



# Secondary use of Clinical Data for Medical Research

Pr Marc CUGGIA

INSERM UMR 936

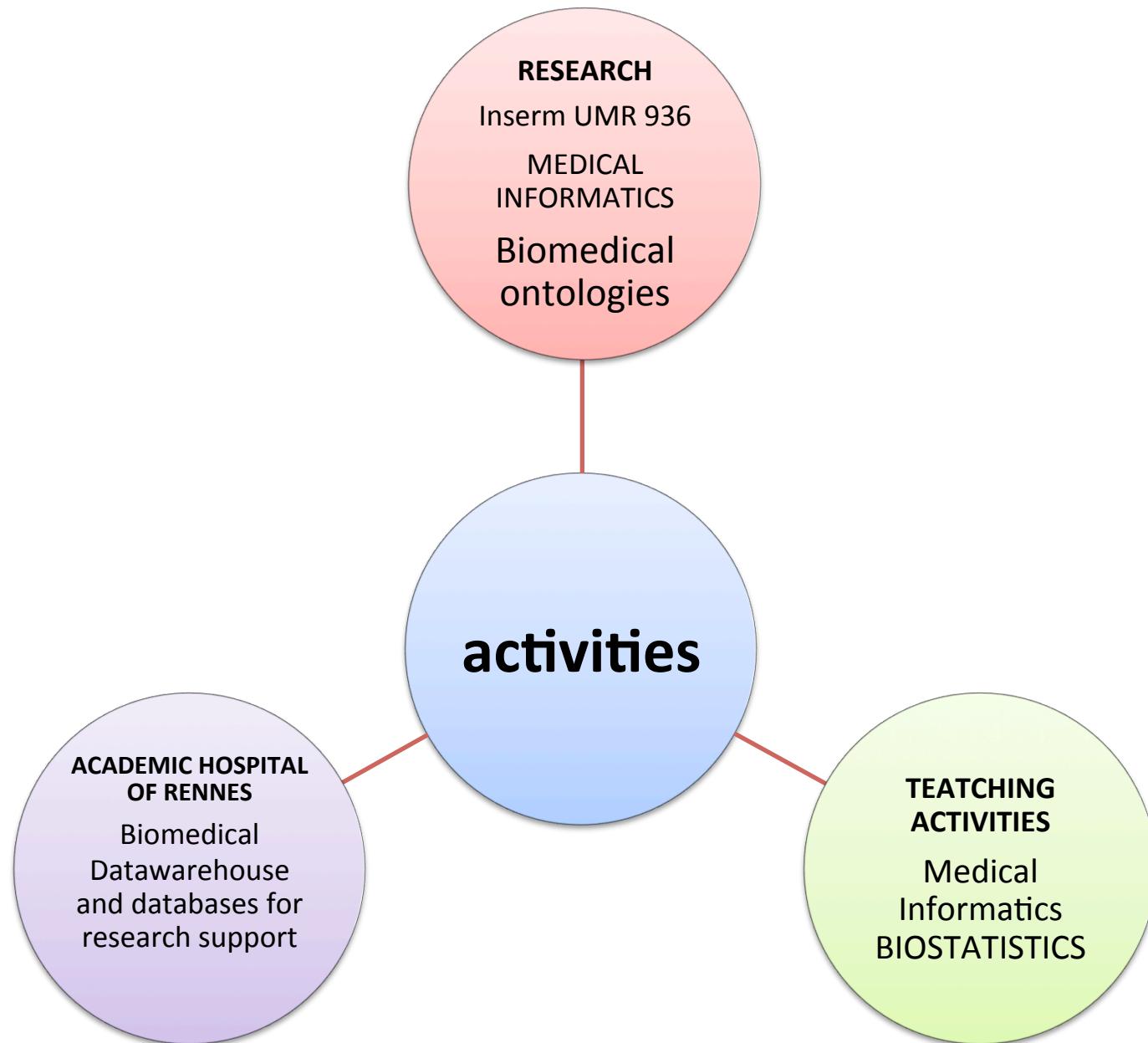
Faculté de médecine

Université de Rennes 1 - CHU de Rennes

CrREDIBLE Workshop

# Agenda

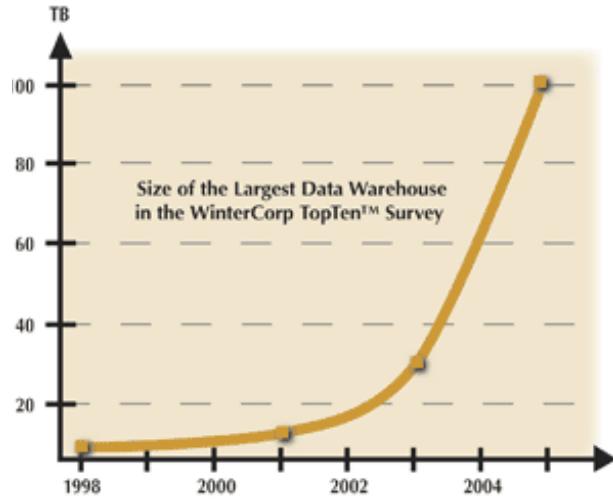
- Context and needs in medical research
- Patient centered approach : ASTEC
- Population centered approach :
  - Clinical DataWarehouse (CDW):
    - GOOGLE
  - CDW Network :
    - EHR4CR / DEBUG IT
- Current issues and perspectives



# Medical context

- Computerization of Patient data
  - Health information systems : HIS, EHR, Health network records, national EHR (DMP)
  - Other : Health insurance data (medication)
  - patients website : forums, tweets...
- Heterogeneity
  - multi domains
  - Variable data quality
  - Different formats :
    - numeric or symbolic, Coded or full text +++
  - No semantic unity : local or international terminologies
  - Components not interoperable

# Technologic context



- Data flood : producing and collecting exponential data
- **Storage capacity**
- **Processing capacity**
- **Data access capabilities (networks)**

Manual data retrieval and analysis is becoming impossible



Mining Biomedical data is now strategic

# Needs in medical research

# Needs : clinical research

- Clinical research : new therapy or diagnostic strategies
- Recruiting patients for studies
  - is expensive : \$250 par candidate
  - is difficult : low recruitment → less studies → no research and less money
- Study in 2006 – Havard
- *Benefits for using clinical datawarehouses*
  - Save \$7 millions on the recruitment cost
  - extrapolation on the clinical research : \$94 à \$136 millions

[1] R. Nalichowski, D. Keogh, H. C. Chueh, and S. N. Murphy, “Calculating the benefits of a Research Patient Data Repository,” AMIA ... Annual Symposium Proceedings / AMIA Symposium. AMIA Symposium, p. 1044, 2006.

# Other medical needs

- Personnalized medicine :
  - In Silico detection of biomarkers for predicting diseases and comorbidities
- Health economics and evaluation of professionnal practices
  - Patient trajectory though the healthcare system
- Public Health :
  - biosurveillances, epidemiology

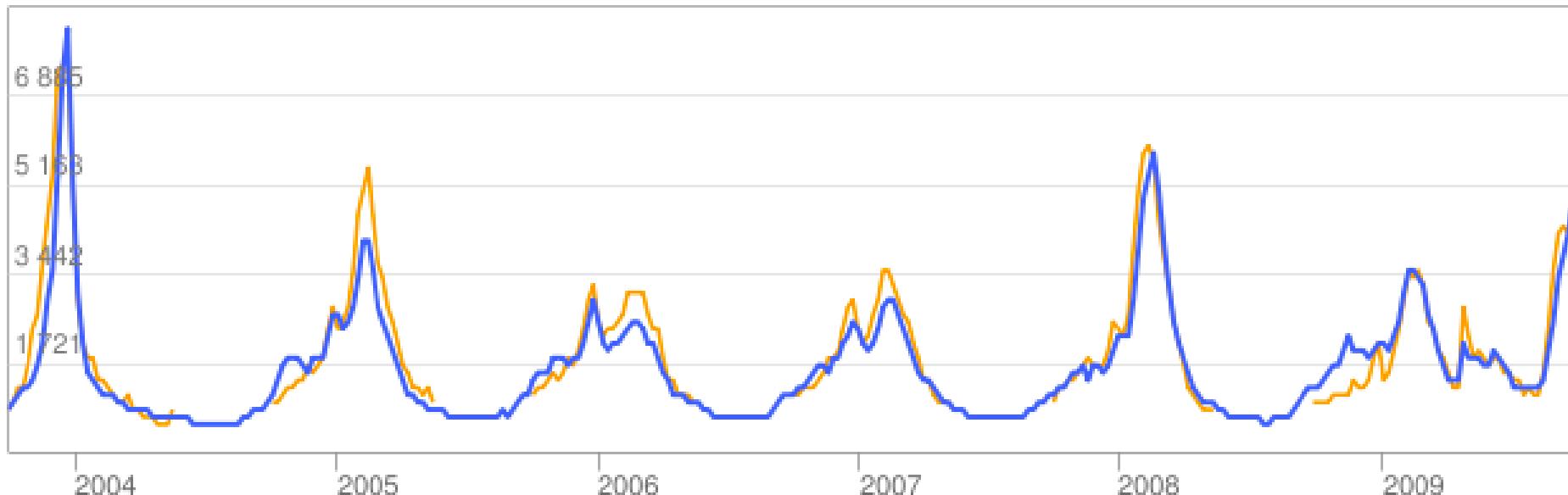
Estimations historiques

Voir les données pour : États-Unis

## États-Unis - Propagation du virus

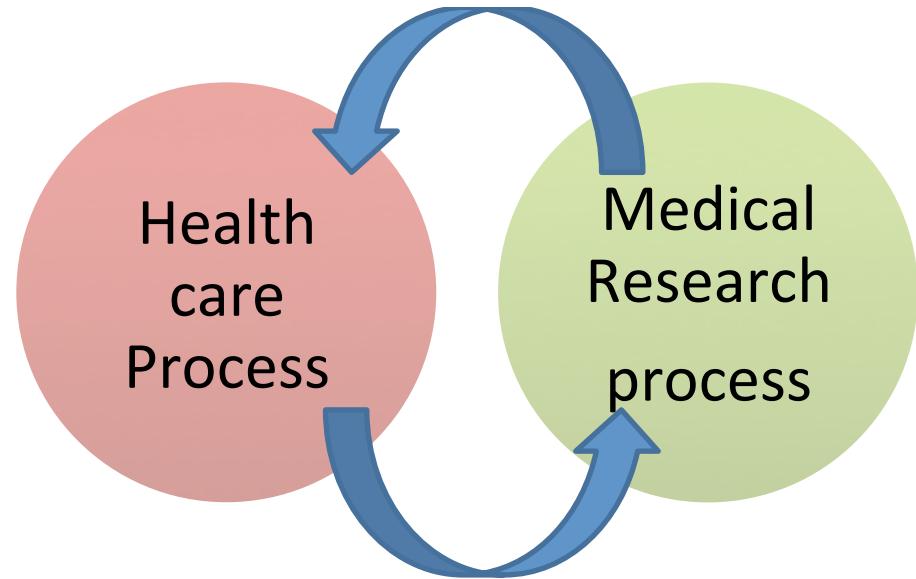
Estimation de la grippe

● Estimation Google Suivi de la grippe ● Données pour : États-Unis

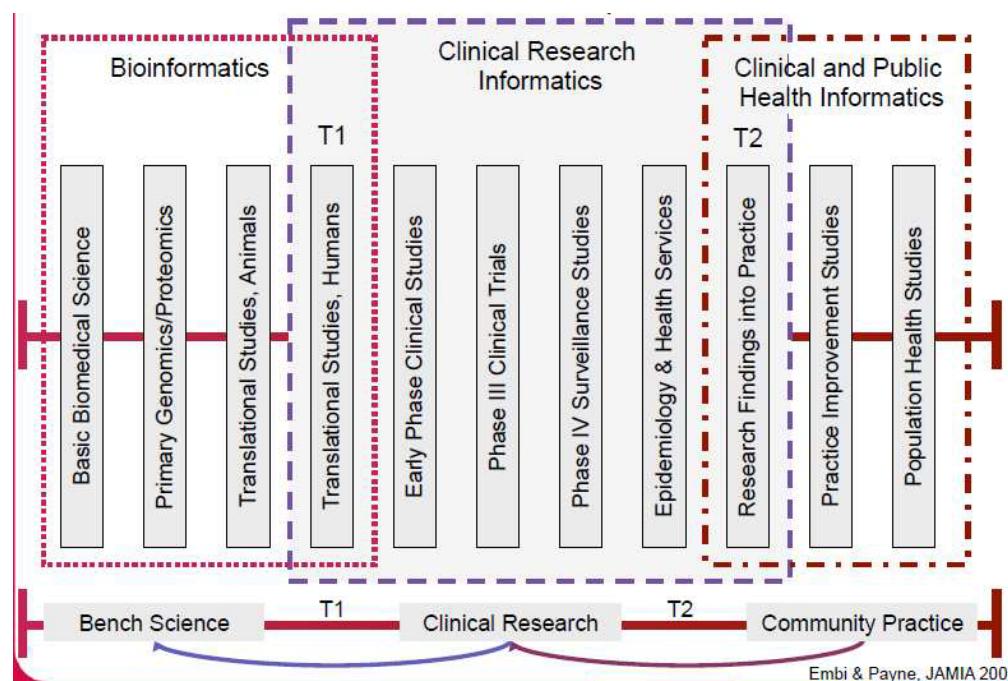


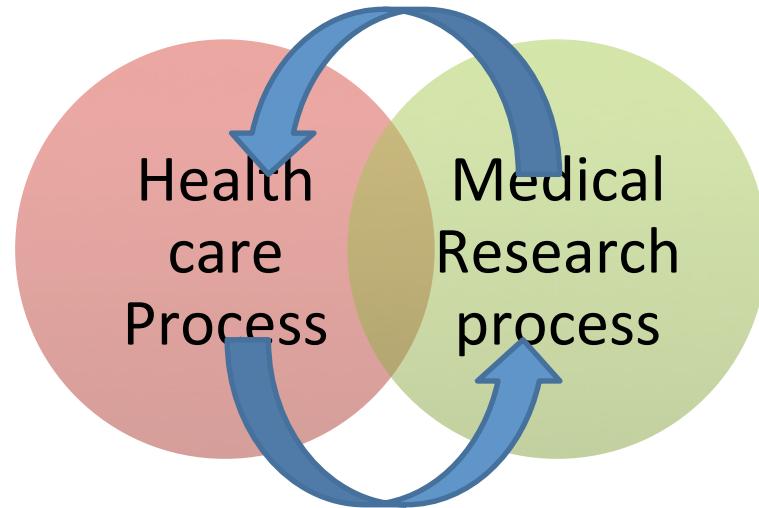
États-Unis : Données publiques sur le syndrome grippal (ILI) fournies par les [Centres américains de prévention et de contrôle des maladies](#).

[1] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, et L. Brilliant, « Detecting influenza epidemics using search engine query data », Nature, vol. 457, n° 7232, p. 1012-1014, févr. 2009.



## Translational medicine



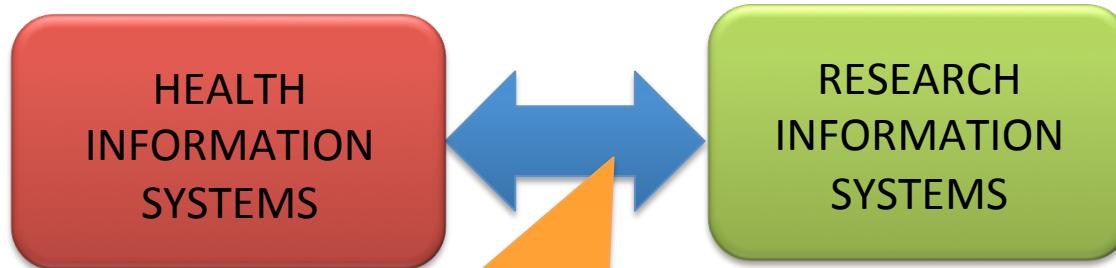
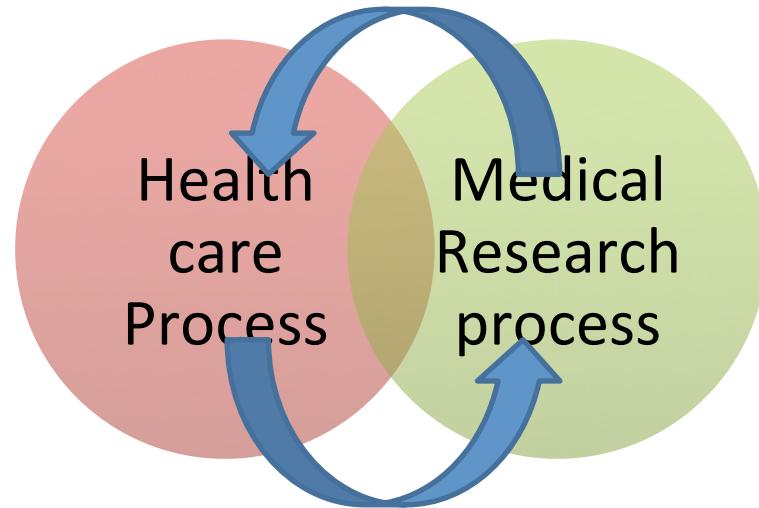


HEALTH  
INFORMATION  
SYSTEMS

HIS and Hospital EHR  
Primicare Care EHR  
National EHR

RESEARCH  
INFORMATION  
SYSTEMS

Epidemiologic Registry databases  
Clinical trial databases  
OMICs and Biobank Databases  
Knowledge databases (e.g MEDLINE)



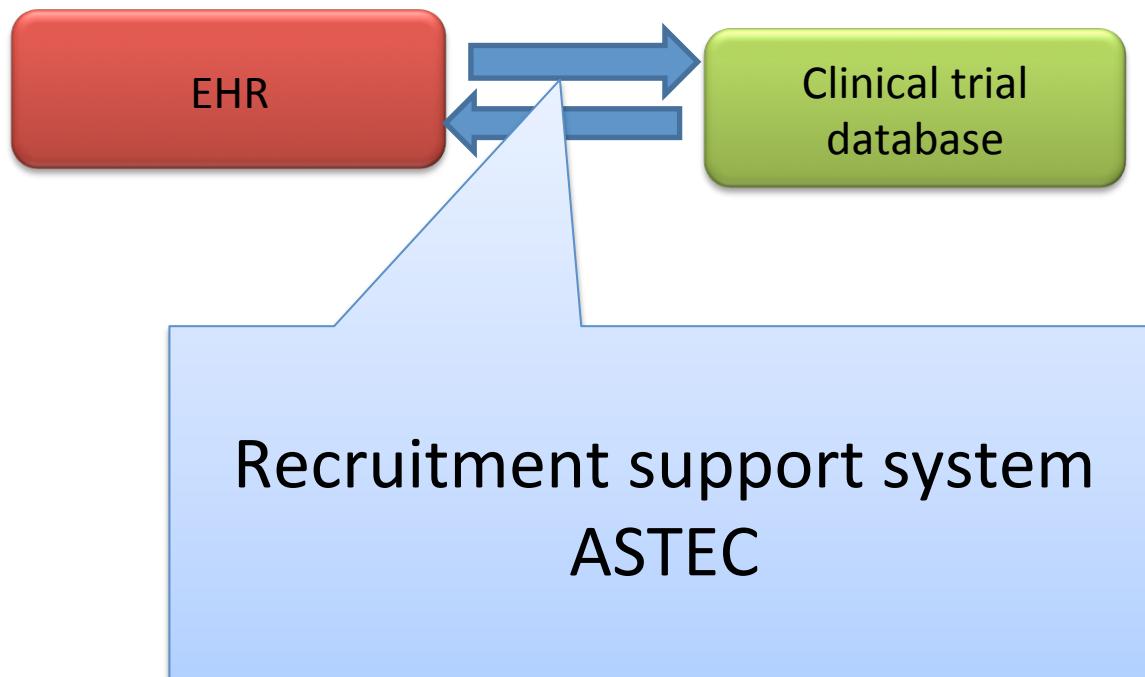
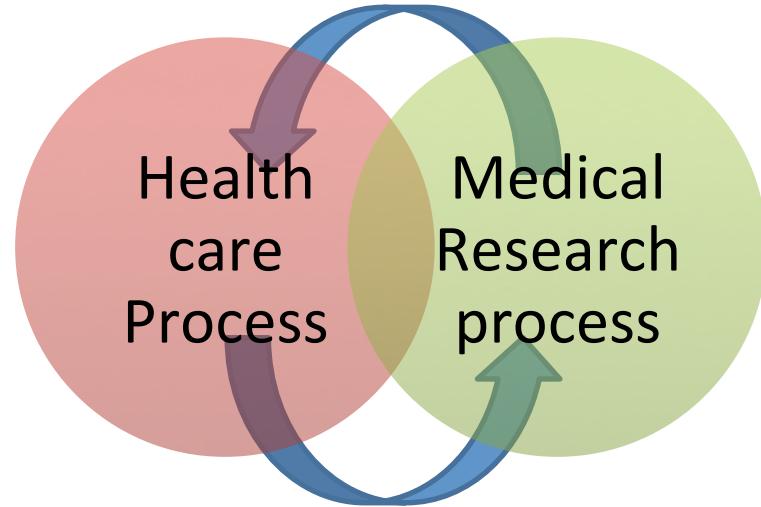
Interoperability (syntactic and semantic) Standards

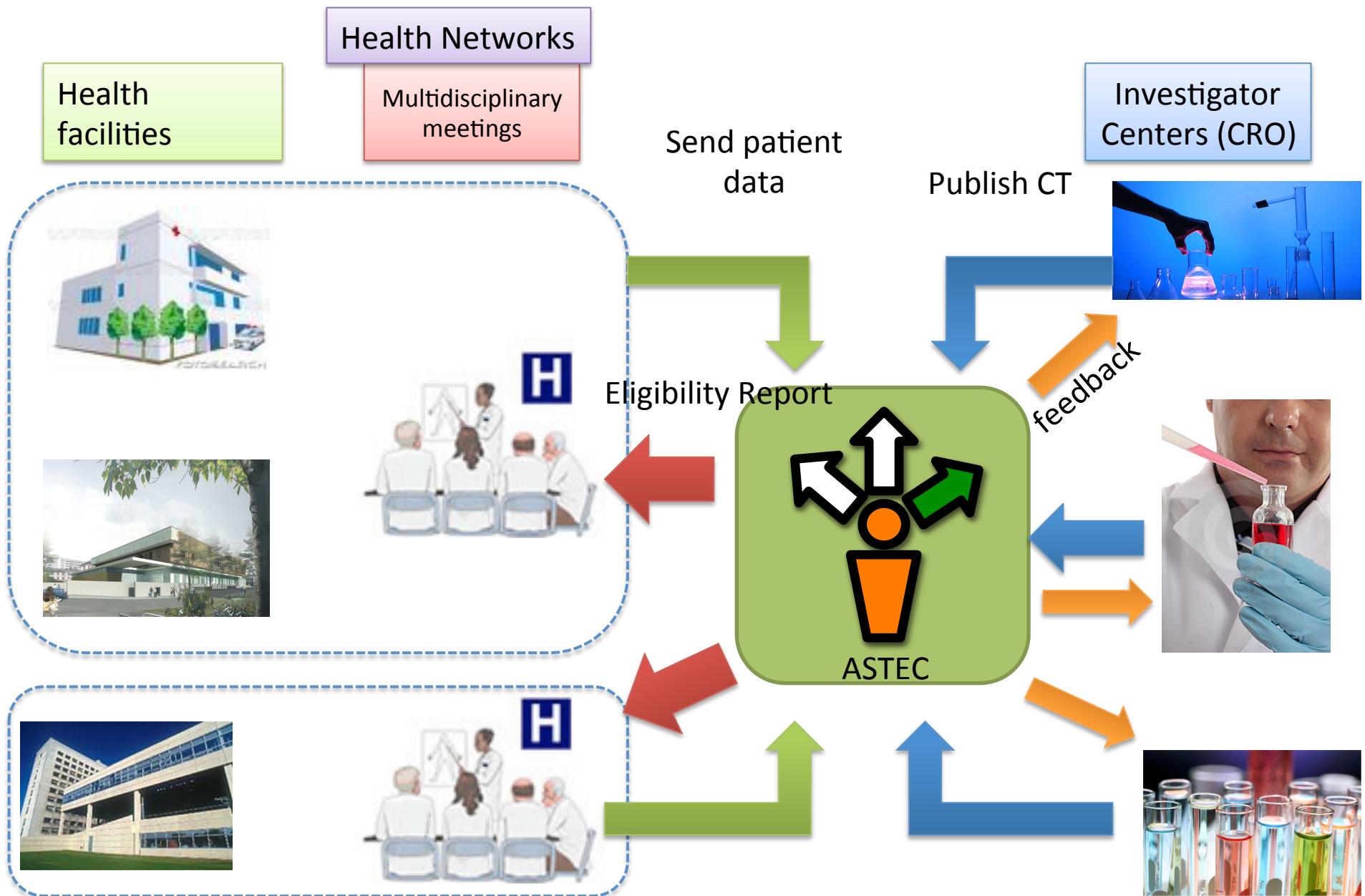
- CDISC
- HL7
- DICOM

IHE : Integrating Health Enterprise

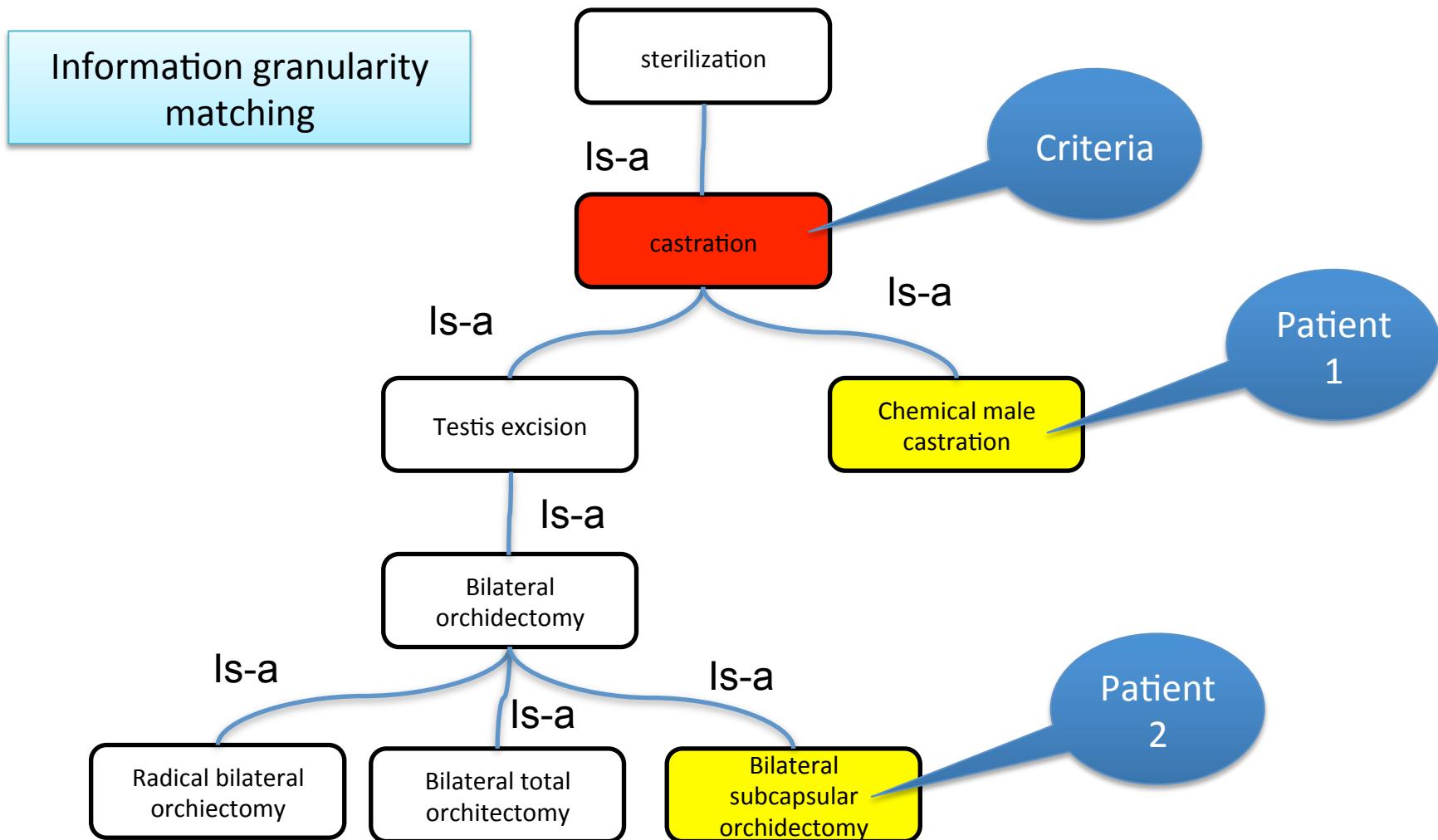
# Examples of tools Patient centered approach

ASTEC : recruitment system support

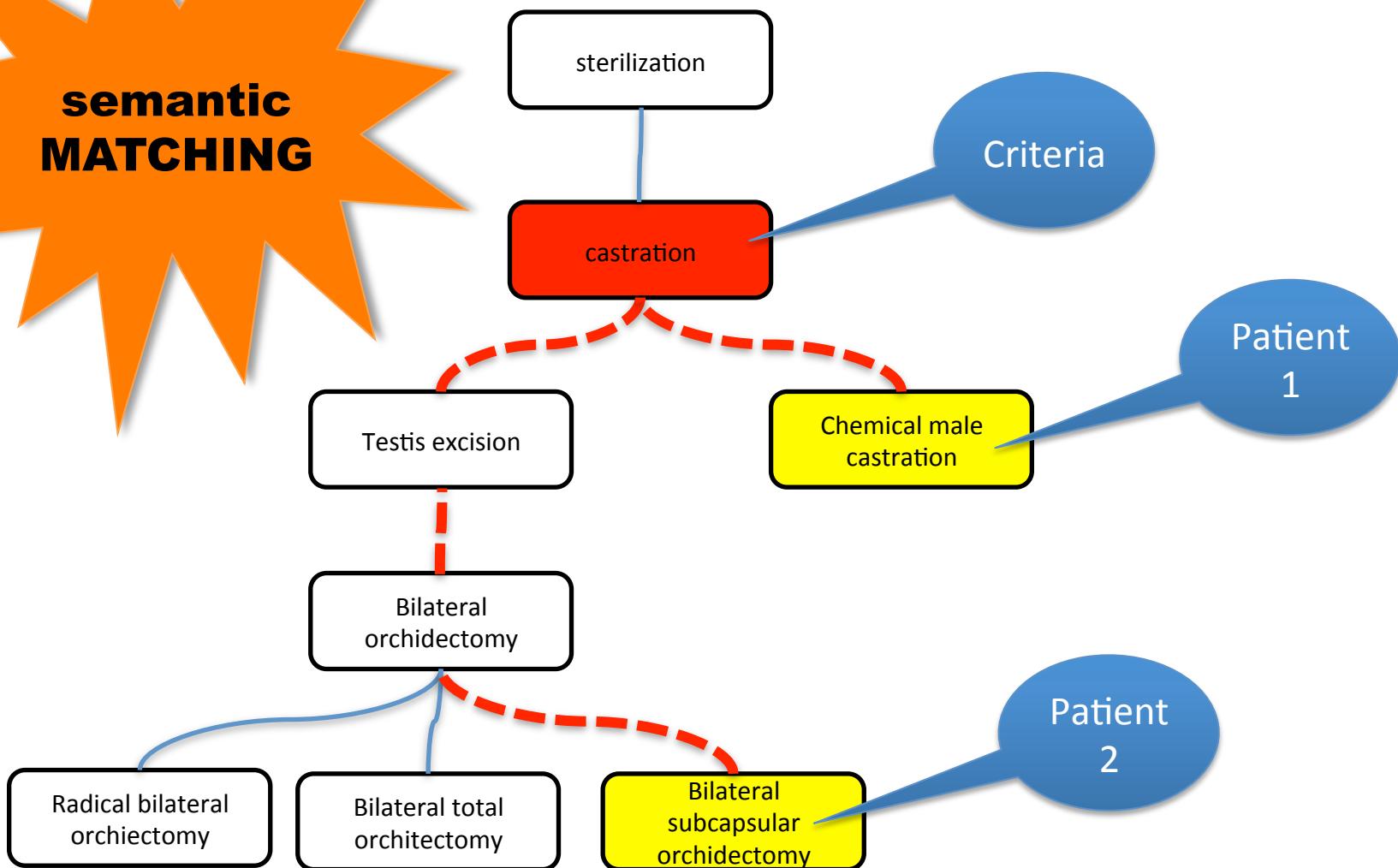




# NCIT cancer ontology for reasonning

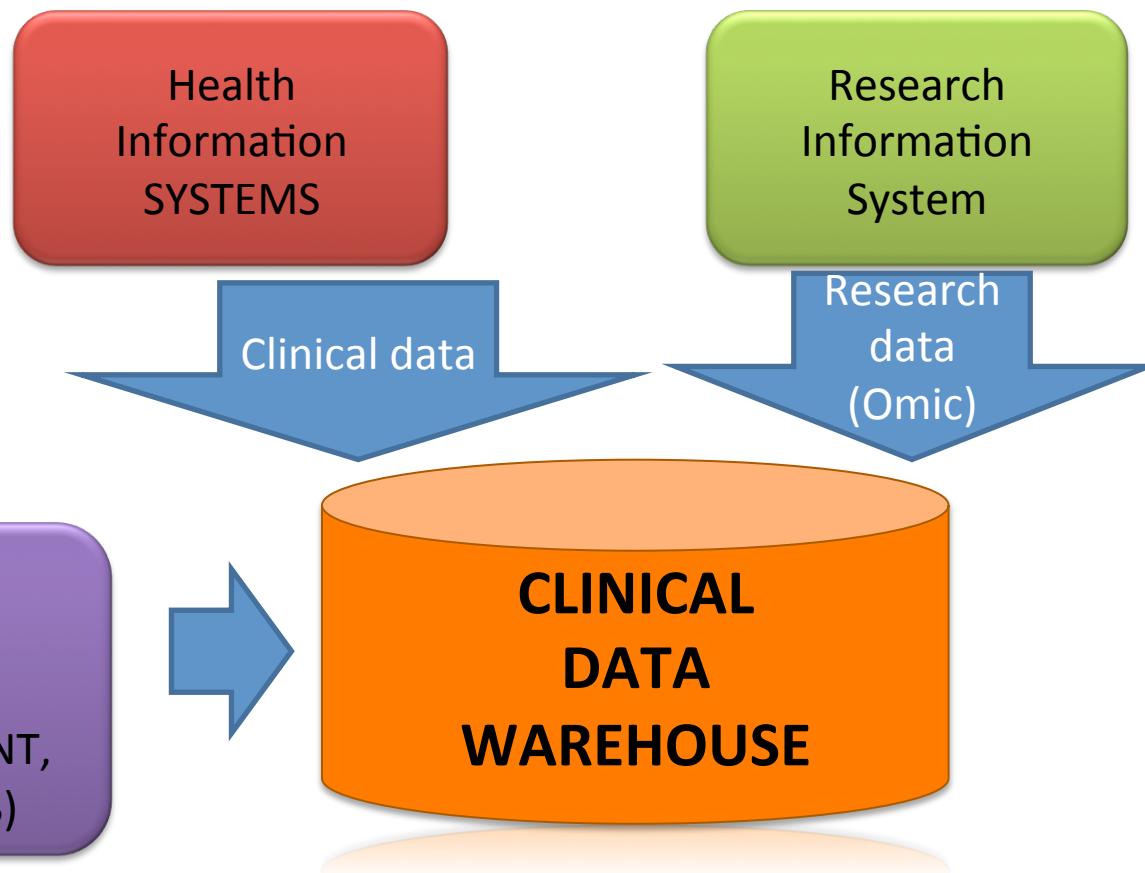
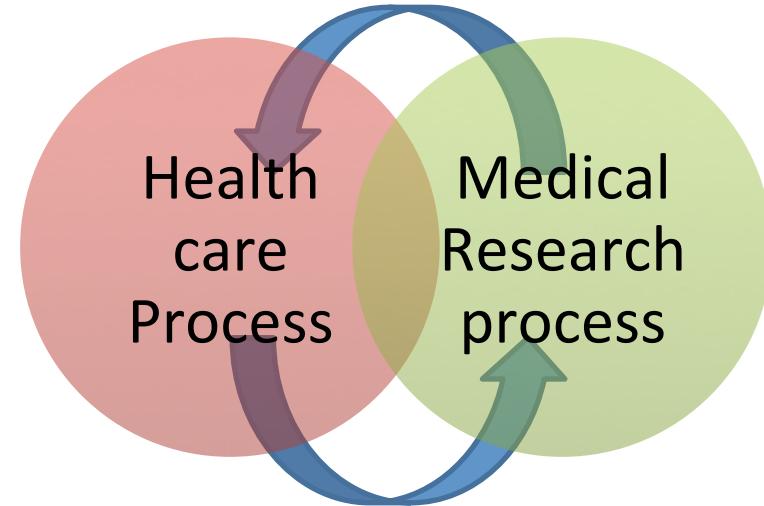


# semantic MATCHING

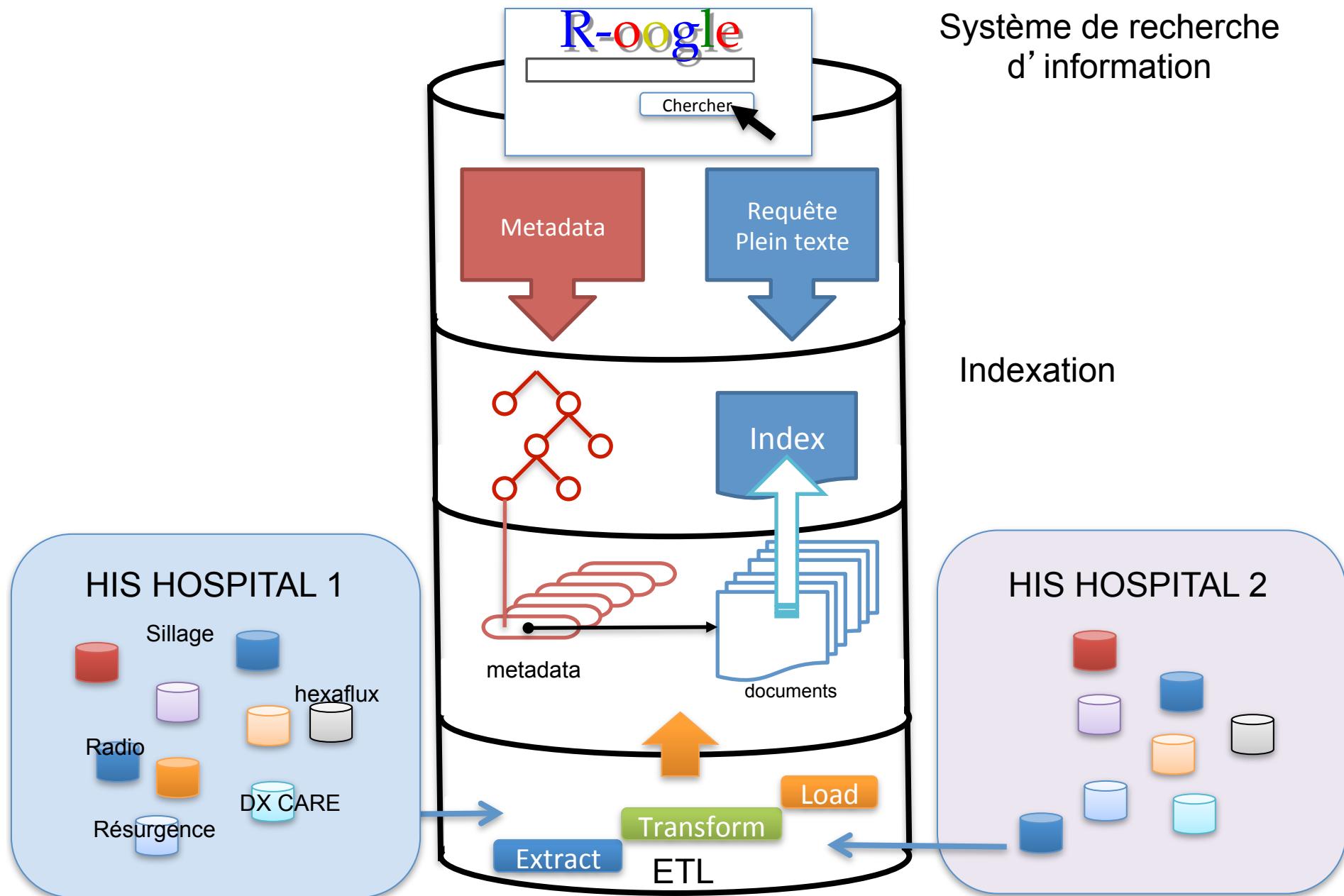


# Examples of tools Population centered approach

Clinical Data Warehouse



# Entrepôt de données: ROOGLE



# Some examples of CDW technologies

- I2B2 Informatics for integrating biology and the bedside
  - Harvard
  - Translationnal medicine
- STRIDE
  - Stanford
  - 3 hospitals
  - Détection de cohorte
- ROOGLE :
  - Rennes / Necker Institut Imagine / BREST

Murphy SN, Mendis ME, Berkowitz DA, Kohane I, Chueh H. Integration of clinical and genetic data in the i2b2 architecture. AMIA Annu Symp Proc. 2006; p.1040. PMID:17238659.

1. Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE--An integrated standards-based translational research informatics platform. AMIA Annu Symp Proc. 2009;2009:391–395.

[1]M. Cuggia, N. Garcelon, B. Campillo-Gimenez, T. Bernicot, J.-F. Laurent, E. Garin, A. Happe, et R. Duvaufier, « Roogle: an information retrieval engine for clinical data warehouse », Stud Health Technol Inform, vol. 169, p. 584-588, 2011.

# Roogle features

- Combine queries on structured and free text data (+++)
- NLP methods
  - Negation or uncertainty detection
  - Semantic annotation, concept extraction
  - Using UMLS (United Medical Language System)
    - Combining all reference vocabularies

## Formulaire de recherche : Intégralité de l'entrepôt (710727 patients)

+ Requêtes enregistrées

- Critères généraux

Date du document(JJ/MM/AAAA):

Unité(s) Fonctionnelle(s):

Age du patient pdt la consultation:

Age du patient maintenant:  30  90

Sexe : H  F

Ne pas prendre en compte les patients décédés a l'hôpital

:

- Recherche en Full-text

Inclure  Type de doc  texte  UF :

rasagiline OR selegiline OR otracel OR deprenyl OR sinemet OR modopar Or Ldopa

Inclure périphrases négatives :

Enrichissement sémantique :

Pour cette expression : 633 patients

Inclure  Type de doc  texte  UF :

dyskinesie

Inclure périphrases négatives :

Enrichissement sémantique :

Pour cette expression : 798 patients

ajouter requête

- Recherche structurée

Filtrer :  hypertensive

liste concepts - arbre

Actes (CCAM)

Requête

Statistiques

Map

R

Utilisez le formulaire à gauche pour créer votre requête

Age du patient : de 30 à 90

Les documents contenant l'expression "rasagiline OR X selegiline OR otracel OR deprenyl OR sinemet OR modopar Or Ldopa" (633 patients patients)

Les documents contenant l'expression "dyskinesie" X (798 patients patients)

Recherche de maladie de Parkinson dans Diagnostics (CIM-10) (2828 patients) X O

Absence de maladies hypertensives dans Diagnostics (CIM-10) (77306 patients) X O

+ Ajouter une contrainte temporelle sur les items ci-dessus

Requête sql

## Requête

Utilisez le formulaire à gauche pour créer votre requête

Les documents contenant l'expression "atrophie cereb%" X

Les documents contenant l'expression "atrophie optique" X

Les documents contenant l'expression "IRM" X

Absence de diabète sucre X

+ Ajouter une contrainte temporelle sur les items ci-dessus

l'évènement  ▾  
Précède ▾ de Maximum ▾ 2 mois ▾  
l'évènement  ▾



TEMPORAL  
CONSTRAINTS

# Résultat : étude de faisabilité

**Formulaire de recherche : Intégralité de l'entrepôt (14660 patients)**

- + Requêtes enregistrées
- + Critères généraux
- Recherche en Full-text

---

Inclure  SUSIE  texte  UF :   
retinopathie

Inclure periphrases négatives :

Enrichissement sémantique :

[ajouter requête](#)

---

+ Recherche structurée

Les documents, de type SUSIE (avec enrichissement sémantique) contenant l'expression "re"

Intersection des critères sur  le patient

Ne pas afficher les patients déjà exclus de l'analyse :  oui  non

---

**Création des paniers**

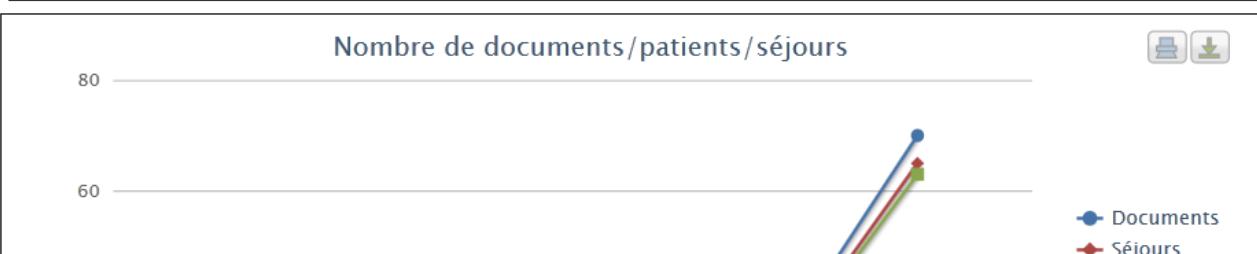
- + Panier en cours
- + Liste des paniers

---

**Alimentation des DataMart/études**

- + Liste des DataMart/études
- + Alimenter / créer un datamart/étude à partir des résultats

**Administration**

Requête	Statistiques	Map	Résultats	Export Excel
<p>Nombre de documents : 78 Nombre de patients : 72 Nombre de séjours : 74</p> <p>Signes / symptômes et diagnostics</p> 				
<p>Nombre de documents/patients/séjours</p>  <p>Documents (blue line) Séjours (red line)</p>				

SEMANTIC  
OVERVIEW  
OF THE  
RESULT

→

**Formulaire de recherche : Intégralité de l'entrepôt (14660 patients)**

+ Requêtes enregistrées

+ Critères généraux

- Recherche en Full-text

Inclure  Type de doc  texte  UF :   
atrophie optique

Inclure périphrases négatives :

Enrichissement sémantique :

[ajouter requête](#)

**+ Recherche structurée**

Intersection des critères sur  le patient  le séjour

Ne pas afficher les patients déjà exclus du panier en cours  
 oui  non

**Création des paniers**

+ Panier en cours

+ Liste des paniers

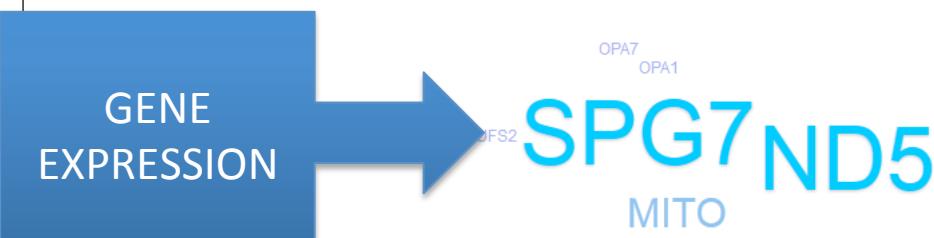
**Alimentation des DataMart/études**

+ Liste des DataMart/études

+ Alimenter / créer un datamart/étude à partir des résultats

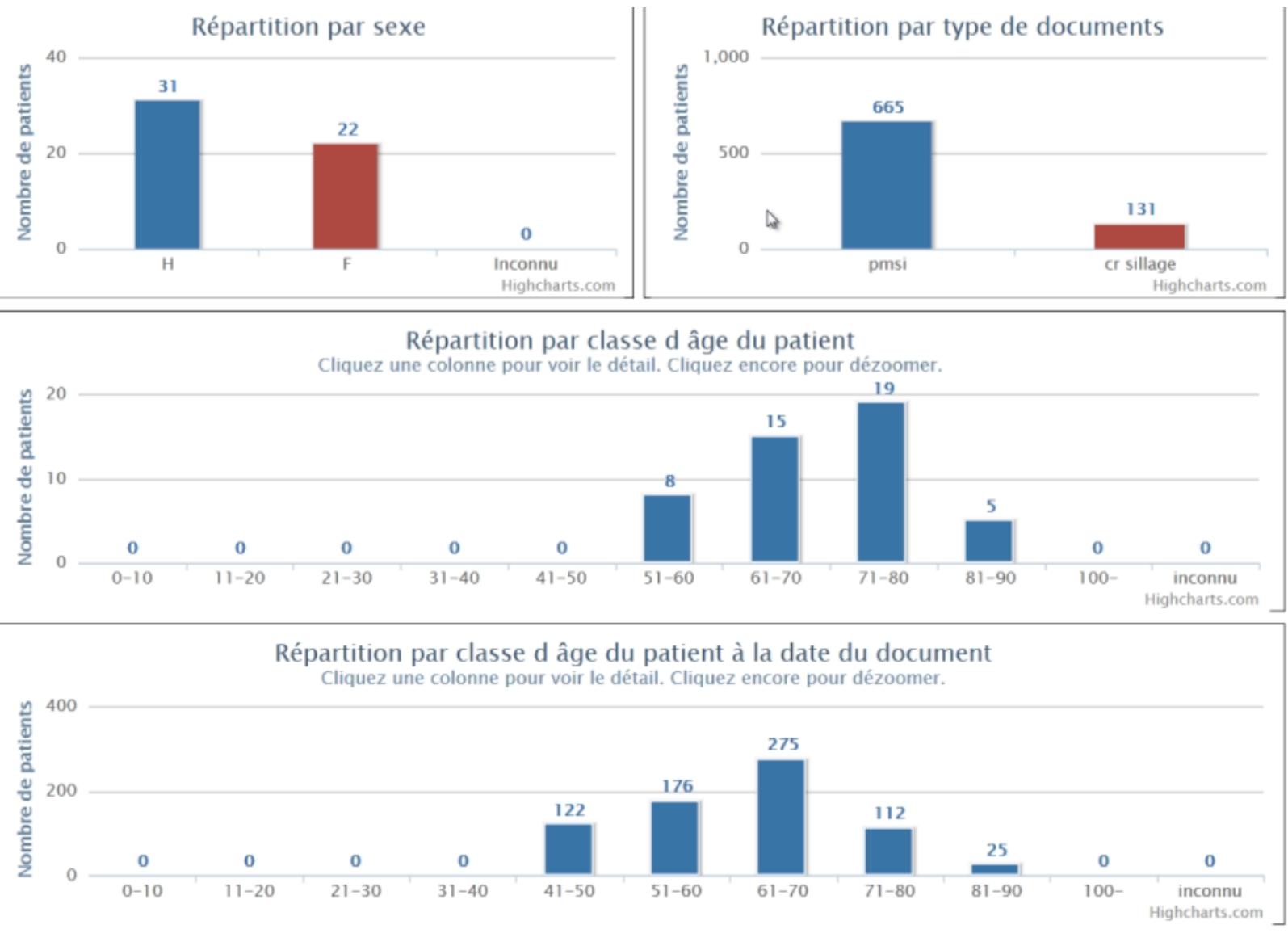
**Administration**

Requête	Statistiques	Map	Résultats	Export Excel
<p>Nombre de documents : 47      Nombre de patients : 46      Nombre de séjours : 46</p> <p><b>Signes / symptômes et diagnostics</b></p> <p>CRISE CONVULSIVE TONIQUE MALAISE      SPLENOMEGALIE CRANIOPHARYNGIOME POLYPHAGIE      NYSTAGMUS OPTOCINETIQUE CRISES SAI THROMBOPHILIE AFFECTION DU SYSTEME MUSCULAIRE      SOUFFRANCE FOETALE VOMISSEMENT RETINITE PIGMENTAIRE HANDICAP VISUEL      ASTHME ALLERGIQUE TROUBLES DU METABOLISME NEUROFIBROMATOSES      NEUROFIBROMATOSE BOUTON NEUROPATHIE AXONALE DIABÈTE HYPOTONIE CRISE      RYTHROBLASTOPENIE NEUROPATHIE SAI EPILEPSIE SAI NEUROPATHIE CATARACTE      EXCITABILITE ENCEPHALOPATHIES TOUX EPILEPSIE SURDITE SPASME DNAI      ICELE DESEQUILIBRE CECITE TOTALE CECITE AIT ALOPECIE NYSTAGMUS      RECIDIVE MYOCLONIE PETITE TAILLE HYPERLACTACIDEMIE      NOCTURNE BRONCHIOLITE GLIOME CICATRI      RDITE DE PERCEPTION DIABÈTE INSIPIDE SPASMES DYSTONIE ENCEPHALOPATHIE COMPLICATION      BRONCHIOLITE GLIOME SEMEIOLOGIE OPA1 SIDEROSE      RETARD DE CROISSANCE REGIME CETOGENE MALNF1 PARAPLEGIE SPASTIQUE ATTEINTE DU NERF OPTIQUE      DIABÈTE INSIPIDE SPASMES DYSTONIE ENCEPHALOPATHIE HYPER SURDOSAGE      CEPHALEES PALE MALADIES METABOLIQUES NEUROPATHIE OPTIQUE LEUCOMALacie PERVENTRICULAIRE      GLIOME DES VOIES OPTIQUES HEPATOSPLENOMEGALIE      SURDITE DE TRANSMISSION NEUROFIBROMATOSE DE TYPE 1 HYPOPITUITARISME      DELUXIEME GROSSESSE ASTHENIE TROUBLE DE LA VISION</p> <p>Gènes</p>				
<pre> graph LR     GE[Gene Expression] --&gt; SPG7[SPG7]     GE --&gt; ND5[ND5]     SPG7 --&gt; MITO[MITO]     ND5 --&gt; MITO     style GE fill:#0070C0,color:#fff     style SPG7 fill:#0070C0,color:#fff     style ND5 fill:#0070C0,color:#fff     style MITO fill:#0070C0,color:#fff   </pre>				



Requête	Statistiques	Map	Résultats	Export Excel																																																												
<p>cherche : Intégralité de 7 patients)</p> <p>es</p> <p>AAAA): <input type="text"/> <input type="text"/></p> <p>itation: <input type="text"/> <input type="text"/> : 30 90</p> <p>H <input type="checkbox"/> F <input type="checkbox"/></p> <p>les patients <input type="checkbox"/></p> <p>ext</p> <p>texte UF : <input type="text"/></p> <p>acel OR deprenyl OR <input type="checkbox"/></p> <p>es : <input type="checkbox"/></p> <p>texte UF : <input type="text"/></p> <p>es : <input type="checkbox"/></p> <p>e</p> <p><input type="checkbox"/> liste concepts - arbre</p> <p>P) <input type="checkbox"/> ts.org</p>	<p>Nombre de documents : 796 Nombre de patients : 53 Nombre de séjours : 473</p>		<p>Nb</p> <p>Nombre de documents/patients/séjours</p> <table border="1"> <caption>Data for Nombre de documents/patients/séjours</caption> <thead> <tr> <th>Année</th> <th>Documents</th> <th>Séjours</th> <th>Patients</th> </tr> </thead> <tbody> <tr><td>1998</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>1999</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>2000</td><td>10</td><td>10</td><td>10</td></tr> <tr><td>2001</td><td>10</td><td>10</td><td>10</td></tr> <tr><td>2002</td><td>15</td><td>15</td><td>15</td></tr> <tr><td>2003</td><td>20</td><td>20</td><td>20</td></tr> <tr><td>2004</td><td>35</td><td>35</td><td>10</td></tr> <tr><td>2005</td><td>40</td><td>40</td><td>10</td></tr> <tr><td>2006</td><td>50</td><td>45</td><td>20</td></tr> <tr><td>2007</td><td>75</td><td>60</td><td>30</td></tr> <tr><td>2008</td><td>120</td><td>70</td><td>10</td></tr> <tr><td>2009</td><td>360</td><td>140</td><td>30</td></tr> <tr><td>2010</td><td>60</td><td>50</td><td>10</td></tr> <tr><td>2011</td><td>50</td><td>40</td><td>10</td></tr> </tbody> </table>	Année	Documents	Séjours	Patients	1998	0	0	0	1999	0	0	0	2000	10	10	10	2001	10	10	10	2002	15	15	15	2003	20	20	20	2004	35	35	10	2005	40	40	10	2006	50	45	20	2007	75	60	30	2008	120	70	10	2009	360	140	30	2010	60	50	10	2011	50	40	10	<p>Highcharts.com</p>
Année	Documents	Séjours	Patients																																																													
1998	0	0	0																																																													
1999	0	0	0																																																													
2000	10	10	10																																																													
2001	10	10	10																																																													
2002	15	15	15																																																													
2003	20	20	20																																																													
2004	35	35	10																																																													
2005	40	40	10																																																													
2006	50	45	20																																																													
2007	75	60	30																																																													
2008	120	70	10																																																													
2009	360	140	30																																																													
2010	60	50	10																																																													
2011	50	40	10																																																													
			<p>Répartition des documents au cours du temps</p> <p>Click and drag in the plot area to zoom in</p>																																																													

**Nombre d'études**  
Nombre d'étude à partir des



Requête
Statistiques
Map
Résultats
Export Excel

## Patients

Voir le fichier 7368

**Recherche:**

**Pathologie** 53 (missing this field)

**Sexe** 22 F 31 M

**Vivant/Décédé** 1 décédé 52 vivant

53 Patients

33 résultats sur 53 ne peuvent pas être tracés.

Identifiant Patient : IPP  
 Nom : 29670,  
 Adresse : 29670,  
 Sexe : M  
 Vivant/Décédé : vivant  
 Né le : 69 ans

Utilisateur : Denis DELAMARRE - Projet : Parkinson - Données anonymisées - Deconnexion

cherche : Intégralité de 27 patients)	Requête	Statistiques	Map	Résultats	Export Excel																																																																																																																																																																																																																
<p><b>éees</b></p> <p>/AAAA): <input type="text"/> <input type="text"/></p> <p>ultation: <input type="text"/> <input type="text"/></p> <p>t: <input type="text"/> 30 <input type="text"/> 90 H <input type="checkbox"/> F <input type="checkbox"/></p> <p>les patients <input type="checkbox"/></p> <p><b>ext</b></p> <p>texte <input type="text"/> UF : <input type="text"/> racel OR deprenyl OR ina <input type="checkbox"/></p> <p>res : <input type="checkbox"/> e : <input type="checkbox"/></p> <p>texte <input type="text"/> UF : <input type="text"/> <input type="checkbox"/></p> <p>res : <input type="checkbox"/> e : <input type="checkbox"/></p> <p><b>re</b></p> <p><input type="checkbox"/> liste concepts - arbre</p> <p>AP) (P)</p>	<p>agrandir - reduire</p> <p><input checked="" type="radio"/> ET <input type="radio"/> OU</p> <p>Search: <input type="text"/></p> <table border="1"> <thead> <tr> <th colspan="2">Concepts</th> <th># Docs</th> </tr> </thead> <tbody> <tr> <td><input type="checkbox"/></td> <td>rasagiline or selegine or otracel or deprenyl or sinemet or modopar or l-dopa</td> <td>122</td> </tr> <tr> <td><input type="checkbox"/></td> <td>dyskinesie</td> <td>62</td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th>ipp</th> <th>nom</th> <th>prenom</th> <th>Age aujourd'hui</th> <th>S</th> <th>CP</th> <th>Actions</th> <th>Panier</th> <th>Doc</th> </tr> </thead> <tbody> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 120755740 13/11/2006 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 121767472 19/01/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 122046170 20/03/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 122046170 21/03/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 122046170 28/03/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 122046170 04/04/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 123121362 09/05/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 123789487 10/07/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 125017758 20/10/2007 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 127301832 18/04/2008 PMSI_DIAG pmsi</td> </tr> <tr> <td>IPP</td> <td>NOM_PATIENT</td> <td>PRENOM</td> <td>55 ans</td> <td>M</td> <td>56650</td> <td></td> <td></td> <td>Voir 131182293 4402 29/01/2009 Compte Rendu d'Hospitalisation cr sillage pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 131182293 03/02/2009 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 131182293 03/02/2009 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 131737467 17/03/2009 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 131737467 17/03/2009 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 135196915 4409 18/11/2009 Compte Rendu de Consultation cr sillage pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 137259889 05/03/2010 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 143759343 08/04/2011 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 147255727 24/11/2011 PMSI_DIAG pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 133876056 4409 17/08/2009 Compte Rendu de Consultation cr sillage pmsi</td> </tr> <tr> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>Voir 133876056 09/03/2010 PMSI_DIAG pmsi</td> </tr> </tbody> </table>						Concepts		# Docs	<input type="checkbox"/>	rasagiline or selegine or otracel or deprenyl or sinemet or modopar or l-dopa	122	<input type="checkbox"/>	dyskinesie	62	ipp	nom	prenom	Age aujourd'hui	S	CP	Actions	Panier	Doc									Voir 120755740 13/11/2006 PMSI_DIAG pmsi									Voir 121767472 19/01/2007 PMSI_DIAG pmsi									Voir 122046170 20/03/2007 PMSI_DIAG pmsi									Voir 122046170 21/03/2007 PMSI_DIAG pmsi									Voir 122046170 28/03/2007 PMSI_DIAG pmsi									Voir 122046170 04/04/2007 PMSI_DIAG pmsi									Voir 123121362 09/05/2007 PMSI_DIAG pmsi									Voir 123789487 10/07/2007 PMSI_DIAG pmsi									Voir 125017758 20/10/2007 PMSI_DIAG pmsi									Voir 127301832 18/04/2008 PMSI_DIAG pmsi	IPP	NOM_PATIENT	PRENOM	55 ans	M	56650			Voir 131182293 4402 29/01/2009 Compte Rendu d'Hospitalisation cr sillage pmsi									Voir 131182293 03/02/2009 PMSI_DIAG pmsi									Voir 131182293 03/02/2009 PMSI_DIAG pmsi									Voir 131737467 17/03/2009 PMSI_DIAG pmsi									Voir 131737467 17/03/2009 PMSI_DIAG pmsi									Voir 135196915 4409 18/11/2009 Compte Rendu de Consultation cr sillage pmsi									Voir 137259889 05/03/2010 PMSI_DIAG pmsi									Voir 143759343 08/04/2011 PMSI_DIAG pmsi									Voir 147255727 24/11/2011 PMSI_DIAG pmsi									Voir 133876056 4409 17/08/2009 Compte Rendu de Consultation cr sillage pmsi									Voir 133876056 09/03/2010 PMSI_DIAG pmsi
	Concepts		# Docs																																																																																																																																																																																																																		
	<input type="checkbox"/>	rasagiline or selegine or otracel or deprenyl or sinemet or modopar or l-dopa	122																																																																																																																																																																																																																		
	<input type="checkbox"/>	dyskinesie	62																																																																																																																																																																																																																		
	ipp	nom	prenom	Age aujourd'hui	S	CP	Actions	Panier	Doc																																																																																																																																																																																																												
									Voir 120755740 13/11/2006 PMSI_DIAG pmsi																																																																																																																																																																																																												
									Voir 121767472 19/01/2007 PMSI_DIAG pmsi																																																																																																																																																																																																												
									Voir 122046170 20/03/2007 PMSI_DIAG pmsi																																																																																																																																																																																																												
									Voir 122046170 21/03/2007 PMSI_DIAG pmsi																																																																																																																																																																																																												
									Voir 122046170 28/03/2007 PMSI_DIAG pmsi																																																																																																																																																																																																												
								Voir 122046170 04/04/2007 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 123121362 09/05/2007 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 123789487 10/07/2007 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 125017758 20/10/2007 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 127301832 18/04/2008 PMSI_DIAG pmsi																																																																																																																																																																																																													
IPP	NOM_PATIENT	PRENOM	55 ans	M	56650			Voir 131182293 4402 29/01/2009 Compte Rendu d'Hospitalisation cr sillage pmsi																																																																																																																																																																																																													
								Voir 131182293 03/02/2009 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 131182293 03/02/2009 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 131737467 17/03/2009 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 131737467 17/03/2009 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 135196915 4409 18/11/2009 Compte Rendu de Consultation cr sillage pmsi																																																																																																																																																																																																													
								Voir 137259889 05/03/2010 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 143759343 08/04/2011 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 147255727 24/11/2011 PMSI_DIAG pmsi																																																																																																																																																																																																													
								Voir 133876056 4409 17/08/2009 Compte Rendu de Consultation cr sillage pmsi																																																																																																																																																																																																													
								Voir 133876056 09/03/2010 PMSI_DIAG pmsi																																																																																																																																																																																																													

## RECHERCHE DE L'OCCURRENCE DU TERME **GRIPPE** dans 3 millions de documents

OM] [NOM\_PATIENT], Agée de 71 ans à la date du document, code postal 35200

PP] , IEP : [IEP]

02, document daté du 29/03/2009

Rendu d'Hospitalisation



vasculaires  
isable Médical  
r Thomas RONZIERE  
de Santé  
atricia VOISIN  
ariat [TEL] (84293)

Investigation  
ologique O. Sabouraud  
isable Médical  
ieur Gilles EDAN  
de Santé  
éronique SCHNEIDER

van Gogh  
isable Médical  
r Arnaud BIRABEN  
de Santé  
atricia VOISIN  
ariat [TEL] (84162)

ILLATION

de la S.E.P  
isable Médical  
r Emmanuelle LE

tre coordinatrice  
ène BRUNET  
ariat [TEL] (83709)

ologie  
r Arnaud BIRABEN  
r Anca PASNICU  
ariat [TEL] (84162)

ments Anormaux

dysarthrie. Au niveau cognitif, on n'observe pas de cauchemar ni d'hallucination.

L'examen clinique met en évidence une marche fluide avec un demi-tour normal et un ballant des bras conservé. Il existe un tremblement de repos du membre supérieur et du membre inférieur. La rigidité est cotée à 1F2 au niveau du membre supérieur droit et OF0 au niveau du membre supérieur gauche.

Devant la recrudescence de la symptomatologie hémicorporelle droite, les paramètres de stimulation sont augmentés au niveau du GPI gauche. Les paramètres de stimulation de sortie sont donc :

- GPI droit : 0-, 2.4 v, 90 µs, 130 Hz
- GPI gauche : plot 0-, 2.9 v, 90 µs, 145 Hz

Il est à noter cependant que la recrudescence de la symptomatologie est contemporaine d'une erreur de traitement avec l'arrêt du TRIVASTAL pendant 5 jours.

Le bilan à 6 mois réalisé :

- la consultation psychiatrique met en évidence une labilité thymique et des ruminations névrotiques. La patiente est orientée vers le CMP de son secteur.
- Le bilan neuropsychologique retrouve une Mattis à 137/144. Il existe une stabilisation de l'efficience cognitive globale et des performances aux épreuves évaluant les fonctions exécutives. L'importante impulsivité constatée lors du dernier bilan gêne la patiente dans les tâches de graphisme.

Le Test à la Dopa :

- **LDOPA** OFF STIM OFF 27
- **LDOPA** ON STIM ON 7.5
- **LDOPA** ON STIM OFF 16.5
- **LDOPA** ON STIM ON 3
- UPDRS I est à 4
- UPDRS II est à 5 en ON et 17 en OFF
- UPDRS IV est à 6
- Le stade de Hoehn et Yahr est à 1 en ON et 3 en OFF
- L'échelle de Schwab et England est à 90 % en ON et 60 % en OFF

AU TOTAL :

- bilan à 6 mois d'une stimulation bi-pallidale dans le cadre d'une **maladie de Parkinson**
- augmentation des paramètres de stimulation du GPI gauche devant une récidive du tremblement et de l'akinésie hémicorporelle droite.
- La patiente sera convoquée pour son bilan à 12 mois

# Ajout dans un panier

Utilisateur : Nicolas Garcelon - Projet : Présentation démo - Données anonymisées - [Déconnexion](#)

### Formulaire de recherche : Intégralité de l'entrepôt (14660 patients)

- [+ Requêtes enregistrées](#)
- [+ Critères généraux](#)
- Recherche en Full-text

Inclure

atrophie optique

Inclure periphrases négatives :

Enrichissement sémantique :

[ajouter requête](#)

### + Recherche structurée

Intersection des critères sur  le patient  le séjour

Ne pas afficher les patients déjà exclus du panier en cours  oui  non

[Rechercher](#) [Enregistrer la requête](#) [Vider le formulaire](#)

### Création des paniers

- Panier en cours

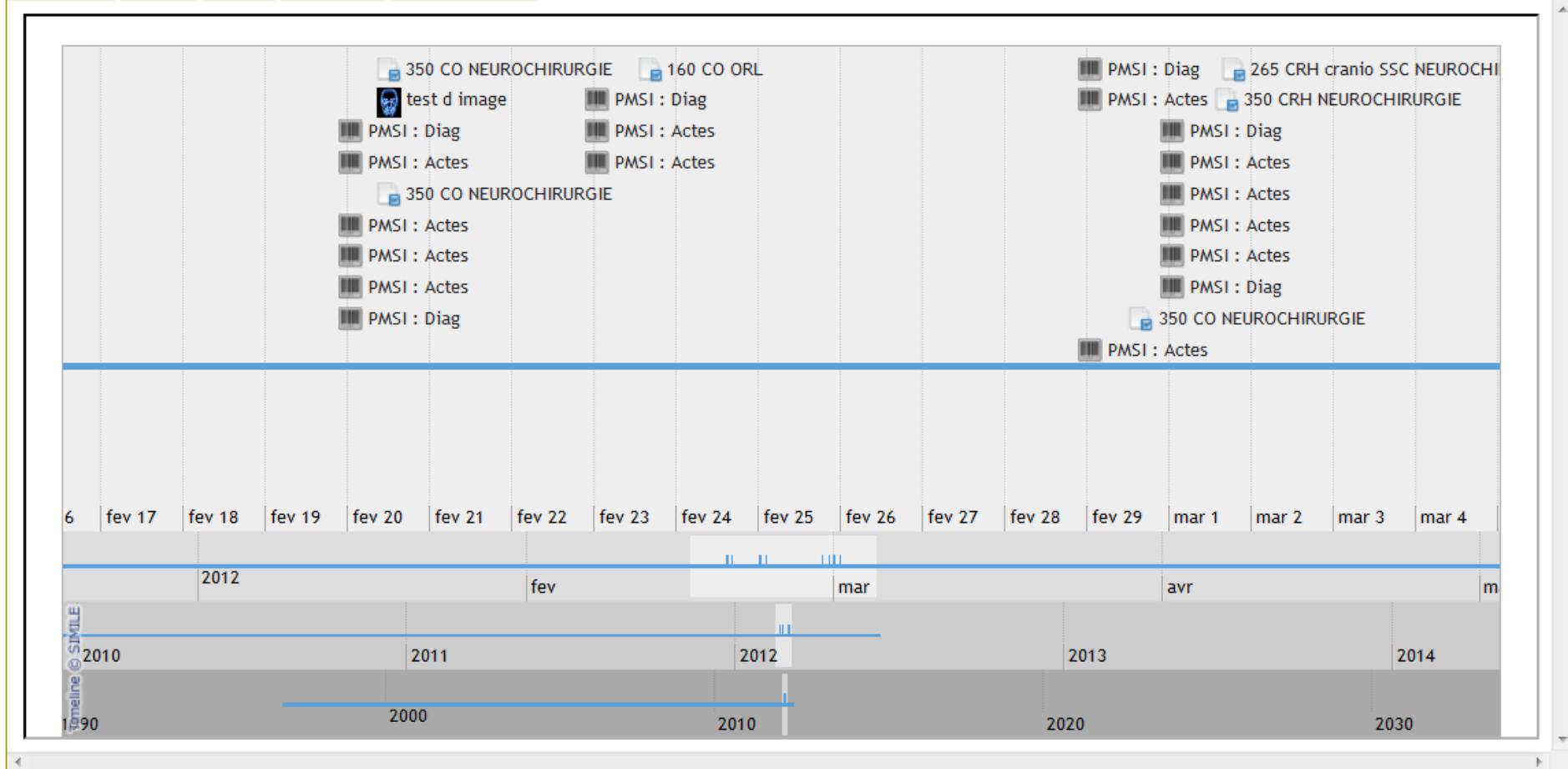
Panier Sauvé : atrophie optique

X
←

Concepts	# Docs	Search:	Concepts	# Docs	Search:
<input type="checkbox"/> atrophie optique	47	Search:	<input type="checkbox"/> entrant	5	Search:
<input type="checkbox"/> atrophie optique, non precisee	47		<input type="checkbox"/> sentiments	5	
<input checked="" type="checkbox"/> encephalopathie	4		<input checked="" type="checkbox"/> encehalopathie	4	

ipp	nom	prenom	Age aujourd'hui	S	CP	Actions	Panier	Doc
NIP	NOM_PATIENT	PRENOM	21 ans	F	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	21 ans	F	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	23 ans	M	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	41 ans	F	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	58 ans	M	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	17 ans	F	-		<a href="#">Voir</a>	06/02/2013 Base IRM Base IRM
NIP	NOM_PATIENT	PRENOM	9 ans	F	-		<a href="#">Voir</a>	301 17/11/2011 301 CION NEUROLOGIE PEDIATRIQUE SUSIRM
NIP	NOM_PATIENT	PRENOM	arrete avant l'age de 6 mois.	-	-		<a href="#">Voir</a>	06/02/2013 Base IRM
NIP	NOM_PATIENT	UF : 301, document daté du 17/11/2011 301 CION NEUROLOGIE PEDIATRIQUE		-	-			IRM
NIP	NOM_PATIENT	Le motif de consultation est la mise en évidence d'une <b>atrophie optique</b> bilatérale en juin 2011, avec une acuité visuelle mesure à 2/10 de façon bilatérale.		-	-			IRM
NIP	NOM_PATIENT	Une IRM cerebrale a été réalisée le 24 octobre dernier, et je propose de la relire en Staff de		-	-			IRM



The screenshot shows a medical software interface with a timeline at the bottom and a detailed view of a diagnosis record in the center.

**Timeline:** The timeline at the bottom spans from 1990 to 2030, with major ticks for 1990, 2000, 2010, 2020, 2030, and specific dates like fev 19, fev 20, fev 21, fev 22, fev 29, mar 1, mar 2, mar 3, mar 4, mar 5, mar 6, avr, and mai.

**Left Panel:** A vertical list of medical records and documents:

- 350 CO NEUROCHIRURGIE
- test d image
- PMSI : Diag
- PMSI : Actes
- 350 CO NEUROCHIRURGIE
- PMSI : Actes
- PMSI : Actes
- PMSI : Actes
- PMSI : Diag

**Center Panel (Detailed View):**

**PMSI : Diag**

Afficher le texte dans une nouvelle fenêtre

Code principal : Z877 - antecedents personnels de malformations congenitales et anomalies chromosomiques

Codes secondaires :

- I899 - atteinte non infectieuse des vaisseaux et des ganglions lymphatiques, sans precision
- Y139 - intoxication par d'autres substances pharmacologiques agissant sur le système nerveux autonome et exposition à ces produits intention non déterminée | lieu, sans précision
- L560 - réaction phototoxique à un médicament
- V315 - occupant d'un véhicule à moteur à trois roues blesse dans une collision avec un cycle | conducteur blesse dans un accident de la

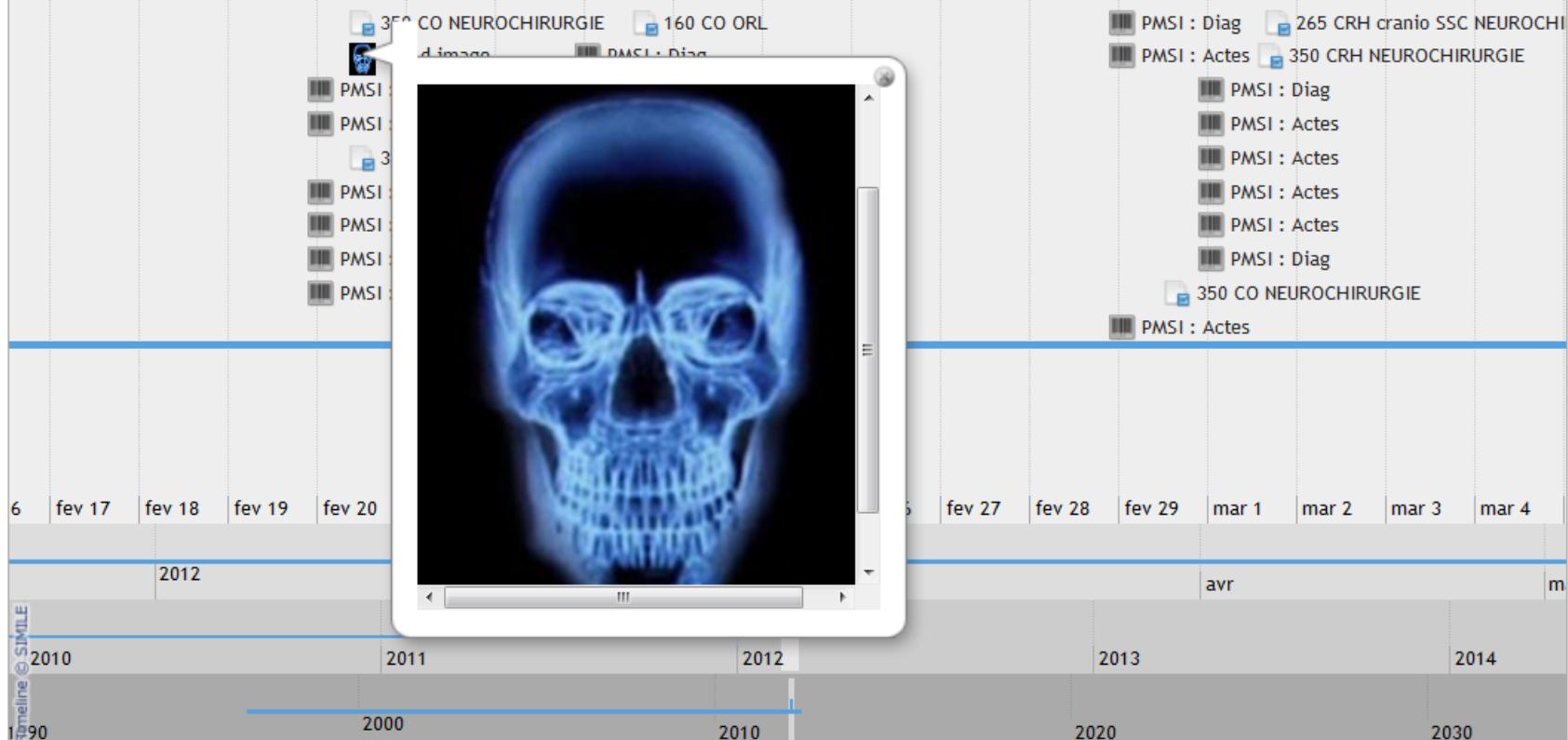
**Right Panel:** A vertical list of medical records:

- 160 CO ORL
- PMSI : Diag
- PMSI : Actes
- 265 CRH crano SSC NEUROCHIR PEDIATRIQUE
- 350 CRH NEUROCHIRURGIE
- PMSI : Diag
- PMSI : Actes
- PMSI : Actes
- PMSI : Actes
- PMSI : Actes
- PMSI : Diag

**Bottom Panel:** A vertical list of medical records:

- 350 CO NEUROCHIRURGIE
- PMSI : Actes

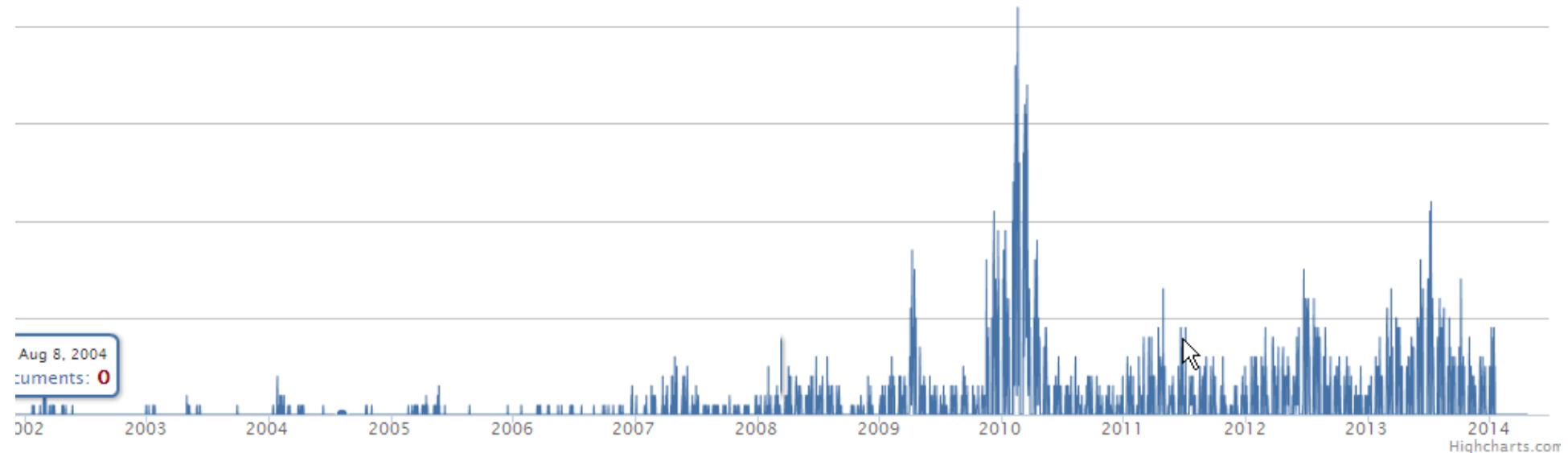
Documents Histoire Paniers Recherche Journal d'accès



## RECHERCHE DE L'OCCURRENCE DU TERME **GRIPPE** dans 10 millions de documents

Répartition des documents au cours du temps

Click and drag in the plot area to zoom in



Next step :  
Querying several hospitals

→ CDW Network

# next step

- Most of studies are multicentric
  - Epidemiology, clinical research, pharmacovigilance
- Need to query different hospitals in the same time
- much more complex
  - heterogeneity of data and CDW
  - Cross language barriers
  - Security
  - Business model (Pharma Compagnies)
- 3 examples :
  - EHR4CR (Clinical Research)
  - DEBUG IT (Biosurveillance)

## Electronic Health Record for Clinical Research

Feasibility

Screening

e-CRF

Pharmaco  
vigilance



## Electronic Health Record for Clinical Research

Etude  
Faisabilité

Recrutement

E-CRF

Pharamaco  
vigilance



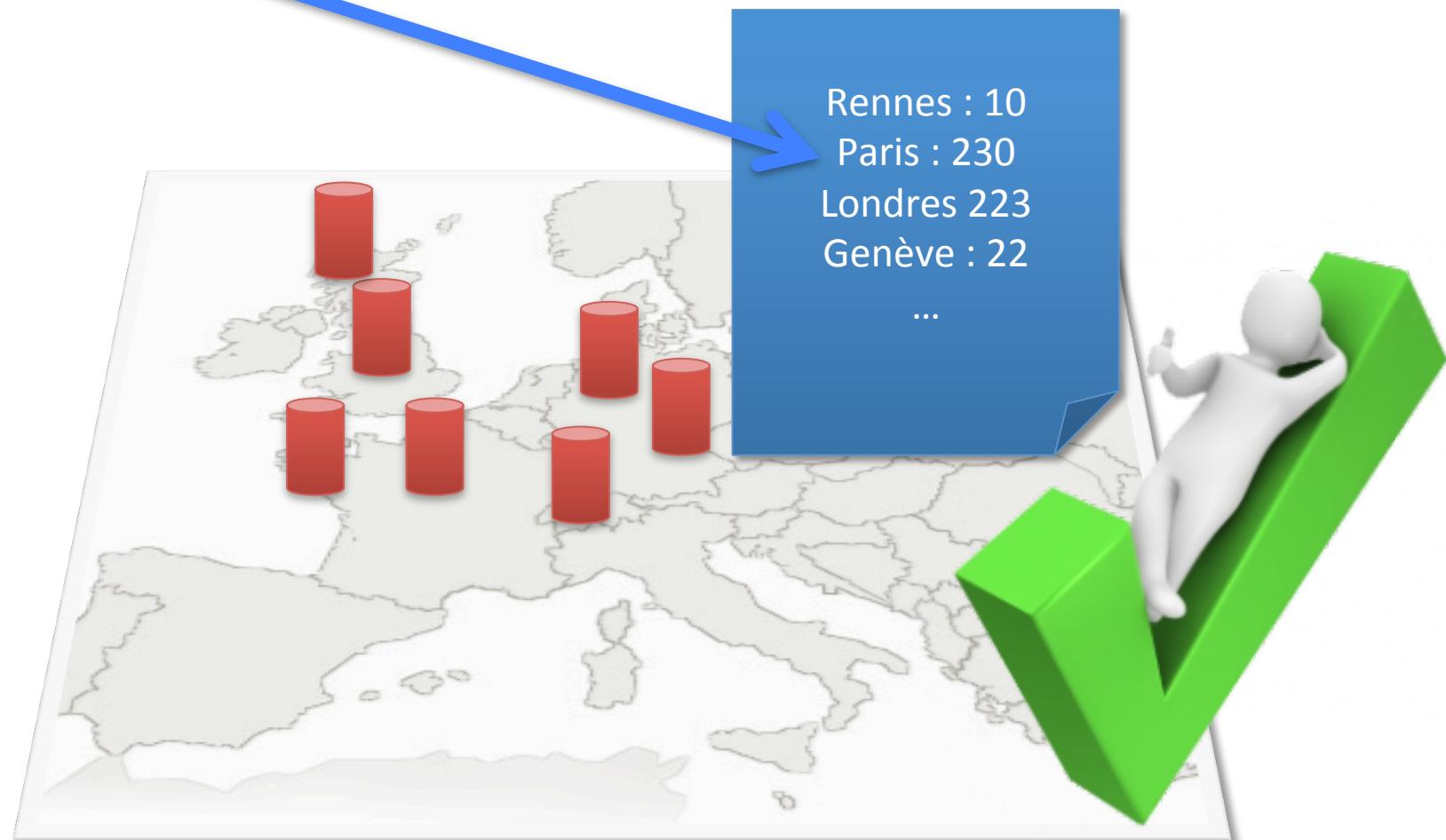
## Electronic Health Record for Clinical Research

Etude  
Faisabilité

Recrutement

E-CRF

Pharamaco  
vigilance



**Terminology**

Browse Search

Expand the terminology to search a concept.  
The draggable elements are colored with blue.

[Collapse all](#)

- [-] EHR4CR terminology
  - [+] Demographic history detail [CAT]
  - [+] Final diagnosis (discharge)[VS]
  - [+] Result, lab.- general (observable entity)[VS]
  - [+] Morphologically altered structure[VS]
  - [+] Drug or medicament (substance)[VS]
  - [-] Procedure [VS]
    - [-] Procedure
      - [+] Procedure by method
      - [-] **Procedure by site**
        - [+] Introduction of substance by body site
        - [+] Procedure on trunk
        - [+] Destructive procedure by anatomic site
        - [+] Procedure on head AND/OR neck
        - [+] Radiotherapy by body site
        - [+] Imaging by body site
        - [+] Surgical repair procedure by body site
        - [+] Procedure on soft tissue
        - [+] Procedure on body system
        - [+] Procedure on organ
        - [+] Procedure on extremity
        - [+] Procedure categorized by device involved
        - [+] Procedure by intent
    - [+] Observable entity [VS]
    - [+] Clinical Finding [VS]
    - [+] deceased

Constraints : Age Before now X Before Y Occurrence

Male

AND

More than 17 Years Old

AND

{ Malignant neoplasms of ill-defined, }

{ At Most 3 Month Before }

{ Malignant neoplasm of prostate }

AND

Enlarged lymph nodes

AND NOT

{ Cryotherapy }

{ At Most 6 Month Before now }

Add a new criterion

Delete All Save the query Run

Options Query Run Status and Results Query in Eclectic Format

**Label :** Malignant neoplasm of prostate  
**Category :** diagnosis  
**Code :** ICD-10 : C61

Please specify if you want the lastest or the first entry in the patient EHR :

[Go Back to the Query Builder](#)**Query : acute myocardial****QUERY STATUS: [COMPLETE]**

Query was distributed at: 04/15/2013 09:16:31

Current statistics: 2 endpoints expected and 2 completed

Expected endpoints: APHP, U936

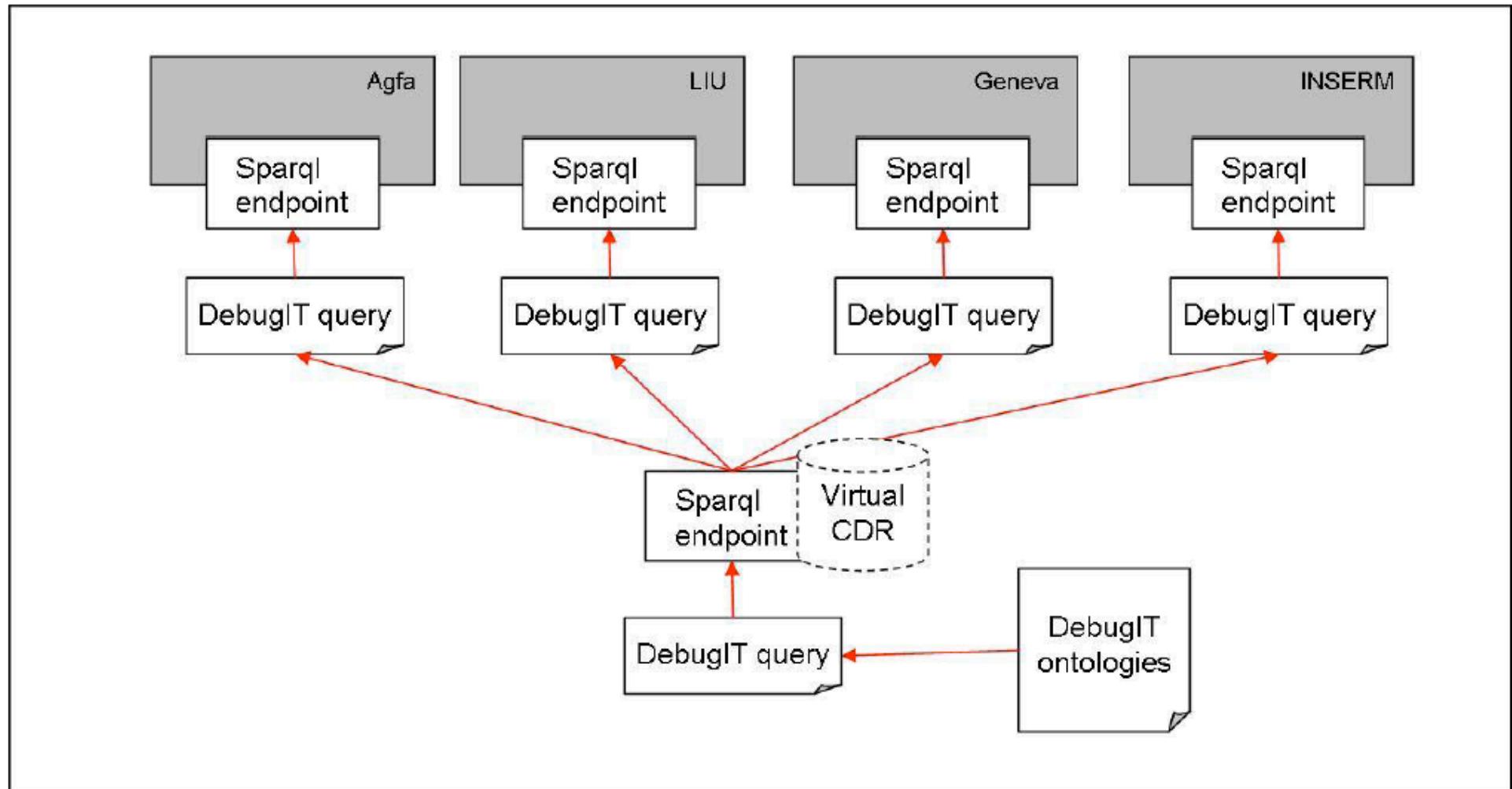
Finished endpoints: APHP, U936

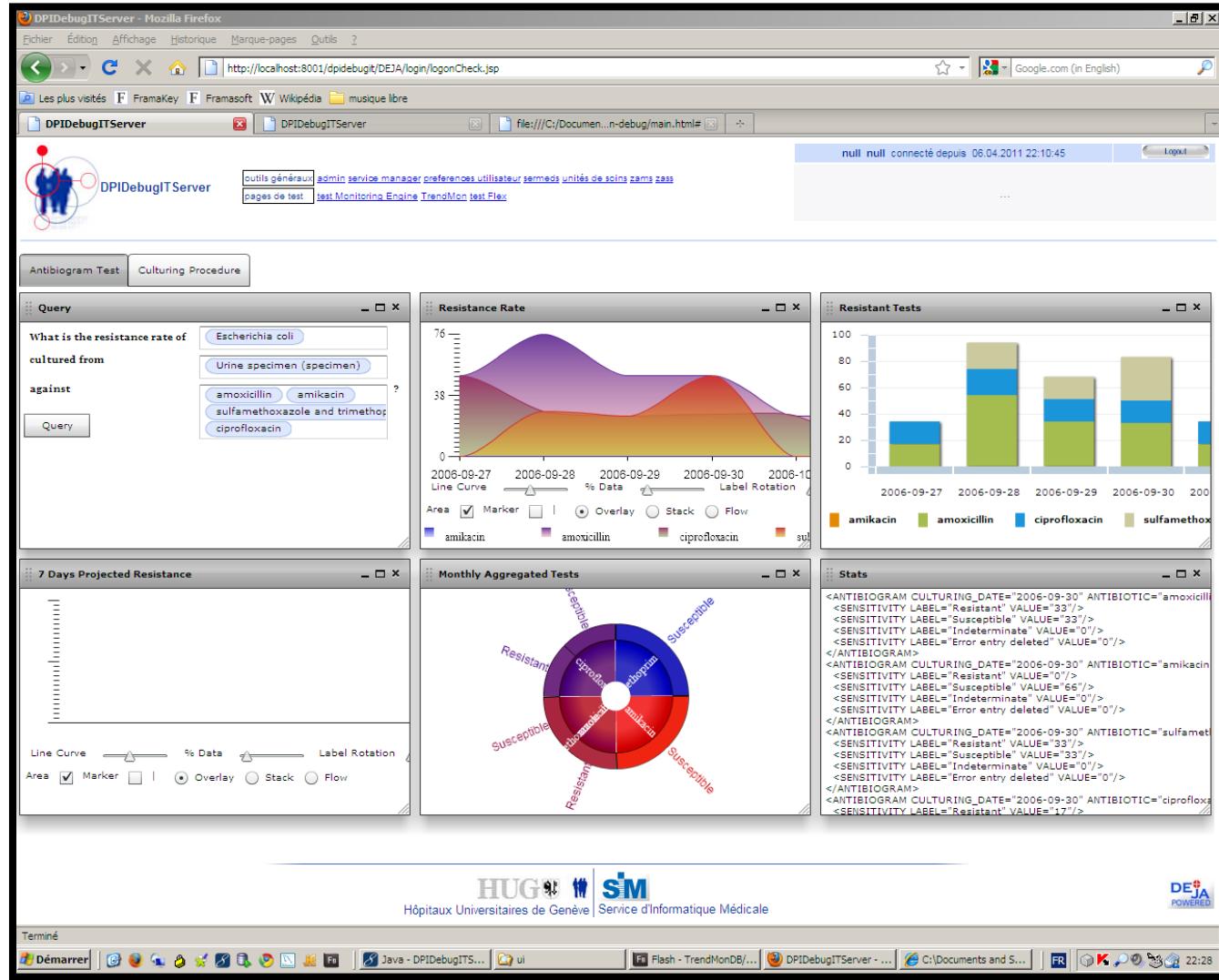
[See detailed status...](#)[See Results](#)**Result :**[◊ Aggregated \(705\)](#)[◊ APHP \(38\)](#)[◊ U936 \(667\)](#)[◊ 667 patients match all criteria](#)[◊ 667 patients match: last diagnosis\(\[ICD-10:I21,"Acute myocardial infarction"\]\)](#)

Age	Male	Female	Undefined	Total
0 to 9	0	0	0	0
10 to 19	0	0	0	0
20 to 29	1	0	0	1
30 to 39	12	2	0	14
40 to 49	50	15	0	65
50 to 59	87	17	0	104
60 to 69	103	28	0	131
70 to 79	107	46	0	153
80 +	102	97	0	199

# Biosurveillance : DEBUGIT project

- FP7
- International survey of bio-resistance of bacteria
- Integration of antibiograms and lab test

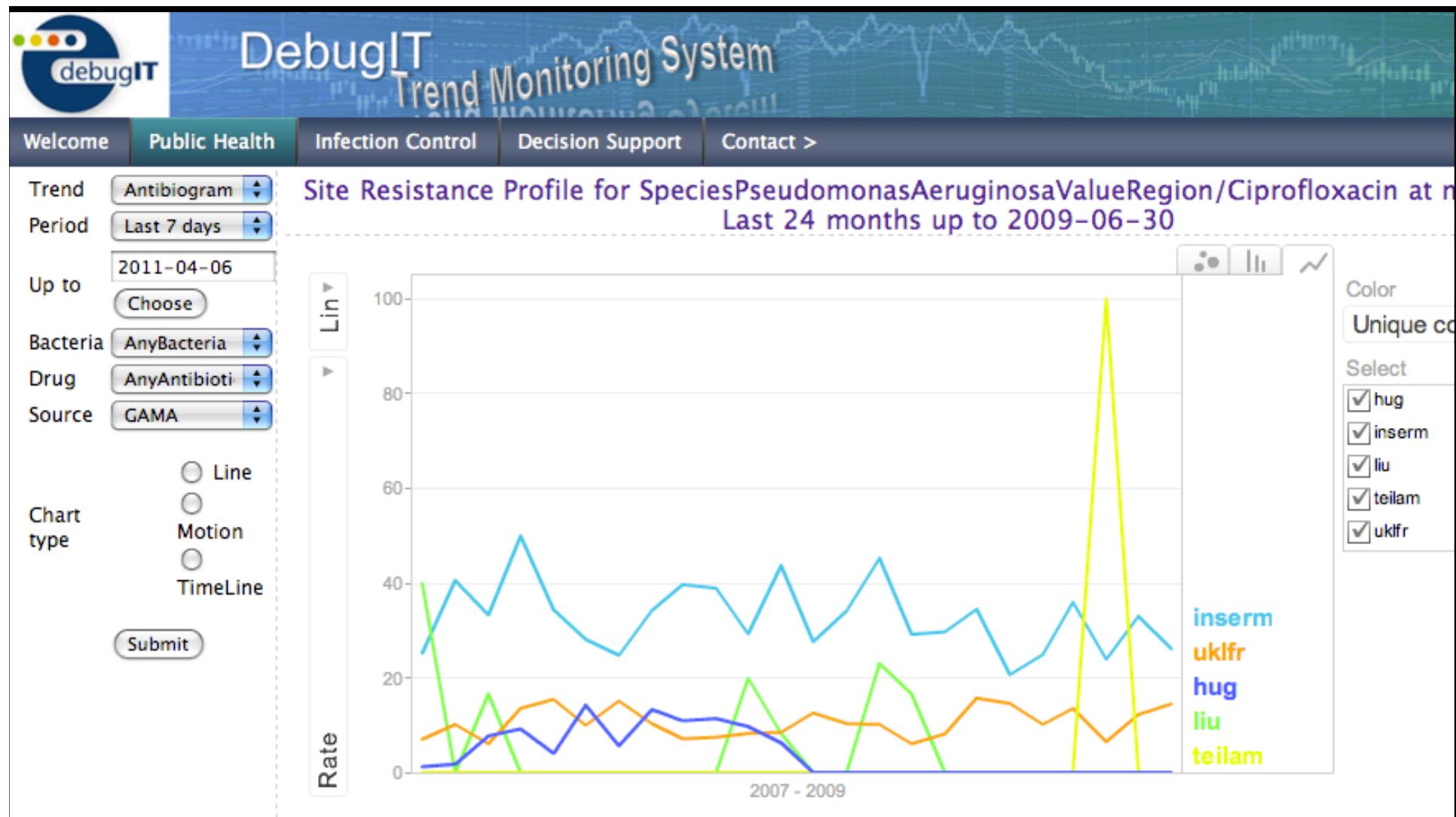




1. Clinical question
2. Resistance rate evolution
3. Antibiogram tests performed (to give confidence in the rate)
4. Forecasting
5. Monthly aggregated
6. Statistics (mean, standard deviation, etc.)

What is the **evolution of the resistance** of *Pseudomonas Aeruginosa* to ciprofloxacin at HUG, INSERM, LIU, TEILAM and UKLFR from 01.06.2007 to 30.06.2009 ?

(use of historic data because the CDRs are not on real time production systems yet)



# Current issues

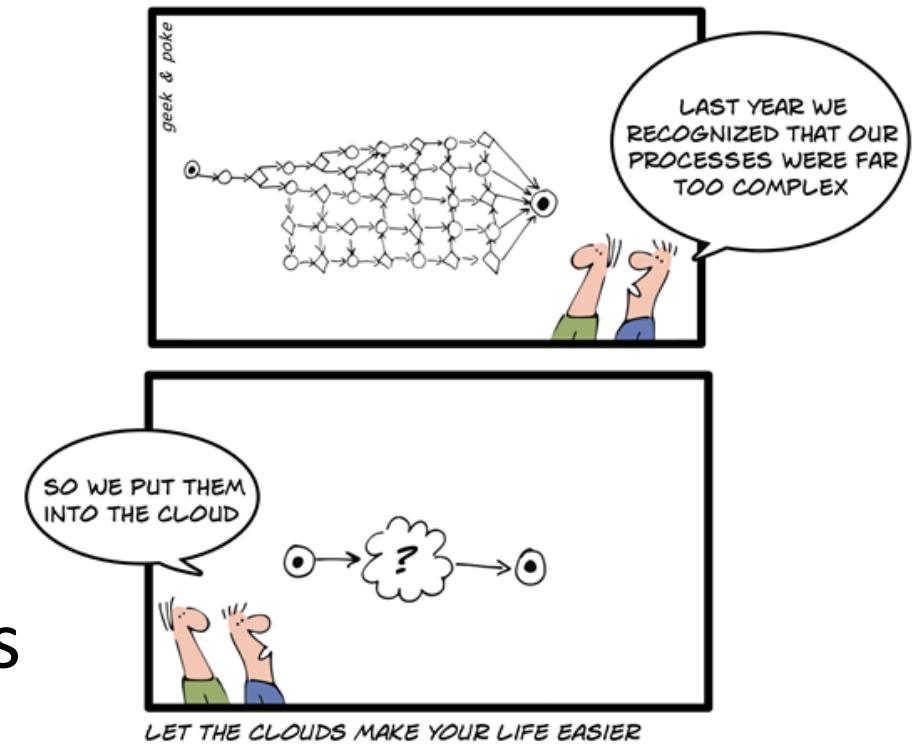
- Semantic interoperability
- Data Quality : Garbage In → Garbage Out
  - «technical » data ☺
  - Interpreted data ☹
  - Missing data ☹
- Semantic integration
  - Ontology for a large domain
    - Not sufficiently rich or formal
    - Hard to maintain

# Current issues

- Reinvent new methods for querying and mining medical data
  - Mining a CDW is not simple as using Google but much more complex
- Scalability and efficiency
  - Optimizing the Architecture and the Query method
  - A query execution should not last more than a coffee break
- Data protection – Traceability+++

# Next steps

- Semantic web
  - Triple Store CDW
  - SPARQL Endpoints
  - CLOUD computing
- NLP and indexing
- Integrating new sources



# Acknowledgment

- Financeurs:

- ANR TECSAN
- IMI FP7
- CRITT Santé Bretagne
- Cengeps



- Collaborateurs :

- Nicolas Garcelon
- Denis Delammare
- Pascal Van Hill
- Olivier Dameron
- Wassim Jouini
- Annabel Bourde
- Régis Duvaufier
- André Happe
- Isabelle Stévant