotics this list goes on) then barn it falls out of grace and suddenly no one can be prescribed the medication. I've watched extremely ill patients where I work suffer in agonising pain because it became unpopular to prescribe pain killers after crack down on abuse. These are people who need the meds but now can't get them. Is it to much to ask of the medical community to prescribe on a per patient basis and not on what's popular. What happens if tramadol falls from grace? Can I no longer use it ?
Moins
19 RÉPONDRE
Afficher les 2 réponses

INSTAGRAM

Commentaires

kimcankickit Keytruda infusion #28.
This drug was the beginning of everything. I had 40+ tumors covering my lungs, and two in my liver. After only four infusions one tumor was left in my lungs plus the two that remained in my liver. Because of this I became eligible for surgery. If Dr. Z would have seen my case before Keytruda he wouldn't have touched me. I asked Dr. Stinnett if he was thinking about taking me off of it now that my scans are clean and he replied "I know you better then that. You would get all mad and go find another doctor to give it to you." You better believe I would and I love that he knows that. I don't want to find out what might happen if we take me off of Keytruda. As far as I'm concerned I plan to remain on it for life.
3sem

ALGORITHME D'ANALYSE DES PARCOURS DE SOINS

Identifier les messages relatant un parcours de soins avec les acteurs de santé consultés

ALGORITHME D'ARRÊT DE TRAITEMENT

Identifier les messages relatant un arrêt de traitement et en identifier les motifs. Analyse de l'observance

ALGORITHME DE PRESENCE

Identifier les messages contenant une liste de mots clés spécifiés dans une équation de recherche

ALGORITHME DE PRISE

Identifier les messages relatant une prise de médicaments par la personnes ayant posté un message relatif à ce médicament

ALGORITHME D'EVALUATION DE LA SEVERITE D'UNE PATHOLOGIE

Identifier les messages relatant un arrêt de traitement et en identifier le motif

suis pré diabétique depuis longtemps, soignée p
mine matin et soir. Depuis deux mois environ, je suis rév
5h du matin par une sensation désagréable d'hypoglycémie.
rtiges, grosse faim. C'est tout nouveau pour moi. Je n'ai rien chang
ans les habitudes, je marche 13000 pas par jours depuis très
gtemps. Ce matin, sensation d'hypoglycémie et très grosse soif. J
ma diabeto fin avril. Es ce qu'il faut que je mange plus le soir
ce à être très fatiguée de me réveiller comme ça. J'al
Internet mais c'est pour les personnes qui

ALGORITHME DE DETECTION DES CONCEPTS MEDICAUX

MedDRA

Identifier dans les messages les concepts médicaux exprimés par les patients et les coder dans la terminologie MedDRA

ALGORITHME DE DETECTION DES GROSSESSES

Identifier les messages postés par des femmes enceintes ou relatant un épisode relatif à leurs grossesses

ALGORITHME DE DETECTION DES EFFETS INDESIRABLES

Identifier les messages relatant un effet indésirable

ALGORITHME DE TERATOGENECITE

Identifier les messages relatant une complication ou des malformations fœtales

ALGORITHMES D'IMPACT SUR LA QUALITE DE VIE

Identifier les messages relatant un impact d'une pathologie sur des dimensions de qualité de vie, sur le travail, la vie sociale…

ALGORITHME D'IDENTIFICATION DU GENRE

Identifier automatiquement le genre de la personne ayant posté

RESEAUX SOCIAUX ET PHARMACOVIGILANCE

KAP CODE
CONNECT YOUR HEALTH

PROFIL DE SECURITE

EI potentiel	Fréquence	Attendu
Effet indésirable	424	Non
Poids	384	Non
Fatigue	371	Non
Poids fluctuant	245	Non
Poids augmenté	151	Non
Tremblement	105	Oui
Malformation	92	Non
Somnolence	89	Non
Dépendance	88	Non
Anxiété	88	Non

Produits associés	Nombre d'apparition
lamictal	1715
tegretol	646
epitomax	520
urbanyl	361
rivotril	194

signal	PRR
Encéphalopathie	1017
Anomalie du comportement	872
Défauts visuels	654
Atteinte de la mémoire	523
Rétention liquidienne	509
Trouble du langage	509
Effet indésirable	446
Idées suicidaires	363
Symptôme dépressif	363
Chute de cheveux	327

LEVOTHYROX

Doctissimo

29 015 messages
8 398 internautes

* REPARTITION PAR AGE ET PAR SEXE *

Répartition géographique des messages

Evolution temporelle du nombre de messages

Observations mensuelles – 2000-2017

SURVEILLANCE DE L'ACTIVITÉ

** L'algorithme de veille modélise l'évolution du nombre de messages et classe les observations selon les volumes historiques observés pour le produit considéré.*

THÈMES DES DISCUSSSIONS 2017

Le changement de formule au cœur de discussions en 2017

Proportions de messages postés en 2017 associés à chaque thème*

** Un message peut être associé à plusieurs thèmes*

Exemples de messages

Actualités du Levothyrox

Anny Duperey vient d'écrire une lettre ouverte à la ministre de la santé. J'espère que cela va faire bouger les choses. Elle prend elle aussi du Lévothyrox depuis une douzaine d'années et depuis qu'elle prend le nouveau elle a plein d'effets secondaires qu'elle a décrits dans sa lettre : ce sont les mêmes que les miens.

Commandes à l'étranger

Bonsoir, j'ai **pu acheter du Lévothyrox en Espagne**. Pas besoin de la carte vitale. Par contre, la pharmacie a conserve mon ordonnance. Il s'agit de boîte de 100 comprimes (donc 3 mois de traitement) environ 4.50€ la boîte."

Effets secondaires

Bonjour, je prends le Lévothyrox depuis 45 ans et j'étais au dosage maximum, tout allait bien jusqu'a ce qu'ils changent la posologie [...]. **Je suis en souffrance depuis le nouveau médicament**.

Débats sur le changement de formule

Arreter le Lévothyrox, va provoquer des symptômes mille fois plus important que les symptômes que nous avons actuellement avec la nouvelle formule. Arrêter le Lévothyrox, ca veut dire vivre sans thyroïde. En 6 mois il y a des comas et ensuite la mort. Ne pas arrêter le Lévothyrox a moins que le médecin vous donne soit en France ou en Europe

EVOLUTION TEMPORELLES DES DÉCLARATIONS D'EFFETS INDÉSIRABLES

DÉTECTION DE SIGNAUX – NOUVEAUX SIGNAUX 2017

RESEAUX SOCIAUX ET MÉSUSAGE DES MÉDICAMENTS

KAP CODE
CONNECT YOUR HEALTH

USAGE DE LA CODEINE

Corpus d'analyse : 12 656 messages

PHARMACODÉPENDANCE – AVANT/APRÈS L'ARRÊTÉ DU 12 JUILLET 2017

Filtre : Messages faisant référence à de la dépendance à la codéine

Corpus : (avant/après arrêté) : 110/ 1 421 messages

DÉTECTION DE SIGNAUX – CODÉINE

Top 20 des signaux significatifs : PRR ≥ 2, Chi² ≥ 4, N ≥ 3

Fréquence (N)

RESEAUX SOCIAUX ET MALADIES RARES / QOL / PARCOURS DE SOINS

KAP • CODE
CONNECT YOUR HEALTH

MALADIES RARES / IMPACT SUR LA QUALITE DE VIE

Figure 11 – Description of Quality of life impacts. Orange bars correspond to messages with treatments, yellow bars correspond to messages without treatments.

RESSENTI PATIENT / PROFESSIONNELS DE SANTE / MALADIE RARES

PARCOURS DE SOINS

Parcours de soins (uniquement ceux qui ont consulté au moins une deuxieme fois)

VACCINS

SYNTHÈSE DES RÉSEAUX SOCIAUX
AOÛT 2018

LES CHIFFRES-CLÉS

2691 messages analysés
4 pays : France, Belgique, Suisse, Canada
30 sources surveillées*
8 relations identifiées

ÉVOLUTION DU NOMBRE DE MESSAGES

TOP 10 DES CONCEPTS MÉDICAUX

1. DOULEUR
2. GRIPPE
3. GROSSESSE
4. DÉPENDANCE
5. TUMEUR MALIGNE
6. FATIGUE
7. INFECTION
8. HÉPATITE
9. MORT
10. FIÈVRE

COMMUNAUTÉ DE PATIENTS

CARACTÉRISTIQUES

1413 Internautes
1,9 Posts/auteur
20 Auteurs actifs

THÈMES DES TOPS USERS

1 – OBLIGATION DE VACCINATION
2 – VACCINATION DE BÉBÉ
3 – MALADIE AUTO-IMMUNE

THÈMES DES DISCUSSIONS

TOP 5 DES VACCINS CITES

1. GARDASIL
2. PREVENAR
3. POLIO
4. PNEUMO
5. MENINGITEC

DETEC'T PROPOSE

L'étude de l'observance de la vaccination
Identification des leviers d'observance
Identification des freins à la prise du produit

L'étude des effets indésirables rapportés par les internautes

*Sources surveillées : Atoute, Doctissimo, Au féminin, E-santé, Santé médecine, Bebe patient, Patient Co UK, Steady Health, Planete santé, Creapharma, Futura Sciences…

CONCLUSION (PROVISOIRE)

- Les réseaux sociaux représentent une nouvelle source de données de vie réelle
- Les technologies sont aujourd'hui matures pour des analyses scientifiques
- Positionnement de l'infodémiologie avec les autres types d'études à définir
- Importance attendue des études multi-sources

MERCI

Dr Warehouse

imagine
INSTITUT DES MALADIES GÉNÉTIQUES

Necker
ENFANTS MALADES
HÔPITAL UNIVERSITAIRE

Inserm
Institut national
de la santé et de la recherche médicale

Dr Warehouse, un accélérateur pour la recherche translationnelle hospitalière ?

Nicolas Garcelon

nicolas.garcelon@institutimagine.org

Shire

DONNEES MEDICALES DU PATIENT

=

GISEMENT DE CONNAISSANCE

- ✓ **La recherche clinique :** inclusion dans des essais, clustering de patients
- ✓ **La recherche Epidémiologique :** constitution de cohorte
- ✓ **Le pilotage médico-économique** : analyse de l'activité, contrôle qualité
- ✓ **L'évaluation des pratiques** : Etude sur la qualité de la prise en charge des patients
- ✓ **Les vigilances :** détection des EI médicamenteux, détections des infections nosocomiales
- ✓ Etc.

REUTILISATION DES DONNEES MEDICALES DU PATIENT, EN PRATIQUE ?

Quels sont les patients éligibles pour ce projet de recherche ?

Revenez dans 1 mois …

Mais c’était avant…

Une informatisation des données mais

Des données dispersées et cloisonnées au sein de l'hôpital

Hétérogénéité :

- De format
- De transmission (flux, bdd)
- De volume
- De technologie

Flux HPRIM ou HL7

```
H|~^\&|EFS09042100006.HPR||000000031026~EFS PM||ORU|||~Purpan|||H2.1~L|200904210015
P|1|0448964160||0448964160|HEMIE~ANNE~~~~|SOEUR|19000101|F||FOUSSAN~~SUR
    TARN~~81999~||||||||||||||~~2416~||||||||
OBR|1|10256852~10256852|10256852~10256852|ABO-RH1~GROUPE ABO-
    D~L~~~|S|200904202355|200904202045|200904202045|1||R|||200904202355|S&&L~~|2416~H2
    416 Urgence gynéco~L~33111~C.H.U  PURPAN~L|||||200904210013||TLPIH~Immuno-|I
A||||||||~||||
OBX|1|TX|ABO-RH1~Groupe~L~~~||||||||||||HR~303~ABO-RH1~0030~1~|~
OBX|2|TX|RHD~Antigène D (RH1)~L~~~||||||||||||HR~303~ABO-RH1~0030~2~|~
OBR|2|10256852~10256852|10256852~10256852|PHEN-RHK~PHENOTYPE RH-
    KELL~L~~~|S|200904202355|200904202045|200904202045|1||R|||200904202355|S&&L~~|2416
    ~H2416 Urgence gynéco~L~33111~C.H.U  PURPAN~L|||||200904210013||TLPIH~Imm
A|uno-||||||||~||||
OBX|1|TX|RH2~Antigène C (RH2)~L~~~||||||||||||HR~303~PHEN-RHK~0046~1~|~
OBX|2|TX|RH3~Antigène E (RH3)~L~~~||||||||||||HR~303~PHEN-RHK~0046~2~|~
OBX|3|TX|RH4~Antigène c (RH4)~L~~~||||||||||||HR~303~PHEN-RHK~0046~3~|~
OBX|4|TX|RH5~Antigène e (RH5)~L~~~||||||||||||HR~303~PHEN-RHK~0046~4~|~
OBX|5|TX|KEL1~Antigène Kell (KEL1)~L~~~||||||||||||HR~303~PHEN-RHK~0046~5~|~
L|1||1|14|
```

Les bases de données « recherche » à l'hôpital :

Souvent pas d'hypothèse de départ
Réelle raison :
Pouvoir se réapproprier les données pour rechercher dedans…
Et éventuellement trouver de nouvelles questions scientifiques.

Problèmes :
Bases sont souvent non exhaustives, voire vides : trop long à remplir
Il va manquer LE champ qui était finalement hyper important (2 ans après le début de l'étude)
On ne va coder que ce que l'on connaît

Entrepôt de données biomédicales

W.H. Inmon, Building the Data Warehouse, John Wiley & Sons, Inc., New York, NY, USA, 1992.

Entrepôts de données existants

i2b2	(Murphy, 2006)	Données codées (Thesaurus)
Stride	(Lowe, 2009)	
Vanderbilt	(Danciu, 2014)	
Starmaker	(Krasowski, 2015)	
Meteor	(Puppala, 2015)	
OHDSI	(Hripcsak, 2015)	
Emerse	(Hanauer, 2006)	Données texte libre

Les données structurées suffisent-elles ?

Dans une cohorte de 741 patients atteints d'une maladie coeliaque, les comorbidités associées sont présentes dans

Données codées

Codées et textuelles

Données textuelles

81 %

Les comorbidités sont présentes uniquement dans le texte libre

Escudié J-B et al. A novel data-driven workflow combining literature and electronic health records to estimate comorbidities burden for a specific disease: a case study on autoimmune comorbidities in patients with celiac disease. BMC Med Inform Decis Mak. 2017;17:140.

Les données textuelles sont indispensables

Description phénotypique plus riche [1,2]

Description des patients non diagnostiqués

Expression du doute et de l’absence de signes

Dans le contexte des maladies rares : évolution des connaissances

Faites de la littérature !!!

*[1] Hanauer DA. EMERSE: The Electronic Medical Record Search Engine. AMIA Annu Symp Proc 2006;**2006**:941.*

*[2] Raghavan P et al. How essential are unstructured clinical narratives and information fusion to clinical trial recruitment? AMIA Jt Summits Transl Sci Proc 2014;**2014**:218–23.*

Dr Warehouse

Un entrepôt de données orienté autour du document hospitalier :

- Pour les médecins
- Pour tous les types de données

3 défis identifiés

*Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. Methods Inf Med 2009;**48**:38–44.*

Dr Warehouse à Necker

640 307 patients
5 000 145 documents

26 sources de données
De 1996 à 2018

240

9000

53000

470

Principaux cas d'usage :

Rechercher des patients pour :

① Réaliser des études de faisabilité

② Recruter des patients dans des études cliniques

③ Tester des associations de signes, de médicaments, etc.

④ Retrouver des patients similaires à un cas complexe (RCP) :

- Pour adapter le traitement / la prise en charge
- Trouver un diagnostic

Les bases de données « recherche » à l'hôpital :

Réutiliser les données produites au cours du soin pour trouver de nouvelles hypothèses

Créer la base de données pour répondre à ces hypothèses :

- Efficacité
- Gain de temps

Fouille de données sur les données de l'hôpital

Questions scientifiques

Bases de données « propres »

Recherche en texte libre

Recherche sur les données codées : résultats de laboratoire, code PMSI etc…

Ajout de contraintes temporelles entre les critères de recherche

Critères démographiques : Sexe, Age, décès

Garcelon N, et al. *Improving a full-text search engine: the importance of negation detection and family history context to identify cases in a biomedical data warehouse* . Journal of the American Medical Informatics Association 2016; doi: 10.1093/jamia/ocw144

Dr WareHouse ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Outils | Contact | Patient Nom / IPP | Admin | Nicolas Garcelon Déconnexion

Rechercher des patients

Sur tout l'entrepôt

epidermolyse bulleuse 341/341

Etendre aux synonymes :

+ Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Filtre patient

LANCER LA RECHERCHE ?

Requêtes sauvées

panniculite and calcification

RETT et MECP2

dyscalculie

heterotopie

test merçjl retti

heterotopie NEUROLOGIE PED

Historique requête

Show 10 entries Search:

date	Requêtes
29/10/2015 22:20	Documents contenant hypertrichose non hypertrichose 0 , en excluant les négations
15/10/2015 18:56	Documents contenant cardiopathie a lga et purpura rhumatoïdes and biopsie rénale , en excluant les négations
26/08/2015 14:52	Documents contenant Persistance du canal artériel or hypoplasie de l' arche aortique et Coarctation aortique or sténose de l' arche

Résultat | Cohortes | Données stat | Concepts | PMSI | Biologie | Genes | Map | Clustering | Datamart

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse' , en excluant les négations SAUVER LA REQUÊTE

Refaire une recherche sur le résultat :

Exporter les patients :

Filtrer le résultat ci-dessous : ok annuler

Interface adaptée à la recherche en texte libre

[redacted]
11 ans

Echographie abdominale du [redacted] - RADIOLOGIE PEDIATRIQUE :
INDICATION : Sd inflammatoire bilogique avec fievre, douleurs abdominales et vomissements chez un enfant presentant une epidermolyse bulleuse [...]

[redacted]
10 ans

CRH HdJ du [redacted]
Madame **[PRENOM_PATIENT]** , nee le **[DATE_NAIS_PATIENT]**, a ete hospitalisee en hopital de jour / seance dans le service de [redacted] pour cure de Venofer dans le contexte d'epidermolyse bulleuse dystrophique recessive avec absence de collagene 7 [...] -Epidermolyse bulleuse dystrophique recessive avec absence totale de collagene 7 [...]

Ostéodensitométrie du corps entier- Enfant du [redacted] (PACS) par [redacted]
Premier Examen dans le cadre d'une epidermolyse bulleuse [...]

630 CRH Dermatologie du [redacted] - DERMATOLOGIE :
Motif d'hospitalisation : reevaluation d'une epidermolyse bulleuse dystrophique recessive avec absence totale de collagene VII chez une enfant de [redacted] ans [...] Epidermolyse bulleuse dystrophique recessive confirmee sur le plan moleculaire (mutation homozygote PARG613X sur l'exon 14 du collagene 7) [...] Epidermolyse bulleuse dystrophique recessive chez une jeune fille de [redacted] ans [...]

[redacted]

+ afficher les autres documents

[redacted]
17 ans

302 COURRIER du [redacted] METABOLISME-NEUROLOGIE PEDIATRIQUE :
J'ai recupere le bilan endocrinien effectue a **[PRENOM_PATIENT]** que tu suis pour une epidermolyse bulleuse dystrophique severe depuis sa naissance [...]

CRH de pédiatrie du [redacted]
Dilatation oesophagienne chez un patient atteint d'une epidermolyse bulleuse dystrophique recessive [redacted] [...] Epidermolyse bulleuse dystrophique recessive [...] [redacted] epidermolyse bulleuse dystrophique recessive par absence totale de collagene VII [...]

DIAG PMSI du [redacted] :
Code principal : Q812 epidermolyse bulleuse dystrophique ; Codes significatifs : D509 anemie par carence en fer, sans precision ; R53+2 Fatigue [asthenie] ; K9380 Atteintes de la cavite buccale au cours de

Dr WareHouse ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Outils | Contact | Patient Nom / IPP | Admin | Nicolas Garcelon Déconnexion

Rechercher des patients

Sur tout l'entrepôt

epidermolyse bulleuse 341/341

Etendre aux synonymes :

+ Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Filtre patient

LANCER LA RECHERCHE ?

Requêtes sauvées

[illegible]

Historique requête

Show 10 entries Search:

date	Requêtes
29/10/2015 22:20	[illegible]
15/10/2015 18:56	[illegible]
26/08/2015 14:52	[illegible]

Résultat | Cohortes | Données stat | Concepts | PMSI | Biologie | Genes | Map | Clustering | Datamart

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse' , en excluant les négations SAUVER LA REQUÊTE

Refaire une recherche sur le résultat :

Exporter les patients :

Filtrer le résultat ci-dessous : ok annuler

[illegible]
11 ans

Echographie abdominale du [illegible] - RADIOLOGIE PEDIATRIQUE :
INDICATION : Sd inflammatoire bilogique avec fievre, douleurs abdominales et vomissements chez un enfant presentant une epidermolyse bulleuse [...]

[illegible]
10 ans

CRH HdJ du [illegible]
Madame [PRENOM_PATIENT [illegible] pour
cure de Venofer dans le con [illegible] phique recessive avec absence totale de collagene 7 [...]

Ostéodensitométrie du co
Premier Examen dans le cac

630 CRH Dermatologie du [illegible] ans [...] Epidermolyse bulleuse dystrophique recessive
Motif d'hospitalisation : ree [illegible] chez une jeune fille de [illegible] ans [...]
confirmee sur le plan molec

[illegible]

+ afficher les autres document

MAHAMED HASSANALY Asma 28/11/2005 F 10 ans

CRH HdJ - 07/03/2016

1/3
CRH HDJ NCK DERMATOLOGIE PED, [illegible]
MOTIF D'HOSPITALISATION
PARIS CEDEX 15, [illegible]
References : CPH/CPH
Compte-rendu d'hospitalisation de jour du [illegible] a ete hospitalisee en hopital de jour /
seance dans le service de NCK HJ3 DERMATOLOGIE - HJ CLINIQUE ROBERT DEBRE RDC le [illegible] pour cure de Venofer dans le contexte d'**epidermolyse bulleuse** dystrophique recessive
avec absence de collagene 7..
HISTOIRE DE LA MALADIE
TRAITEMENT a L'ENTReE
Antecedents medicaux :
Antecedents chirurgicaux :
-**Epidermolyse bulleuse** dystrophique recessive avec absence totale de collagene 7.

[illegible]
17 ans

302 COURRIER du [illegible]
J'ai recupere le bilan endoc [illegible] nce [...]

CRH de pédiatrie du [illegible]
Dilatation oesophagienne ch [illegible] phique recessive [...]
epidermolyse bulleuse dyst

DIAG PMSI du [illegible] :
Code principal : Q812 epidermolyse bulleuse dystrophique ; Codes significatifs : D509 anemie par carence en fer, sans precision ; R53+2 Fatigue [asthenie] ; K9380 Atteintes de la cavite buccale au cours de

Détecter l’expression de la négation et des antécédents familiaux

Objectif : **conserver le texte libre**

Notre approche : **découper le texte en syntagmes et de les classer en fonction du contexte et de la certitude**

Compte rendu

classification

Texte	Certitude	Contexte
	Affirmation	Patient
	Négation	Patient
	Affirmation	Famille
	Négation	Famille

Dr. WareHouse ©Imagine

Entrepôt de données

Home | Search engine | My Requests | My Cohorts | Tools | Patient Nom / IPP | Notificati

Search for patients

Throughout the warehouse

lupus 1041/1041

Extending to synonyms :

+Advanced - Rewrite the query

+ Add a filter Full text

+ Add a structured filter

+ Time constraints

+ Patient filter

START A RESEACH ?

Saved queries

Search:

title
No data available in table

History query

Search:

Dated	queries
28/07/2017 09:07	Filtered 1 : Documents containing 'lupus' , Excluding negations
27/07/2017 16:31	Filtered 1 : Documents containing 'infection% and eczema and thrombopenie' , Excluding negations
27/07/2017 16:28	Filtered 1 : Documents containing 'MECP2' , Excluding negations
27/07/2017 15:06	Filtered 1 : Documents containing 'lupus and cicatrice' , Excluding negations

Result | Cohort : Test 0 x | Stats data | Concept | Biology | Map | clustering | Datamart

1041 patients
5972 Documents

Filtered 1 : Documents containing 'lupus' , Excluding negations

SAVE QUERY | SEARCH ON RESULT | EXPORT PATIENTS TO EXCEL | FIT THE RESULT BELOW | SHARE YOUR QUERY | FEEDING A COHORT

+ Create a cohort

Or Select an existing cohort to feed it to patients :

Choose a Cohort

Open Cohort : Test x

Include all of the following patients in the cohort
Exclude all patients below from the cohort
In doubt all patients below
Import all the patients below into the cohort
Do not display patients already included or excluded

Patient F, 42 years

Lettre médicale du [redacted] - NEPHROLOGIE ADULTE :
Pour ce qui est du lupus, les articulations font mal comme depuis toujours, pas de lesion cutanee en rapport [...]
280 HDJ NEPHROLOGIE ADULTE du [redacted], par [redacted] - NEPHROLOGIE ADULTE :
LUPUS ERYTHEMATEUX DISSEMINE [...] Lesions cutanees brulees a l'azote liquide, decrites par la patiente comme « cicatrices de lupus » [...] 2/ sur le plan du lupus : [...] Absence d'argument pour u
lupus, bilan immunologique en attente [...] [redacted] qui avait conclu a l'absence d'arguments pour un neuro-lupus [...]

+ View other documents

Patient F, 14 years

090 CION RHUMATOLOGIE du [redacted], par [redacted] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :
L'histoire familiale revele un LUPUS avec atteinte renale chez la maman [...] Lupus aenticoagulant et anticardiolipine [...]

Patient F, 37 years

373 CRH HEMATOLOGIE ADULTE du [redacted], par [redacted] - HEMATOLOGIE CLINIQUE ADULTE :
- Lupus cutane et hematologique ([redacted]), puis glomerulopathie lupique stade V ([redacted]) [...]
HEMATOLOGIE ADULTES SCID du [redacted], par [redacted] - HEMATOLOGIE CLINIQUE ADULTE :
HDM [redacted]: lupus ([...]
DIAG PMSI du [redacted] par [redacted] - HEMATOLOGIE CLINIQUE ADULTE :

Faciliter la sélection de patients pour des études

Dr. WareHouse ©Imagine

Entrepôt de données

Home | Search engine | My Requests | My Cohorts | Tools | Patient Nom / IPP | Notification

Search for patients

Throughout the warehouse

lupus 1041/1041

Extending to synonyms :

+Advanced - Rewrite the query

+ Add a filter Full text

+ Add a structured filter

+ Time constraints

+ Patient filter

START A RESEACH ?

Saved queries

Search:

title
No data available in table

History query

Search:

Dated	queries
28/07/2017 09:07	Filtered 1 : Documents containing 'lupus' , Excluding negations
27/07/2017 16:31	Filtered 1 : Documents containing 'infection% and eczema and thrombopenie' , Excluding negations
27/07/2017 16:28	Filtered 1 : Documents containing 'MECP2' , Excluding negations
27/07/2017 15:06	Filtered 1 : Documents containing 'lupus and cicatrice' , Excluding negations

Result | Cohort : Test 1 x | Stats data | Concept | Biology | Map | clustering | Datamart

1041 patients
5972 Documents

Filtered 1 : Documents containing 'lupus' , Excluding negations

SAVE QUERY | SEARCH ON RESULT | EXPORT PATIENTS TO EXCEL | FIT THE RESULT BELOW | SHARE YOUR QUERY | FEEDING A COHORT

+ Create a cohort

Or Select an existing cohort to feed it to patients :

Choose a Cohort

Open Cohort : Test x

Include all of the following patients in the cohort
Exclude all patients below from the cohort
In doubt all patients below
Import all the patients below into the cohort
Do not display patients already included or excluded

Patient F, 42 years

Lettre médicale du [redacted] - NEPHROLOGIE ADULTE :
Pour ce qui est du lupus, les articulations font mal comme depuis toujours, pas de lesion cutanee en rapport [...]
280 HDJ NEPHROLOGIE ADULTE du [redacted], par [redacted] - NEPHROLOGIE ADULTE :
LUPUS ERYTHEMATEUX DISSEMINE [...] Lesions cutanees brulees a l'azote liquide, decrites par la patiente comme « cicatrices de lupus » [...] 2/ sur le plan du lupus : [...] Absence d'argument pour une lupus, bilan immunologique en attente [...] [redacted] qui avait conclu a l'absence d'arguments pour un neuro-lupus [...]
+ View other documents

Patient F, 14 years

090 CION RHUMATOLOGIE du [redacted] par [redacted] - IMMUNO-HEMATOLOGIE PEDIATRIQUE :
L'histoire familiale revele un LUPUS avec atteinte renale chez la maman [...] Lupus aenticoagulant et anticardiolipine [...]

Patient F, 37 years

373 CRH HEMATOLOGIE ADULTE du [redacted] par [redacted] - HEMATOLOGIE CLINIQUE ADULTE :
- Lupus cutane et hematologique [redacted], puis glomerulopathie lupique stade V [redacted] [...]
HEMATOLOGIE ADULTES SCID du 16/09/14 (KDOS) par SUAREZ Felipe - HEMATOLOGIE CLINIQUE ADULTE :
HDM [redacted] lupus ([...]
DIAG PMSI du [redacted] 'PMSI DIAG) par [redacted] - HEMATOLOGIE CLINIQUE ADULTE :

Faciliter la sélection de patients pour des études

Principaux cas d'usage :

Dr WareHouse ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Outils | Contact | Patient Nom / IPP | Admin | Nicolas Garcelon Déconnexion

Rechercher des patients

Sur tout l'entrepôt

epidermolyse bulleuse 341/341

Etendre aux synonymes :

+ Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Filtre patient

LANCER LA RECHERCHE ?

Requêtes sauvées

[illegible]

Historique requête

Show 10 entries Search:

date	Requêtes
29/10/2015 22:20	[illegible]
15/10/2015 18:56	[illegible]
26/08/2015 14:52	[illegible]

Résultat | Cohortes | Données stat | Concepts | PMSI | Biologie | Genes | Map | Clustering | Datamart

341 Patients
2396 Documents

Documents contenant 'epidermolyse bulleuse' , en excluant les négations SAUVER LA REQUÊTE

Refaire une recherche sur le résultat :

Exporter les patients :

Filtrer le résultat ci-dessous : ok annuler

[illegible]
11 ans

Echographie abdominale du [illegible] - RADIOLOGIE PEDIATRIQUE :
INDICATION : Sd inflammatoire bilogique avec fievre, douleurs abdominales et vomissements chez un enfant presentant une epidermolyse bulleuse [...]

[illegible]
10 ans

CRH HdJ du [illegible]
Madame [PRENOM_PATIENT] , nee le [DATE_NAIS_PATIENT], a ete hospitalisee en hopital de jour / seance dans le service de [illegible] pour cure de Venofer dans le contexte d'epidermolyse bulleuse dystrophique recessive avec absence de collagene 7 [...] -Epidermolyse bulleuse dystrophique recessive avec absence totale de collagene 7 [...]

Ostéodensitométrie du corps entier- Enfant du [illegible] (PACS) par [illegible]
Premier Examen dans le cadre d'une epidermolyse bulleuse [...]

630 CRH Dermatologie du [illegible] - DERMATOLOGIE :
Motif d'hospitalisation : reevaluation d'une epidermolyse bulleuse dystrophique recessive avec absence totale de collagene VII chez une enfant de [illegible] ans [...] Epidermolyse bulleuse dystrophique recessive confirmee sur le plan moleculaire (mutation homozygote PARG613X sur l'exon 14 du collagene 7) [...] Epidermolyse bulleuse dystrophique recessive chez une jeune fille de [illegible] ans [...]

[illegible]

+ afficher les autres documents

[illegible]
17 ans

302 COURRIER du [illegible] METABOLISME-NEUROLOGIE PEDIATRIQUE :
J'ai recupere le bilan endocrinien effectue a [PRENOM_PATIENT] que tu suis pour une epidermolyse bulleuse dystrophique severe depuis sa naissance [...]

CRH de pédiatrie du [illegible]
Dilatation oesophagienne chez un patient atteint d'une epidermolyse bulleuse dystrophique recessive [...] Epidermolyse bulleuse dystrophique recessive [...] epidermolyse bulleuse dystrophique recessive par absence totale de collagene VII [...]

DIAG PMSI du [illegible] :
Code principal : Q812 epidermolyse bulleuse dystrophique ; Codes significatifs : D509 anemie par carence en fer, sans precision ; R53+2 Fatigue [asthenie] ; K9380 Atteintes de la cavite buccale au cours de

Dr WareHouse ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Outils | Contact | Patient Nom / IPP | Admin | Nicolas Garcelon Déconnexion

Rechercher des patients

Sur tout l'entrepôt

epidermolyse bulleuse 341/341

Etendre aux synonymes : ☐

+ Avancé - Réécrire la requête

+ Ajouter un filtre Full text

+ Ajouter un filtre structuré

+ Filtre patient

LANCER LA RECHERCHE ?

Requêtes sauvées

[redacted]

Historique requête

Show 10 entries Search:

date	Requêtes
29/10/2015 22:20	[redacted]
15/10/2015 18:56	[redacted]
26/08/2015 14:52	[redacted]

Résultat | Cohortes | Données stat | Concepts | PMSI | Biologie | Genes | Map | Clustering | Datamart

Sur tout l'hôpital :
341 Patients
2396 Documents

Afficher le parcours moyen (hospitalisations) des patients
Afficher le parcours moyen complet (hospitalisations et consultations) des patients

Pyramide des âges au 1er document trouvé

Pyramide des âges aujourd'hui

Pyramide des âges < 20 ans au 1er document trouvé

Pyramide des âges < 20 ans aujourd'hui

Analyse du parcours de soin d'une population

Des outils de visualisation des données agrégées

Dr. WareHouse ©Imagine

Entrepôt de données

Home | Search engine | My queries | My Cohorts | Tools | Patient Patient name or ID | Notifications | Nicolas Garcelon | Log out

Search for patients

Across the entire data warehouse

lupus 1041/1041

Extend to synonyms : ☐

+Advanced - Rewrite the query

+ Add a full text filter

+ Add a structured filter

+ Time constraints

+ Patient filter

START A SEARCH ?

Query history

Search:

Date	Queries
29/08/2017 10:09	Filtered 1 : Documents containing 'lupus' , Excluding negations
18/10/2016 13:37	Filtered 1 : Documents containing 'LIPA' , Excluding negations
18/10/2016 13:35	Filtered 1 : Documents containing 'mecp2' , Excluding negations
18/10/2016 13:31	Filtered 1 : Documents containing 'lupus' , Excluding negations
12/10/2016 15:30	Filtered 1 : Documents containing 'syndrome de lowe' , Excluding negations

Result | Stats data | Concept | Biology | Genes | Map | Clustering

Across the entire data warehouse :
1041 patients
5972 Documents

Phénotypage haut débit

Objectifs :

Description phénotypique automatisée d'un groupe de patients.

Enrichissement des bases de connaissance sur les maladies rares.

Notre approche : Réutiliser les comptes rendus hospitaliers, en français, sans phase d'annotation manuelle, pour réaliser un phénotypage haut débit de maladies rares

Phénotypage haut débit

Entrepôt de données
500,000 patients

Requête
“syndrome de Rett”

N Patients retrouvés

Agrégation des concepts et ranking

Phénotypes associés à syndrome de Rett:

UMLS Concepts	Freq	TF-IDF
Stéréotypie	71.8	8.06
Crises	55	8.37
Scoliose	51.2	6.71
Epilepsie	50.7	2.96
Syndrome pyramidal	42.6	2.87
Ostéoporose	42.6	2.79
…		
1022 concepts		

Pour chaque patient de l’entrepôt

Comptes rendus

UMLS Metathesaurus
Termes français

UMLS Concepts	Contexte	Certitude	Fréquence
Lupus	Patient	1	3
Anémie	Patient	-1	2
Insuffisance rénale	Family	1	1
Fièvre	Patient	1	5
Asthénie	Patient	1	4

Phénotypage haut débit

A partir des concepts extraits des comptes rendus médicaux des patients retrouvés

A Necker : 22 millions de concepts extraits
14 000 concepts médicaux

Phénotypage haut débit

Exemple : Activated PI3K Delta Syndrome (APDS)

Prise en compte de l'âge d'apparition des symptômes

Médiane de l'âge d'apparition des premiers symptômes

Garcelon et al. Orphanet Journal of Rare Diseases (2018) 13:85
https://doi.org/10.1186/s13023-018-0830-6

Orphanet Journal of Rare Diseases

RESEARCH Open Access

Next generation phenotyping using narrative reports in a rare disease clinical data warehouse

Nicolas Garcelon[1,2,13*], Antoine Neuraz[2,3], Rémi Salomon[1,4], Nadia Bahi-Buisson[1,5], Jeanne Amiel[1,6,7], Capucine Picard[1,8,9], Nizar Mahlaoui[1,8,10,11], Vincent Benoit[1], Anita Burgun[2,3,12] and Bastien Rance[2,12]

Abstract

Background: Secondary use of data collected in Electronic Health Records opens perspectives for increasing our knowledge of rare diseases. The clinical data warehouse (named Dr. Warehouse) at the Necker-Enfants Malades Children's Hospital contains data collected during normal care for thousands of patients. Dr. Warehouse is oriented toward the exploration of clinical narratives. In this study, we present our method to find phenotypes associated with diseases of interest.

Methods: We leveraged the frequency and TF-IDF to explore the association between clinical phenotypes and rare diseases. We applied our method in six use cases: phenotypes associated with the Rett, Lowe, Silver Russell, Bardet-Biedl syndromes, DOCK8 deficiency and Activated PI3-kinase Delta Syndrome (APDS). We asked domain experts to evaluate the relevance of the top-50 (for frequency and TF-IDF) phenotypes identified by Dr. Warehouse and computed the average precision and mean average precision.

Results: Experts concluded that between 16 and 39 phenotypes could be considered as relevant in the top-50 phenotypes ranked by descending frequency discovered by Dr. Warehouse (resp. between 11 and 41 for TF-IDF). Average precision ranges from 0.55 to 0.91 for frequency and 0.52 to 0.95 for TF-IDF. Mean average precision was 0.79. Our study suggests that phenotypes identified in clinical narratives stored in Electronic Health Record can provide rare disease specialists with candidate phenotypes that can be used in addition to the literature.

Conclusions: Clinical Data Warehouses can be used to perform Next Generation Phenotyping, especially in the context of rare diseases. We have developed a method to detect phenotypes associated with a group of patients using medical concepts extracted from free-text clinical narratives.

Keywords: Data warehouse, Next generation phenotyping, Data mining, Rare diseases, Natural language processing

Principaux cas d'usage :

Des fonctionnalités centrées patient

Faciliter l'exploration du dossier du patient

Dr WareHouse ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Mes demandes | Outils | Patient Nom / IPP | Notifications | Nicolas Garcelon An

Patient M,

Documents | Biologie | TimeLine | Parcours | Cohorte | Concepts | Similarité

biopsie RECHERCHER

38 documents trouvés

ORBIS CR - courrier
*Histologie de la **biopsie** realisee le : lesions de [...]*

DIAMIC Biopsie cutanée avec colo spé pour diag d'affection non carcinologique
*La **biopsie** communiquee fixee a ete analysee sur plusieurs niveaux de coupe [...] Le reste de la **biopsie** est sans particularite [...]*

ORBIS CR - courrier
__Biopsie__ cutanee realisee le : aspect histologique observe montre d'une part des lesions de [...] Realisation de __biopsies__ cutanees au bloc operatoire sous anesthesie generale le . [...] Bilan dermatologique en attente (__biopsies__ cutanees en cours) [...]

ORBIS CR - courrier
Derniere __biopsie__ cutanee en sept [...] __Biopsies__ : aspect de vascularite leucocytoclasique malgre l'absence de necrose fibrinoide [...] Prevoir une __biopsie__ cutanee pour nouvelle histologie + IFD plus ou moins microscopie electronique si [...] realisation dans le meme temps de la FOGD une __biopsie__ cutanee des lesions de purpura des [...]

ORBIS CR - courrier
__Biopsie__ cutanee en 09/2014 (IFD sur lesion ancienne --> difficilement interpretable): depots IgG et C3 [...] __Biopsie__ : aspect de vascularite leucocytoclasique malgre l'absence de necrose fibrinoide [...] AG prevue en pour une endoscopie : prevoir __biopsie__ cutanee pour nouvelle histo + IFD [...]

ORBIS CR - courrier
__Biopsie__ cutanee en (IFD sur lesion ancienne --> difficilement interpretable): depots IgG et C3 [...] __Biopsie__ : aspect de vascularite leucocytoclasique malgre l'absence de necrose fibrinoide [...] AG prevue en pour une endoscopie : prevoir __biopsie__ cutanee pour nouvelle histo + IFD [...]

ORBIS CR - courrier
- __Biopsie__ cutanee : aspect histologique, malgre l'absence de necrose fibrinoide, fait conclure a une vascularite [...] Il n'y a pas eu de __biopsie__ [...] Durant le meme geste anesthesique, une PBH et une __biopsie__ cutanee ont ete realisees [...]

Biopsie cutanée avec colo spé pour diag d'affection non carcinologique, par , le

SERVICE D'ANATOMIE ET DE CYTOLOGIE PATHOLOGIQUES
Groupe Hospitalier Necker-Enfants Malades
149 rue de Sevres 75743 PARIS CEDEX 15

Secretariat : [TEL]
Fax : [TEL]

Chef de Service

Recu le : [DATE]
GASTRO-ENTÉROLOGIE ENF
Patient : [NOM]
[PRENOM]
Sexe : M
NIP : [IPP] / NDA :

NECKER-ENFANTS MALADES
75743 PARIS CEDEX 15age de[DATE]

A la demande du :
Copie au(x) :
NSLE /NCGE

RENSEIGNEMENTS CLINIQUES :
Poussee de purpura non infiltree des membres. Contexte de maladie de Hirschsprung avec thrombopenie fluctuante. Purpura rhumatoide ? Anomalie du tissu elastique ?

NATURE DU PRELEVEMENT :

Des fonctionnalités centrées patient

Intégration des images dans l’entrepôt de données

Des fonctionnalités centrées patient

Visualisation du parcours du patient dans **le temps**

Des fonctionnalités centrées patient

Visualisation du parcours de soin d'un patient

DWH ©Imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Contact | Recherche rapide : | Pa

Patient M,
2 ans

Documents | Biologie | TimeLine | Parcours | Cohorte | Concepts | Similarité

Concepts

Affichage simplifié - Affichage pour validation

Search:

Concepts	Nb documents
Epidermolyse bulleuse simple	5
vesicule	4
dowling meara	4
Surinfection	3
epidermolyse bulleuse	3
Epidermolyse bulleuse dystrophique	3
Prise de poids	3
Retrait	3
endocardite d'osler	2
DYSPHAGIE	2
immobilisation prolongee	2
hyperkeratose	2
strabisme	2
infection a staphylocoque	2
hypermetropie	2
pachyonychie	2
malnutrition	2

Les concepts médicaux sont automatiquement extraits à partir des comptes rendus hospitaliers

Quand le moteur de recherche ne suffit plus

Patients non diagnostiqués

Un patient diagnostiqué

Phénotype complexe

Moteur de recherche

Tous les critères phénotypiques

Similarité avec le patient index

Proximité phénotypique

Dr WareHouse ©imagine

Entrepôt de données

Accueil | Moteur de recherche | Mes requêtes | Mes cohortes | Mes demandes | Outils | Contact | Patient Nom / IPP | Admin | Nicolas Garc

Nom Prénom né le 01/01/2000

Documents | Biologie | TimeLine | Parcours | Cohorte | Concepts | Similarité

Publication : Garcelon et al, 2017, JBI

Intégration du calcul de similarité phénotypique dans Dr WH

Patient	Similarité	Concepts communs	Concepts patient
Patient 1 (statut cohorte APDS : 1) (statut cohorte APDS2 : 1) (statut cohorte DIH inclus CEREDIH : 1) Afficher le comparateur	81	Conjonctivite, Toux, DYSPHAGIE, dyspnee, Fievre, Hyperplasie, deficit immunitaire, Lymphopenie, mycose, Otite, Douleur, Sinusite, syndrome d'apnee du sommeil, Surinfection, Sinusite chronique, larmoiement, Rhinorrhee purulente, syndrome hyper igm, Adenopathie, Intervention chirurgicale, Gene respiratoire, rhinosinusite, Cicatrice, Recidive	Vegetations, **stenose pharyngee**, Gaz, Epanchement, Retrait, hypertrophie des amygdales, suppuration, pi3k, Thrombus, deficit immunitaire humoral, Ptosis, Croute, septicemie, dacryocystite, hypertrophie, Observation, obstruction laryngee, Bouton, Respiration, Aphte, tor, Complication, Thrombose, Hernie inguinale bilaterale, obstruction chronique des voies aeriennes, nutrition, Stenose laryngee, Oedeme, Hypopnee, vesicule, obstruction nasale bilaterale, Deficit en IgG, Apnee, Obstruction, Lesions muqueuses, Nodule, SAOS, Circulation collaterale, Douleurs abdominales, adherence, hypertrophie des vegetations adenoides, apds, appareil respiratoire
Patient 2 (statut cohorte APDS : 1) (statut cohorte APDS2 : 1) (statut cohorte DIH inclus CEREDIH : 1) Afficher le comparateur	80	Anorexie, Asthenie, dilatation des bronches, rhinite, infection, Conjonctivite, Toux, Fievre, gastroenterite, cephalees, Hyperplasie, deficit immunitaire, Lymphopenie, Otite, Douleur, syndrome d'apnee du sommeil, Surinfection, baisse, Sinusite chronique, Adenopathie cervicale, Tolerance, Ecoulement, Rhinorrhee purulente, syndrome hyper igm, Adenopathie, Pansinusite, Rhinorrhee, Cicatrice, Recidive	Enuresie, bruits, Aphte, Hypopnee, 3 6, bronchite chronique, Ptosis, Toux nocturne, Hypogammaglobulinemie, somnolence diurne, deviation de la cloison nasale, Gastro-enterite aigue, ANEMIE MICROCYTAIRE, Diphterie, Apnee, 3 4, Anemie, Injection intraveineuse, baisse de l'acuite visuelle, **Conjonctivite chronique**, Ronchi, crepitants, deficit de l'immunite humorale, Conjonctivite purulente, pai, lymphoproliferation, erytheme localise, Infections respiratoires, PIK3R1, cauchemars, ronflement, Obstruction, Enuresie nocturne, Gynecomastie, hypertrophie, Urgences, endormissement, Epistaxis, Rhinite chronique, **pi3k**, **apds**, Irritable, bronchorrhee, Agitation, Prise de poids, Toux chronique, Taux d'anticorps, cavite, vomissement, Sueurs nocturnes, SAOS, Frequence cardiaque maximale, sommeil agite, Respiration buccale, Toux persistante
Patient 3 (statut cohorte DIH inclus CEREDIH : 1) (statut cohorte APDS : 1)		douleur abdominale, Asthme, ATELECTASIE, dilatation des bronches, rhinite, infection, Toux, Diarrhee,	vesicule, Decouverte fortuite, Toux chronique, Douleurs abdominales, Rhinite chronique, regression, bruits, PIK3CD, HPA3, cavite, PIK3, Vegetations, distension, Diarrhee chronique, hypertrophie, Retrecissement, CD31, Purpura, **apds**, ANEMIE MICROCYTAIRE, consideration, eruption, fenetre aorto-pulmonaire, adenopathies hilaires, Sequestration splenique, positive, Hematome, douleur dentaire, Hypersplenisme, Cytolyse hepatique, infection virale, Toux nocturne, Moelle riche, EMPHYSEME, Hypogammaglobulinemie, selles glaireuses, malformation pulmonaire, Respiration, Sueurs nocturnes, 70%, Pancytopenie, pi3k, Detresse

Recherche translationnelle et similarité

Une petite histoire

APDS : 2 gènes
PIK3CD (Agnulo et al, 2013)
PIK3R1 (Deau et al, 2014)

Evaluation sur la cohorte APDS

?

Dernière visite en 2012
Déficit immunitaire non étiquetée

Délétion sur le gène PIK3R1

Dr Warehouse : un entrepôt de données orienté document

Dr Warehouse publié
en février 2018

Dr Warehouse® est open source

Journal of Biomedical Informatics 80 (2018) 52–63

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

A clinician friendly data warehouse oriented toward narrative reports: Dr. Warehouse

Nicolas Garcelon[a,b,*], Antoine Neuraz[b,c], Rémi Salomon[a,d], Hassan Faour[a], Vincent Benoit[a], Arthur Delapalme[a], Arnold Munnich[a,e,f], Anita Burgun[b,c], Bastien Rance[b,g]

[a] Institut Imagine, Paris Descartes Université Paris Descartes-Sorbonne Paris Cité, Paris, France
[b] INSERM, Centre de Recherche des Cordeliers, UMR 1138 Equipe 22, Université Paris Descartes, Sorbonne Paris Cité, Paris, France
[c] Department of Medical informatics, Hôpital Necker-Enfant Malades, Assistance Publique des Hôpitaux de Paris, Paris, France
[d] Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
[e] Département de génétique médicale, Hôpital Necker-Enfants Malades, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France
[f] Centre de Référence des Maladies Osseuses Constitutionnelles, INSERM UMR 1163, Laboratoire de bases moléculaires et physiopathologiques de l'ostéochondrodysplasie, Paris Descartes-Sorbonne Paris Cité University, AP-HP, Institut Imagine, 75015 Paris, France
[g] Hôpital Européen Georges Pompidou, Assistance Publique-Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité, France

ARTICLE INFO

Keywords:
Software
Computational biology
Method
Data warehouse
Rare diseases
Electronic health records
Information storage and retrieval
Text-mining

ABSTRACT

Introduction: Clinical data warehouses are often oriented toward integration and exploration of coded data. However narrative reports are of crucial importance for translational research. This paper describes Dr. Warehouse®, an open source data warehouse oriented toward clinical narrative reports and designed to support clinicians' day-to-day use.

Method: Dr. Warehouse relies on an original database model to focus on documents in addition to facts. Besides classical querying functionalities, the system provides an advanced search engine and Graphical User Interfaces adapted to the exploration of text. Dr. Warehouse is dedicated to translational research with cohort recruitment capabilities, high throughput phenotyping and patient centric views (including similarity metrics among patients). These features leverage Natural Language Processing based on the extraction of UMLS® concepts, as well as negation and family history detection.

Results: A survey conducted after 6 months of use at the Necker Children's Hospital shows a high rate of satisfaction among the users (96.6%). During this period, 122 users performed 2837 queries, accessed 4,267 pa-

Conclusion

Faciliter la réutilisation de la **mémoire collective** à la fois pour la recherche et pour la prise en charge des patients.

Remettre dans les mains des médecins les outils des data sciences, pour leur permettre **d'explorer leurs propres données**.

Décloisonner les données issues de la **clinique et de la recherche**.

Ne plus faire de bases de données sans questions scientifiques claires !!

Accélérer la recherche translationnelle

Imagine – Data science Platform Team :

Nicolas Garcelon
Vincent Benoit
Hassan Faour
Marc Vincent

Une démo existe ici :

http://drwarehouse.org/demonstration

codoc

nicolas.garcelon@institutimagine.org
arthur@drwarehouse.org

Des questions ?

Shire