A search and analysis platform for scientific publications dedicated to Pharmacovigilance

Mathilde Thébault
Senior Text-Mining Consultant – Life Sciences
This project was founded by ANR (National Research Agency)
- ANR-07-TECSAN-026
- 2 years project (from December 2007 to January 2010)

Partners

**Research Institutes**
- Université Jean Monnet
- LM&Bio
- Loria
- Langues O'-
- SPIM
- Inserm

**Hospitals**
- HEGP
- CHU Saint-Étienne

**SME**
- Temis
- Mondeca
Pharmacovigilance: Key Points

- Mission of Regional Centers of Pharmacovigilance (CRPV)
  - To monitor, Evaluate and Prevent drug risks (known or potential) and to promote the right use of drugs
  - The main source of information comes from spontaneous notifications of adverse events
VigiTermes: Objectives

❖ Objectives:
  • To create a semantic portal that gives an access to enriched PubMed abstracts
  • To facilitate the access to similar cases and offer support to the documentation process

❖ Creation of an user interface
  • To send a query with the couple Adverse Event / Active Substance
  • To retrieve corresponding enriched PubMed abstracts
VigiTermes platform

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table

1. Adverse event + Drug + Filters

2. Query: converted and enriched

VigiPubMed Web Service

3. PMID List

4. PMID List

5. PubMed Abstract content

6. Annotations

7. Annotated and enriched abstracts

Pharmacovigilant

PubMed

ITM

Ontologies

Terminologies Thesaurus

Knowledge Base

Content Augmentation Manager

LUXID

Annotation Factory

542x34 Copyright © 2010 TEMIS - All Rights Reserved - Slide 5
Retrieval of Pubmed abstracts

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table

1. Adverse event + Drug + Filters

2. Query: converted and enriched

VigiPubMed Web Service

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

Pharmacovigilant

PubMed

ITM

- Ontologies
- Terminologies Thesaurus
- Knowledge Base

Content Augmentation Manager

LUXID

Annotation Factory

Copyright © 2010 TEMIS - All Rights Reserved - Slide 6
VigiPubMed: Queries

- **Active Substance (AS)**
  - Based on International Nonproprietary Name (INN or DCI for France)
  - Or based on trade name

- **Adverse Event (AE)**
  - 4 ways to send a query
    1. Direct query to PubMed
    2. Query with Normalization
    3. Query with Normalization and MeSH extension
    4. Query with OntoEIM extension

OntoEIM = Ontologie des Évènements Iatrogènes Médicamenteux

=> Ontology for Adverse Drug Event
VigiPubMed overview

- Different methods are available to send a query
- They are based on the couple Active Substance/Adverse Event

http://www.med.univ-rennes1.fr/vigitermes/ref/vigipubmed.php

Adverse Event: Normalized query

Adverse Event
Term mapped to MedDRA (coding of signs and symptoms)

Terms Normalization
Web Service UMLS: findCUIByNormString

Identification of a closed MeSH term:
Table RestrictToMesh from UMLS

PubMed: Translated Term [MeSH Terms] or raw term

Raw Term = non refined query

Example:
"cardiac failure" → CUI "C0018801" → "Heart Failure"

PubMed:
Heart Failure[MeSH Terms]
Or cardiac failure

CUI = Concept Unique Identifier
UMLS = Unified Medical Language System
AE: Query extension with OntoEIM

- **OntoEIM: Adverse Events ontology**
  - Sources
    - WHO-ART: 5 798 classes
    - MedDRA: 37 892 classes
    - SNOMED-CT: 14342 classes (4.6 %)
  - Alignment, concepts network
    - UMLS Metathesaurus

- **PharmARTS: Adverse Events ontology**
  - Tool for queries and navigation

- **Web Services – VigiTermes**
  - Terms association (search of similar cases)

WHO-ART = World Health Organization - Adverse Reactions Terminology
MedDRA = Medical Dictionary for Regulatory Activities
SNOMED-CT = Systematized Nomenclature of Medicine - Clinical Terms
AE: Query extension with OntoEIM

Example: "cardiac failure"

Term mapped to MedDRA (coding of signs and symptoms)

PharmARTS / OntoEIM Web Service (MedDRA to MedDRA)

Cardiac_failure,
Cardiac_failure_NOS,
Cardiac_failure_aggravated,
Cardiac_function_failure,
Cardiac_insufficiency,
Decompensation_cardiac,
Decompensation_myocardial,
Failure_heart, Heart_failure,
Heart_failure_NOS,
Heart_failure_unspecified,
Heart_insufficiency,
Insufficiency_cardiac,
Myocardial_contraction_decreased,
Myocardial Decompensation,
Heart_failure_NYHA_class_I,
Heart_failure_NYHA_class_II,
Heart_failure_NYHA_class_III,
Heart_failure_NYHA_class_IV,
Decompensated_heart_failure,
Recurrent_cardiac_decompensation
Filters used to complete the query

- Various filters can enrich the query: age-group, gender, organism types, publication time, publication types, ...
- These filters are described into PubMed in the « Advanced Search » mode
- They help to target the search

Example:

Retrieval of PubMed abstracts

Query: "carbamazepine + alopecia" is converted to:

(carbamazepine) AND (alopecia [MeSH Terms] OR alopecia) AND 1900[DP]:2009[DP] AND (adverse effects[SH])

Results format (XML) for VigiTermes workflow

```
<projetvigitermes.queryPubmed.Results_TEMIS_data>
  <mvrVersion>V1</mvrVersion>
  <mvrDate>2009-12-13 12:00:37.116 CET</mvrDate>
  <mvrAdverseDrugReaction>alopecia</mvrAdverseDrugReaction>
  <mvrActiveIngredient>carbamazepine</mvrActiveIngredient>
  <mvrQuery>"carbamazepine"[MeSH Terms] OR "carbamazepine"[All Fields]) AND '(("alopecia"[MeSH Terms] OR "alopecia"[All Fields]) AND 1900[PDAT] : 2009[PDAT] AND "adverse effects"[Subheading])</mvrQuery>
  <mvrPmids>18779906;15293764;12124512;10798824;9343147;8899137;7901970;1486112;3107169;4085351;4265242</mvrPmids>
</projetvigitermes.queryPubmed.Results_TEMIS_data>
```
Retrieval of PubMed abstracts

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table

1. Adverse event + Drug + Filters

2. Query: converted and enriched

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

Pharmacovigilant

VigiPubMed Web Service

PubMed

ITM

Ontologies

Terminologies Thesaurus

Knowledge Base

Content Augmentation Manager

LUXID

Annotation Factory

Content Augmentation Manager

Copyright © 2010 TEMIS - All Rights Reserved - Slide 14
**Query:** "carbamazepine + alopecia" :

<table>
<thead>
<tr>
<th>PublicationType</th>
<th>Case Reports, English Abstract, Journal Article, Review,</th>
</tr>
</thead>
</table>
| MeshHeading     | DescriptorName (MajortopicYN) 
                  | QualifierName (MajortopicYN) |
| Epilepsy, Temporal Lobe (N) | drug therapy (N) |
| Humans (N)      | Male (N) |
| Middle Aged (N) | Alopecia (N) |
| Carbamazepine (N) | adverse effects (Y) |
| Anticonvulsants (N) | chemically induced (Y) |
| CHEMICAL        | NameOfSubstance |
|                 | (RegistryNumber) |
|                 | Anticonvulsants (O) |
|                 | Carbamazepine (298-46-4) |

**RESUME:**

We report a 52 year-old man presenting with an acute considerable hair loss induced by carbamazepine (CBZ). The remarkable scalp hair loss started within a week after CBZ administration. There was no evidence of dermatitis or allergic reaction, or other cause for the hair loss. The serum concentration of CBZ was 8.6 microg/ml (therapeutic range 8-12 microg/ml). CBZ was discontinued, and the hair loss stopped within several days with new hair growth. Medication-induced hair loss is an occasional adverse effect of many drugs used for neuropsychological diseases. CBZ also induces hair loss and its frequency was reported below 2%. Only a limited number of detailed case reports describing CBZ-induced hair loss were available, and we found these cases could divide into two groups with regard to a delay in starting hair loss after administration of CBZ. In one group, the hair loss started within a week suggesting anagen effluvium and in another it started after two or three months suggesting telogen effluvium. This finding suggests the causative mechanism of CBZ-induced hair loss is not unitary.
Annotation of PubMed abstracts

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

Search Interface  - Optional Filters (gender, age, etc...)

Results Interface  - By article or by concept or cross-table

1. Adverse event + Drug + Filters

2. Query: converted and enriched

VigiPubMed Web Service

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

LUXID Annotation Factory

7. Annotated and enriched abstracts

Copyright © 2010 TEMIS - All Rights Reserved - Slide 16
The Skill Cartridge™ Pharmacovigilance

- **Annotation module**
  - A dedicated Skill Cartridge™ is able to detect within documents:
    - *The entities*
    - *The relationships between these entities*

- **Annotations**
  - This Skill Cartridge™ is based on the Skill Cartridge™ dedicated to medical information (MER)
  - It was enriched with additional concepts and relations specific to the field of Pharmacovigilance
  - It extracts semantic relationships between an Adverse Event and Active Substances
Concepts overview

- Entities and Relationships from MER
  *Integration of MeSH 2009 for Diseases, Symptoms and Diagnostic Methods*
- Enriched entities
  *Adverse Event: MedDRA integration*
  *Drug Class and Active Substances: integration of several lexicons from Thesorimed, Drug Bank, RXNorm*
Lexicons integrated into the SC

- MeSH 2009...
  - ... is natively integrated in the MER product for Diseases, Symptoms and Diagnostic Methods

- Thesorimed...
  - ... is a database for drugs available in France. Contains regulatory information and bibliographic references

- MedDRA (Medical Dictionary for Regulatory Activities)...
  - ... integrates the medical terminology used as benchmark by regulatory authorities

- RXNorm...
  - ... contains normalized names of drugs and links to vocabularies of active molecules (First Databank, Micromedex, MediSpan, Gold Standard Alchemy and Multum)

- Drug Bank...
  - ... is a Bioinformatics and Chemoinformatics resource that combines active molecules information (chemical, pharmacological and pharmaceutical) with associated targets (structure, sequence, pathway)
 Annotation of PubMed abstracts

**Pharmacovigilant**

1. Adverse event + Drug + Filters

2. Query: converted and enriched

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

**VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance**

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table
We report a 52 year-old man presenting with an acute considerable hair loss induced by carbamazepine (CBZ). The hair loss started within a week after CBZ administration. There was no evidence of dermatitis or allergic reaction, or other cause for the hair loss. The serum concentration of CBZ was 8.6 microg/ml (therapeutic range 8-12 microg/ml). CBZ was discontinued, and the hair loss stopped within several days with new hair growth. Medication-induced hair loss is an occasional adverse effect of many drugs used for neuropsychological diseases. CBZ also induces hair loss and its frequency was reported below 2%. Only a limited number of detailed case reports describing CBZ-induced hair loss were available, and we found these cases could divide into two groups with regard to a delay in starting hair loss after administration of CBZ. In one group, the hair loss started within a week suggesting anagen effluvium and in another it started after two or three months suggesting telogen effluvium. This finding suggests the causative mechanism of CBZ-induced hair loss is not unitary.
Annotation of PubMed abstracts

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

1. Adverse event + Drug + Filters

2. Query: converted and enriched

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

Pharmacovigilant

PubMed

VigiPubMed Web Service

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table

ITM

Ontologies

Terminologies Thesaurus

Knowledge Base

Content Augmentation Manager

LUXID

Annotation Factory

Copyright © 2010 TEMIS - All Rights Reserved - Slide 22
Data sent to Mondeca

Patient

Symptom

Drug

Adverse Effect

Carbamazepine (CBZ)

52 Year-old

Report a 52 year-old man presenting with an acute considerable hair loss induced by carbamazepine (CBZ).
Annotations consolidation

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

1. Adverse event + Drug + Filters

2. Query: converted and enriched

PubMed

3. PMID List

VigiPubMed Web Service

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

Pharmacovigilant

ITM

Ontologies

Terminologies Thesaurus

Knowledge Base

Content Augmentation Manager

LUXID

Annotation Factory

Copyright © 2010 TEMIS - All Rights Reserved - Slide 24
 Annotations consolidation

❖ Two preliminary steps:
  • Modeling of a pharmacovigilance ontology
    
    *It is then loaded in the referential with all standard medical terminologies*

  ✓ It is based on CRPV database from HEGP
  ✓ It is a way to present links between AE and AS
  ✓ Integration of complementary linguistic resources is required
Referentials in application

- **Integration of Thesorimed (from LIM&Bio)**
  - Symptoms and Therapies
  - 39,855 substances with roles to Chemical Classes and Pharmacological Classes
  - Drawback: terms are mainly in French

- **Integration of MeSH**
  - Use of descriptors to normalize symptoms and diseases

- **Integration of OntoEIM (SPIM)**
  - 7,357 classes (with 1,597 defined classes)

- **Integration of drugs vocabularies from AFSSAPS**
  - Specialties and active substances

- **Integration of HEGP database (CRPV)**
  - Database sample with 2005 report cases
  - Alignment between Active Substances and Therapies (from Thesorimed)
  - Alignment between Adverse Events and Symptoms (from Thesorimed and WHO-ART)
Ontology overview: graphical view
Annotations consolidation

- Two preliminary steps:
  - *Modeling of a pharmacovigilance ontology*
    
    *It is then loaded in the referential with all standard medical terminologies*
  
  - Definition of knowledge acquisition rules (RAC)
    
    *Used to transform annotations to instances for pharmacovigilance ontology*

- RAC usage
  
  - The rules are applied to text annotations
  
  - This process generates a semantic network (RDF format)

- And after...
  
  - To check the generated annotations
  
  - And to apply consolidation rules
NomRègle: CasPharmaco01
TypeConcept: Classe
ConceptURI: http://www.vigitermes.com/n#Cas_PharmacoVigilance
NoeudIndicateur: AdverseEffect
IndicesContextuels:

{Existe: [EspaceRechercheArbre: enfant]
[NoeudIndice: Symptom]
}

{Existe: [EspaceRechercheArbre: enfant]
[NoeudIndice: Treatment]
}

{Existe: [EspaceRechercheArbre: descendant]
[NoeudIndice: LivingBeing]
{Existe: [EspaceRechercheArbre : enfant]
[NoeudIndice: Patient]
}

Valeur: text()
Position: true
Confiance: élevé
finRègle
Example with the PubMed abstract

- **Patient**: report a 52 year-old man presenting with an acute considerable **hair loss** induced by **carbamazepine (CBZ)**
Annotations rebuilding

/AdverseEffect (report a 52 year-old man presenting with an acute considerable hair loss induced by carbamazepine (CBZ))

/Animal
(52 year-old man)

/LivingBeing
(52 year-old man)

/Age
(52 year-old)

/MalePatient
(man)

/Symptom
(hair loss)

/External
(hair loss)

/Entity/Medical Entity
(Alopecia)

/Adverse Events
(Alopecia)

/Meddra
(Alopecia)

/Treatment
(carbamazepine (CBZ))

/Drug-Related
(Carbamazepine)

/Potential Pharmacological Substance
(Carbamazepine)
Evaluations

- PubMed Abstracts retrieval
- Semantic Annotations
### Evaluations

- **PubMed Abstracts retrieval**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstracts</td>
<td>480</td>
<td>184</td>
<td>218</td>
<td>1780</td>
</tr>
<tr>
<td>Relevant Abstracts</td>
<td>60</td>
<td>48</td>
<td>64</td>
<td>74</td>
</tr>
<tr>
<td>Recall</td>
<td>63%</td>
<td>51%</td>
<td>67%</td>
<td>78%</td>
</tr>
<tr>
<td>Precision</td>
<td>13%</td>
<td>26%</td>
<td>29%</td>
<td>4%</td>
</tr>
</tbody>
</table>

**DELAMARRE D, LILLO-LE LOUËT A, GUILLOT L, JAMET A, SADOU E, OUAZINE T, BURGUN A, JAULENT M.C (2010).**

Documentation in Pharmacovigilance: using an ontology to extend and normalize Pubmed queries. MedInfo 2010

Copyright © 2010 TEMIS - All Rights Reserved - Slide 34
Evaluations

- **PubMed Abstracts retrieval**

- **Semantic Annotations**
  - Collaboration with CRPV for validation
  - Selection of sentences to validate if they contain at least one AE and one AS
  - A sentence is declared wrong if:
    - *An AE or an AS is missing*
    - *One term is not well-typed*
    - *One term should not have been annotated*
  - Evaluations should be recomputed with a weight-system on annotations results
    - *Some sentences marked as wrong are nevertheless relevant*
    - *No measure of recall/precision for this validation process*

**Example :**

[Acute encephalopathy caused by sulfadiazine and trimethoprim-sulfamethoxazole in a patient with AIDS]

Adverse Events: *AIDS ; Encephalopathy*
Drug Class and Active Substance: *sulfadiazine ; trimethoprim-sulfamethoxazole*
Relation: *Adverse Effect*
Evaluations

- **PubMed Abstracts retrieval**
- **Semantic Annotations**

<table>
<thead>
<tr>
<th>Changes</th>
<th>Validation Phase I</th>
<th>Validation Phase II</th>
<th>Validation Phase III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard Annotation Module</td>
<td>Removal of noisy terms in lexicons</td>
<td>Complementary cleanup of lexicons, modification of the annotation module</td>
</tr>
<tr>
<td><strong>Global Validation</strong></td>
<td>225 sentences</td>
<td>108 sentences</td>
<td>208 sentences</td>
</tr>
<tr>
<td><strong>Correct Sentences</strong></td>
<td>54 sentences 20 %</td>
<td>45 sentences 41,7 %</td>
<td>95 sentences 45,7 %</td>
</tr>
<tr>
<td><strong>Wrong Sentences</strong></td>
<td>171 sentences 80 %</td>
<td>63 sentences 58,3 %</td>
<td>113 sentences 54,3 %</td>
</tr>
</tbody>
</table>
VigiTermes Platform

VIGITERMES SEMANTIC PORTAL – User Interfaces for Pharmacovigilance

1. Adverse event + Drug + Filters

2. Query: converted and enriched

3. PMID List

4. PMID List

5. Pubmed Abstract content

6. Annotations

7. Annotated and enriched abstracts

PubMed

VigiPubMed Web Service

Search Interface
- Optional Filters (gender, age, etc...)

Results Interface
- By article or by concept or cross-table

Pharmacovigilant

ITM

Terminologies Thesaurus

Ontologies

Knowledge Base

LUXID

Annotation Factory

Content Augmentation Manager

Copyright © 2010 TEMIS - All Rights Reserved - Slide 37
Médicament(s) utilisé(s)  Effet(s) indésirable(s) produit(s)

- targetin
- medic

- classe d'effet indésirable
- classe d'effet indésirable
- effet indésirable
- effet indésirable
- effet indésirable
- effet indésirable

TARGETETIN 75 mg Capsule molle Flacon de 100

traitement des manifestations cutanées des lymphomes cutanés T épidermo-tropes (LCT), au stade avancé et refractaires à un traitement systélique.

indications
- Hypersensibilité canute au benzotriène ou à l'un des excipients du médicament.
- Antécédents de pancreatite.
- Hypercholestérolémie incontrolée.
- Hypertriglycéridémie incontrolée.
- Hypovitaminose A.
- Maladie thyroïdotte incontrolée.

contre-indications

Cas grave
- 19/06/08 medic associe medic associe
- 19/06/08 medic associe medic associe
- 19/06/08 medic associe medic associe

Derniers articles
- Generation and validation of a revised classification for oesophageal and functional adenocarcinoma.
- Forequarter amputation for malignancy.
- Beth's palsy preceding Parkinson's disease: A Case-control study.
- Proteomic approach to studying Parkinson's disease.

Copyright © 2010 TEMIS - All Rights Reserved - Slide 38
Services integration: Actual Status

Functional demonstration platform

Operational Workflow

- VigiPubMed: operational
- PharmARTS: operational
- ITM: operational
- LUXID: operational
- SignalMiner: operational

- Terminologies + AE ontology
- Luxid® Annotations integrated
- Terminologies + drugs ontology
- Pharmaco-Vigilance ontology

Pharmacovig. Portal

- Search Interface
- Notification Interface
- Results Interface

Registered examples (HEGP DB+ manual)
Conclusion

- The Pharmacovigilance platform is validated: functionally and technically
- Some services have been daily used by the CRPV for testing and validating the tools
  - The results are partly conclusive but the platform has a real interest for pharmacovigilance professionals.
- Some technical points need to be addressed
  - Monitoring of the services
Promotion phase

- Different presentations and publications around the platform...

- ... in 2009
  - MIE 2009 (*Medical Informatics Europe Conference*)
    A knowledge management platform for documentation of case reports in pharmacovigilance (Amardeilh and al.)
  - ISoP 2009 (*International Society of Pharmacovigilance*)
    Rationale and design of a semantic portal for analysis and documentation of spontaneous reports (Amardeilh and al.)
  - AMIA 2009 (*American Medical Informatics Association*)
    Using semantic web technologies to extract knowledge about adverse drug reactions from PubMed (Bousquet and al.)
  - GDR Stic Santé (December 2009)
... in 2010

- Inauguration of Institut de la Vision / Cap Digital
  
  Presentation to Nathalie Kosciusko-Morizet (in charge of Forward Planning, Assessment of Public Policies and Development of the Digital Economy in the actual french government)

- MedInfo 2010 (13th International Congress on Medical Informatics)
  
  Documentation in Pharmacovigilance: using an ontology to extend and normalize Pubmed queries (Delamarre and al.)

- IC2010 (21es Journées francophones d’Ingénierie des Connaissances)
  
  VigiTermes : une plateforme de recherche et d'analyse des publications scientifiques au service de la pharmacovigilance (Thébault and al.)

- Text Mining Symposium 2010
Conclusion

- A 6-months promotion phase from January to June 2010
- Since June 30th, the project is over.
- We would like to go to an industrialization of the platform
  - Two different scenarios are possible depending on the client
    - Industrialization for CRPVs
    - Industrialization for pharmaceutical industries
Conclusion

- Industrialization for CRPVs – French scale
  - Today...
    
    *Our work was achieved in collaboration with one CRPV*
    *Each center addresses this problematic of documentation with its own resources*
  - Tomorrow...
    
    *We could offer a global and unified platform to the 31 CRPVs*
    *This could contribute to make the answering process for each new pharmacovigilance case more efficient*
Conclusion

- Industrialization for "Big Pharma" – World scale
  - It’s a known problematic for this type of industry
  - This platform is not the dedicated solution, but...
  - ... some VigiTermes modules could be used to help on this topic

*The dedicated Skill Cartridge™ exists and can be re-used*
VigiTermes
A search and analysis platform for scientific publications dedicated to Pharmacovigilance

Mathilde Thébault
mathilde.thebault@temis.com
info@temis.com