INFORMATIQUE ET SANTÉ

Collection dirigée par P. Degoulet et M. Fieschi

17

RISQUES, TECHNOLOGIES DE L'INFORMATION POUR LES PRATIQUES MÉDICALES

RÉDACTEURS
M. FIESCHI | P. STACCINI | O. BOUHADDOU | C. LOVIS

0101 01010101010

Springer

Risques, technologies de l'information pour les pratiques médicales

Springer
Paris
Berlin
Heidelberg
New York
Hong Kong
Londres
Milan
Tokyo

Informatique et Santé

Collection dirigée par P. Degoulet et M. Fieschi

17

Rédacteurs

M. Fieschi, P. Staccini, O. Bouhaddou, C. Lovis

Risques, Technologies de l'Information pour les Pratiques Médicales

Comptes rendus des treizièmes Journées francophones d'informatique médicale, Nice, 28-30 avril 2009

Organisées par la Société tunisienne d'informatique médicale (STIM), l'Assocation française d'informatique médicale (AIM), la Société belge d'informatique médicale (MIM), la Société suisse d'informatique médicale (SSIM), la Société québécoise d'informatique biomédicale et de la santé (SoQibs), la Société malienne d'informatique biomédicale et de santé (SoMIBS)

Springer

Rédacteurs

Pr Marius Fieschi
Service de l'Information Médicale
Hôpital La Timone
254, rue Saint-Pierre
13385 Marseille, France
marius.fieschi@ap-hm.fr

Dr Pascal Staccini
Laboratoire STIC Santé, UFR Médecine,
Université de Nice Sophia Antipolis,
06200 Nice, France
pascal.staccini@unice.fr

Dr Omar Bouhaddou
EDS (an HP company) and Department of Veterans Affairs
PO Box 3351
Rancho Santa Fe, CA 92067, États-Unis
omar.bouhaddou@va.gov

Christian Lovis
Unité d'informatique clinique
Hôpitaux Universitaires de Genève
24, rue Micheli-du-Crest
1211 Genève 4
christian.lovis@hcuge.fr

ISBN-13 : 978-2-287-99304-6

© Springer-Verlag France, Paris, 2009
Imprimé en France
Springer-Verlag France est membre du groupe Springer Science + Business Media

Cet ouvrage est soumis au copyright. Tous droits réservés, notamment la reproduction et la représentation, la traduction, la réimpression, l'exposé, la reproduction des illustrations et des tableaux, la transmission par voie d'enregistrement sonore ou visuel, la reproduction par microfilm ou tout autre moyen ainsi que la conservation des banques de données. La loi française sur le copyright du 9 septembre 1965 dans la version en vigueur n'autorise une reproduction intégrale ou partielle que dans certains cas, et en principe moyennant les paiements des droits. Toute représentation, reproduction, contrefaçon ou conservation dans une banque de données par quelque procédé que ce soit est sanctionnée par la loi pénale sur le copyright. L'utilisation dans cet ouvrage de désignations, dénominations commerciales, marques de fabrique, etc., même sans spécification ne signifie pas que ces termes soient libres de la législation sur les marques de fabrique et la protection des marques et qu'ils puissent être utilisés par chacun. La maison d'édition décline toute responsabilité quant à l'exactitude des indications de dosage et des modes d'emplois. Dans chaque cas il incombe à l'usager de vérifier les informations données par comparaison à la littérature existante.

Maquette de couverture : Nadia Ouddane

Dans la même collection

VOLUME 1 : *Informatique et Gestion des Unités de Soins* (Rédacteurs : P. Degoulet, J.-C. Stéphan, A. Venot et R.-J. Yvon), 1989.

VOLUME 2 : *Informatique et médicaments* (Rédacteurs : A. Venot et P. Degoulet), 1989.

VOLUME 3 : *Informatique au service des soins infirmiers* (Rédacteurs : L. Dusserre et M.-P. Florin), 1990.

VOLUME 4 : *Nouvelles technologies et traitement de l'information en médecine* (Rédacteurs : H. Ducrot, E. Martin et J.-R. Scherrer), 1991.

VOLUME 5 : *Nouvelles méthodes de traitement de l'information médicale* (Rédacteurs : P. Degoulet, A. Hénault et P. Corvol), 1992.

VOLUME 6 : *Information médicale et stratégies hospitalières* (Rédacteurs : M. Fieschi, P. Dujols et R. Beuscart), 1993.

VOLUME 7 : *Informatisation de l'unité de soins du futur* (Rédacteurs : J. Demongeot, P. Le Beux et G. Weil), 1994.

VOLUME 8 : *Information médicale : aspects déontologiques, juridiques et de santé publique* (Rédacteurs : L. Dusserre, M. Goldberg et R. Salamon), 1996.

VOLUME 9 : *Informatique et gestion médicalisée* (Rédacteurs : F. Kohier, M. Brémond et D. Mayeux), 1997.

VOLUME 10 : *Santé et réseaux informatiques* (Rédacteurs : A. Albert, F. Roger-France, P. Degoulet et M. Fieschi), 1998.

VOLUME 11 : *L'informatisation du cabinet médical du futur* (Rédacteurs : A. Venot et H. Falcoff), 1999.

VOLUME 12 : *L'informatique au service du patient* (Rédacteurs : M. Fieschi, O. Bouhaddou, R. Beuscart et R. Baud), 2000.

VOLUME 13 : *Télémédecine et eSanté* (Rédacteurs : R. Beuscart, P. Zweigenbaum, A. Venot et P. Degoulet), 2002.

VOLUME 14 : *Formation médicale et technologies de l'information et de la communication* (Rédacteurs : P. Staccini, M. Fieschi, D. Benchimol et R. Beuscart), 2002.

VOLUME 15 : *Présent et avenir des systèmes d'information et de communication hospitaliers* (Rédacteurs : P. Degoulet, L. Marin, L. Kleinebreil et B. Albigès), 2003.

VOLUME 16 : *Technologies de l'information et de la communication pour les pratiques médicales* (Rédacteurs : A. Harmel, R. Hajromdhane, M. Fieschi), 2004.

À la mémoire de notre collègue le professeur Bernard Huet

Liste des personnes qui ont participé activement

à la sélection des articles de ces journées

Beuscart Regis

Beuscart-Zephir Marie-Christine

Boire Jean-Yves

Bouhaddou Omar

Boyer Célia

Bricon-Souf Nathalie

Burgun Anita

Charlet Jean

Cuggia Marc

Darmoni Stéphan

De Clercq Étienne

Dufour Jean-Charles

Fieschi Marius

Geissbuhler Antoine

Jaulent Marie-Christine

Joubert Michel

Le Beux Pierre

Lebozec Christel

Lepage Éric

Lovis Christian

Rialle Vincent

Rodrigues Jean-Marie

Salamon Roger

Séroussi Brigitte

Staccini Pascal

Testa Jean

Thiessard Frantz

Venot Alain

Zweigenbaum Pierre

Qu'elles trouvent ici l'expression de nos remerciements pour la qualité du travail accompli.

Préface

Les différentes éditions des Journées francophones d'informatique médicale (JFIM) témoignent des progrès et de l'importance grandissante de notre domaine de recherche dans le champ de la qualité de la prise en charge des malades et dans le champ de la santé publique.

Les thèmes choisis pour ces treizièmes rencontres des JFIM sont au centre des débats sur l'amélioration continue de la qualité des prises en charge et de la continuité des soins, ainsi que de la maîtrise des risques. De l'importance accordée par les gouvernements et leurs agences à la transparence et à la qualité des soins jusqu'à l'annonce du dossier personnel de santé informatisé, les enjeux sont clairs. Dans tous les pays, pour tous les décideurs et les professionnels de santé, il est admis aujourd'hui que ces progrès ne pourront se réaliser sans une meilleure utilisation des technologies de l'information et de la communication (TIC) intégrées dans des systèmes d'information cohérents. Si l'objectif ne fait pas de doute, il reste à préciser les méthodes, les moyens, et à développer des expériences significatives pour progresser dans cette voie. Les progrès se construiront sur des applications innovantes basées sur des fondations scientifiques robustes et des faits probants (« *evidence based management* »). Notre communauté francophone contribue largement à ce mouvement. Le nombre et la qualité des travaux présentés dans cet ouvrage mais aussi dans les séances de posters commentés lors des journées le prouvent.

Le comité scientifique et les nombreux experts qui ont participé à la sélection des articles ont été confrontés à une tâche difficile. Nous avons reçu 49 soumissions d'articles de bonne ou de très bonne qualité. Tous les articles ont été relus et évalués par au moins deux experts, le plus souvent trois. Les 28 travaux retenus pour une présentation orale (soit 57% des soumissions) font l'objet d'une publication dans le présent ouvrage.

Les travaux présentés concernent des thématiques de recherche et de développement désormais classiques dans le domaine médical et de la santé publique. Le découpage en chapitres est toujours un exercice complexe tant les travaux sont riches et présentent de multiples facettes intéressant à la fois des aspects techniques, des résultats d'évaluations, des contributions organisationnelles... L'organisation adoptée pour ces journées a retenu un découpage en sept sessions que suivent les chapitres de l'ouvrage :

-Le premier chapitre regroupe les articles présentés lors de la session « systèmes d'information hospitaliers et systèmes d'information cliniques ». Ils traitent des systèmes d'information et des questions que sous-tendent leurs développements : leur organisation, leur architecture, leur évolution et l'interopérabilité des systèmes, leur intégration au processus de soins, le retour sur investissement.

- Les référentiels, les serveurs de terminologie, les standards d'interopérabilité et de représentation des connaissances sont au cœur de la conception des systèmes d'information et de leur évolutivité. Ce thème fait l'objet notamment du chapitre 2 intitulé « SNOMED, interopérabilité sémantique et multiterminologies ».

- La représentation des connaissances et la recherche d'information est le thème principal du chapitre 3. Le thème de la représentation des connaissances donne un socle solide sur le plan scientifique et méthodologique au développement des applications innovantes, dédiées aux professionnels comme au grand public, que nous citions plus haut.

- Le chapitre 4, « Modèles, méthodes et outils pour l'amélioration des pratiques » regroupe des travaux illustrant les problématiques de l'amélioration des pratiques et le rôle pivot que doivent jouer les TIC dans le domaine de la santé.
- Le vaste sujet de l'aide à la décision dans le contexte clinique a toujours été un domaine de recherche très actif. Les bénéfices concernant la qualité des soins escomptés pour de tels systèmes ont été perçus depuis longtemps. L'évolution des systèmes d'information vers des approches centrées sur les processus permet aujourd'hui d'intégrer efficacement les recommandations et les guides de bonnes pratiques informatisés. Le chapitre 5 regroupe des travaux sur ces aspects.
- L'usage des informations médicales à des fins de santé publique est un thème d'autant plus important que de nombreux et volumineux entrepôts de données existent aujourd'hui dans les établissements de santé. Qu'il s'agisse de méthodes et d'applications de fouilles de données ou d'usages des informations pour la recherche clinique, l'intérêt de ce thème est bien illustré par les travaux présentés au chapitre 6.
- Le développement du Web doit permettre d'améliorer la diffusion de l'information et la connaissance auprès des professionnels comme des citoyens sur les problèmes de santé. La qualité de cette diffusion, l'accessibilité et la crédibilité des informations constituent des enjeux importants. Les applications Web et les informations en ligne dignes de confiance sont les thèmes des travaux présentés dans le chapitre 7.

Avec ce 17ème numéro de la collection « Informatique et Santé », le dynamisme de la communauté scientifique francophone d'informatique médicale contribue de façon significative aux réflexions et aux développements visant au bon usage des technologies de l'information et de la communication en santé.

M. Fieschi
P. Staccini
O. Bouhaddou
C. Lovis

Table des matières

Préface IX

CHAPITRE 1

Systèmes d'information hospitaliers, systèmes d'information cliniques **1**

L'approche économétrique pour évaluer l'intérêt financier des systèmes d'information hospitaliers
R. Meyer, P. Degoulet 3

Systèmes d'information cliniques : le défi de l'intégration de systèmes commerciaux
P. Trolliard, D. Bandon, A. Geissbühler, C. Lovis 15

Une procédure d'anonymisation à deux niveaux pour créer un corpus de comptes rendus hospitaliers
C. Grouin, A. Rosier, O. Dameron, P. Zweigenbaum 23

Systèmes d'information cliniques : au cœur des informatiques hospitalières
C. Lovis 35

CHAPITRE 2

SNOMED, interopérabilité sémantique et multiterminologies **45**

SMTS® : Un serveur multiterminologies en santé
M. Joubert, B. Dahamna, J. Delahousse, M. Fieschi, SJ. Darmoni 47

F-MTI : outil d'indexation multiterminologique : application à l'indexation automatique de la SNOMED International
S. Pereira, P. Massari, A. Buemi, B. Dahamna, E. Serrot, M. Joubert, SJ. Darmoni 57

Vers une version française de la SNOMED CT
P. Fabry, R. Lemieux, A. Grant 69

Projection des relations SNOMED CT entre les termes de deux terminologies (CIM10 et SNOMED 3.5)
T. Merabti, H. Abdoune, T. Lecroq, M. Joubert, SJ. Darmoni 79

CHAPITRE 3

Représentation des connaissances et recherche d'information 89

Définition d'une structure catégorielle pour les effets indésirables liés aux médicaments
C. Bousquet, B. Trombert, P. Gasperina, L. Clavel, JM. Rodrigues 91

Recherche d'information multiterminologique en contexte : étude préliminaire
AD. Dirieh Dibad, S. Sakji, E. Prieur, S. Pereira, M. Joubert, SJ. Darmoni 101

Enrichissement des bases de connaissances en biologie par extraction de marqueurs de confiance dans la littérature scientifique
I. Jilani, MC. Jaulent 113

Interopérabilité sémantique : comparaison de la représentation du score d'APGAR en HL7V3 et OpenEHR
M. Cuggia, S. Bayat, P. Poulain, P. Pladys, H. Robert, R. Duvauferrier..... 125

CHAPITRE 4

Modèles, méthodes et outils pour l'amélioration des pratiques..... 137

Étude contrastive des documents vulgarisés et scientifiques de santé : sur la piste de la morphologie
J. Chmielik, N. Grabar 139

Un service Web pour l'annotation sémantique de données biomédicales avec des ontologies
C. Jonquet, N. Shah, MA. Musen 151

Les aspects organisationnels et logistiques de la télémédecine : expériences du projet CERTES au Mali
AK Traoré, CO. Bagayoko, M. Niang, ST. Traoré, A. Geissbühler..... 163

Évaluation du portfolio électronique et de son système d'accompagnement dédié à l'évaluation des pratiques professionnelles en technique et médecine transfusionnelles
S. Vessière, S. Jullien, P. Rouger, P. Staccini..... 171

CHAPITRE 5

Aide à la décision en pratique clinique..... 183

Prescription informatisée : de l'ordre à l'itinéraire clinique
M. Tschopp, M. Despond, D. Grauser, JC. Staub, C. Lovis..... 185

Caractérisation de l'évolution d'une base de connaissances suite à la réactualisation de recommandations de pratiques cliniques : une étude de cas sur la prise en charge des tumeurs de la vessie
B. Séroussi, J. Bouaud, A. Brizon, T. Culty, F. Mentre, V. Ravery 191

Gérer et améliorer les décisions des réunions de concertation pluridisciplinaires en oncologie : un cadre conceptuel utilisant des outils informatisés d'aide à la décision
JC. Dufour, F. Barlesi, D. Fieschi, JP. Torre, O. Chinot, M. Fieschi 203

Mise en œuvre d'un processus de conception objet pour l'implémentation de règles de décisions destinées à la validation pharmaceutique
A. Boussadi, C. Bousquet, P. Degoulet 213

CHAPITRE 6

Données médicales et usages en santé publique 225

Projet REUSE : utilisation du dossier patient informatise (DPI) en recherche clinique
N. El Fadly, N. Lucas, PY. Lastic, P. Verplancke, C. Daniel 227

Prédire l'accès à la liste d'attente de transplantation rénale : comparaison de deux méthodes de fouille de données
S. Bayat, M. Cuggia, D. Rossille, L. Frimat 239

Approche sémantique pour l'identification homogène d'effets indésirables et de médicaments dans huit bases de données de patients : une contribution au projet européen eu-ADR
P. Avillach, F. Mougin, M. Joubert, F. Thiessard, A. Pariente, JC. Dufour, M. Fieschi... 251

Constitution d'un référentiel probabiliste d'aide au codage
G. Thillay, L. Lecornu, C. Le Guillou, PJ. Garreau, V. Hery, JM. Cauvin 263

CHAPITRE 7

Web et eSanté 275

CEMARA: une plate-forme Web pour les maladies rares
C. Messiaen, L. Le Mignot, A. Rath, JB. Richard, E. Dufour, M. Ben Said, JP. Jais, A. Verloes, M. Le Merrer, C. Bodemer, G. Baujat, M. Gerard-Blanluet, E. Bourdon-Lanoy, R. Salomon, S. Aymé, P. Landais..... 277

Les informations médicales en ligne dignes de confiance : étude des spécificités en Afrique francophone
CO. Bagayoko, C. Boyer, A. Geissbühler..... 287

Plate-forme pour la surveillance épidémiologique et l'aide à la décision en Santé : Hivern@le-KhiObs : surveillance épidémiologique des pathologies hivernales de la sphère ORL chez l'enfant
L. Toubiana, T. Clarisse, F. Colin, H. Alexandre, P. Landais 299

Méthodes et évolution de la certification de site de santé sur le Web à travers l'expérience du HONcode
A. Gaudinat, C. Boyer, V. Baujard, A. Geissbühler 313

Liste des auteurs 325

Chapitre 1

Systèmes d'information hospitaliers, systèmes d'information cliniques

Informatique et Santé, 2009 (17) : 3-14
© Springer – Verlag France

L'approche économétrique pour évaluer l'intérêt financier des systèmes d'information hospitaliers

Rodolphe Meyer [1,3] et Patrice Degoulet [2,3]

[1] *Hôpitaux Universitaires de Genève, Genève, Suisse*
[2] *Hôpital Européen Georges Pompidou, Paris, France*
[3] *INSERM - UMRS 872, équipe 20, Paris, France*

Abstract

Objectives: To review the different techniques used in the quantification of the return on investment (ROI) of information technologies (IT) in the healthcare sector.

Methods: Papers for this work were obtained directly from their authors or downloaded via the University of Geneva and the INSERM bibliographic portals. Choice was based on Medline results and specific readings in selected journals of both the medical and economic fields. Focus is made on macroeconomic and microeconomic studies using techniques that could provide disruptive results for IT ROI impact evaluation in hospitals.

Results: Accounting and financial methods do not reveal the global IT investment influence in hospital although giving interesting data on specific IT projects. Econometric techniques tend to confirm a positive impact of health care IT (HIT) on hospital productivity and production. The impact seems to be majored in hospitals whose IT investments and integration levels are the highest.

Conclusion: The identification of specific IT benefits remains knowing their interminglement with people and processes. Macroeconomic analyze could provide a new set of tools to compute IT ROI in health care. All types of investments (inputs) and all the returns (outputs) can be taken into account by tools enabling the precise measurement of IT investments impact, breakeven points, and possible threshold levels. These results could lead to higher levels of IT governance in hospitals.

Keywords

Econometrics; Return on investment (ROI); Production function, IT productivity.

1 Introduction

Depuis le début des années 1970, les technologies de l'information et de la communication (TIC) en santé ont acquis une importance stratégique et économique croissante, en particulier après la publication d'études mettant en évidence leur rôle possible dans l'amélioration de la qualité des soins et la réduction des erreurs médicales [1]. Au vingt-et - unième siècle, les systèmes d'information hospitaliers (SIH) sont devenus si intriqués dans le quotidien des soignants qu'il est presque devenu impossible de prendre une décision sans en tenir compte [2]. Mais, même si tous les acteurs s'accordent sur leur importance, ils représentent des lignes budgétaires importantes qui n'ont pas encore prouvé indiscutablement leur profitabilité, ce qui provoque toujours de nombreux débats lorsqu'il

s'agit de justifier les dépenses correspondantes. Sur cette base, des taxonomies de méthodes d'évaluation ont vu le jour et de nombreuses études ont été publiées dans les revues économiques et managériales [3-6]. Malheureusement, trop peu concernent le secteur médical et en particulier les bénéfices que l'ont pourrait attendre des TIC.

L'évaluation des bénéfices indirects (qualité et continuité des soins, satisfaction des utilisateurs, optimisation des processus de soins, etc.) du déploiement d'un SIH (aide à la prescription, aide au diagnostic, archivage d'images (PACS)) concerne la grande majorité des études ; seulement 6,5 % des évaluations sont des analyses cout-bénéfice [2, 4, 7-14]. Peu d'études visent à évaluer les bénéfices directs du déploiement des SIH, et presque aucune étude n'a envisagé ces bénéfices globalement dans un esprit de stratégie de type portfolio. En outre, les résultats contradictoires de certains travaux amènent régulièrement les décideurs du milieu hospitalier à faire des choix basés seulement sur les prévisions des bénéfices indirects possibles/potentiels [15] et peuvent laisser penser que la notion même de bénéfices associés aux TIC serait un oxymore.

Cet article se focalise sur les approches évaluant le retour sur investissement (RSI) qui ont été publiées en vue de quantifier les bénéfices directs des SIH. Il passe en revue différentes approches et méthodes disponibles pour aider à la décision en matière d'investissements en TIC à l'hôpital.

2 Bénéfices des SIH

Les bénéfices des systèmes d'information hospitaliers sont de trois catégories : cliniques, organisationnels et financiers [9-10, 16-18]. Dans ces trois catégories, ils peuvent être tangibles (faciles à identifier) ou intangibles (plus difficiles à mesurer). Tous ces bénéfices sont à prendre en considération lorsque l'objectif est de déterminer la plus-value d'un SIH.

Les bénéfices financiers figurent généralement dans l'une de ces catégories : réduction des coûts, amélioration de la productivité (qui se traduit en augmentation des revenus) et amélioration de la compétitivité (qui se traduit en génération de revenus). Pour les comptabiliser, on utilise couramment le retour sur investissement (RSI ou ROI en anglais) qui est le rapport de la somme d'argent gagné ou perdu lors d'un investissement sur la somme d'argent investie. La somme d'argent gagnée ou perdue est communément dénommée *résultat net* ou *gain/perte* ou encore *profit/perte*. On se référera à l'argent investi avec les termes *capital* ou *actif*. Le RSI peut représenter le retour sur un investissement passé, en cours ou futur, mais il est calculé généralement sur une base annuelle calendaire ou fiscale. Il s'exprime habituellement en pourcentage afin de permettre une meilleure vision du gain proportionnel. Par exemple, un investissement de 1 000 € qui rapporte 50 € d'intérêts génère plus d'actifs qu'un investissement de 100 € qui rapporte 20 € ; toutefois l'investissement de 100 € possède un meilleur RSI (50 €/1000 € = 5 % de RSI *vs.* 20 €/100 € = 20 % RSI) [19]. Lorsque le gain compense le capital investi, on dit que l'on a atteint le seuil de rentabilité.

3 Evaluer le RSI des TIC avec des méthodes financières et comptables

Il existe de nombreux modèles financiers pour calculer le RSI ou essayer de l'estimer le plus correctement possible. Les méthodes les plus répandues sont l'analyse coût-bénéfice (ACB) et la mesure de la valeur actuelle nette (VAN). Ce sont principalement des méthodes comptables qui peuvent être appliquées aux projets d'investissement.

3.1 L'analyse coût-bénéfice

L'analyse coût-bénéfice correspond au quotient des bénéfices totaux apportés par un projet spécifique sur la somme des investissements mis en œuvre pour le réaliser. Un rapport supérieur à 1 indique un RSI positif. Mais aussi simple que ce calcul paraisse, les données nécessaires à sa réalisation sont loin d'être faciles à réunir.

Une étude récente utilisant l'ACB a été commanditée par la commission européenne (direction générale de l'information de la société et des médias, Bruxelles) [11]. Dix sites européens avec des implémentations TIC très variées (dossier patient électronique local ou national, aide à la prescription, aide à la décision clinique, dispatching logistique...) ont été évalués au niveau de l'impact des TIC dans leurs bénéfices. Les coûts et les bénéfices (ramenés à leur valeur en euros) prennent en considération tous les acteurs du milieu hospitalier (citoyens, professionnels, tiers payeurs, débiteurs, prestataires externes, etc.). Les résultats obtenus semblent montrer que les sites ont obtenus un impact positif, à la fois des déploiements effectués et de l'utilisation de leurs SIH, sur des périodes parfois très courtes (moins de une année) [11].

Une autre étude intéressante utilisant le RCA a été réalisée en 2006 aux États-Unis. Le *Brigham and Women's Hospital* de Boston (centre universitaire de 720 lits d'adultes) a voulu évaluer le rapport coût bénéfices de son système d'aide à la prescription. Entre 1993 et 2002 ce site a dépensé 11,8 millions de dollars pour développer, déployer et maintenir son SIH. Sur ces dix années, le système leur aurait permis d'économiser 28,5 millions de dollars. Les plus grandes économies auraient été réalisées sur l'adaptation des doses liées à la fonction rénale, sur l'optimisation du temps de travail des infirmières, sur l'aide à la prescription de médicaments spécifiques et la prévention des effets indésirables. Ces investissements substantiels ont toutefois mis six années pour atteindre leur seuil de rentabilité et les bénéfices mesurés ont été obtenus par analyse de la littérature et non par mesure directe [20].

3.2 La valeur actuelle nette (VAN)

La VAN représente la différence entre les cash-flows actualisés à la date zéro et le capital investi. On calcule pour chaque période future (pendant toute la période d'utilisation prévue de l'investissement) les recettes et les dépenses générées par l'investissement, puis on détermine les flux économiques de l'investissement. On actualise ensuite les flux nets annuels à la date d'investissement et on compare l'ensemble au montant initial. En d'autres termes, c'est la somme des bénéfices nets durant la vie du projet ajustée du coût du capital [2]. La VAN est donc un indicateur de la plus-value qu'apporte un investissement ou un projet à l'entreprise, ce qui permet de déterminer si un investissement est rentable ou non [10]. Le secret est d'être exhaustif dans l'identification des coûts. C'est une méthode bien appropriée aux projets sur le long terme [21]. Peu d'évaluations de projets TIC ayant utilisé cette méthode ont été publiées montrant une VAN positive [22-23]. Toutefois l'une d'elles, publiée en 2000, a montré que la télécardiologie, en plus d'améliorer la qualité des soins, était rentable dans le cadre de la néonatalogie sur la région de Washington DC [24].

3.3 Les autres méthodes

Très proche de la VAN, le taux de rentabilité interne (TRI) correspond à la comparaison entre le montant des fonds investis et la valeur actuelle de l'entreprise rapportée à la durée de l'investissement. Le TRI est un taux de rentabilité propre à chaque projet

d'investissement. Ce taux ne permet pas de déterminer, dans l'absolu, si un investissement est rentable ou non pour un investisseur. C'est un outil d'arbitrage entre différents projets d'investissement. L'investisseur choisira celui pour lequel le TRI est le plus élevé [2, 10, 21].

Il existe aussi des méthodes qui prennent en compte la mesure du risque des investissements d'un projet. Citons l'analyse de rentabilité [14, 25-27], l'analyse de sensibilité [26, 28, 29], l'analyse de scénario [27], la simulation Monte Carlo [26, 30], le coût moyen du capital pondéré [31] et l'approche par option réaliste [32-34]. Par ailleurs, de nombreuses sociétés de consulting ont développé leurs propres méthodes d'analyse de l'impact des projets TIC. Citons l'*economic value added* de Stern Stewart & Co [35], le *business value index* du Hackett Group [36], les *balanced scorecards* de Palladium [37], l'*information economics methodology* du Beta Group [38], l'*IT Performance Management Group Method* [39], le *total economic impact* de Forrester Research [40], et la *total value of opportunity* de Gartner Group [41].

3.4 Les limites des méthodes comptables et économiques

De par leur nature intrinsèque (financière/économique) les méthodes précédentes ne prennent pas en compte les aspects les moins tangibles de l'impact des TIC. L'amélioration de la satisfaction des utilisateurs (professionnels et patients) est l'exemple typique. L'impact potentiel saute aux yeux de tous, mais reste complexe à traduire financièrement [9, 42]. Ces techniques ont aussi tendance à sous-estimer certains aspects du capital TIC. L'exemple le plus évident est représenté par les nombreux éléments du parc informatique de l'hôpital qui ont dépassé leur délai habituel d'amortissement de 5 ans et qui sont toujours en activité bien que non valorisés. Cette approche aura tendance à biaiser les études qui prennent en compte la production liée aux investissements TIC puisqu'une partie de la production analysée sera le fruit d'actifs invisibles sur un plan comptable. Par ailleurs, l'acuité des modèles de calcul de RSI se dégrade en proportion de leur complexité. Il existe au moins 15 modèles d'analyse RSI que les décideurs peuvent utiliser pour les TIC. Cela est finalement problématique, car plus la méthode et compliquée, plus elle est chronophage ; cela se faisant bien sûr au détriment du temps passé sur le projet global [41].

Isoler un projet informatique hospitalier pour essayer d'en mesurer la valeur ajoutée n'est pas toujours pertinent. Ou, comme Steve Ulfeder l'a mentionné dans les colonnes de Computerworld, « l'idée qu'il existe des projets TIC doit être abandonnée. Il n'existe que des projets visant à améliorer les processus commerciaux, à développer de nouveaux produits ou services, à améliorer la prise en charge des utilisateurs ou à améliorer les performances » [43]. En outre, tous les outils financiers et comptables décrits précédemment, lorsqu'ils sont utilisés pour évaluer un investissement futur, ont une tendance nette à freiner les velléités d'innovation. Ou comme l'a dit CM Christensen dans la *Harvard Business Review* de janvier 2008, ils détournent les ressources des investissements dont la rentabilité ne se situe pas dans un horizon immédiat [44].

4 L'approche économétrique pour évaluer le RSI

L'économétrie est une branche des mathématiques qui s'est développée afin d'appliquer des techniques de calcul statistiques aux principes de l'économie moderne [45, 46]. La mise en application des méthodes économétriques au niveau d'une entreprise (microéconomie) est possible. Dans ce cadre, la mesure de l'efficacité du capital que l'on nomme RSI se fera grâce à l'utilisation des fonctions de coût ou de production [47-50].

Utilisées depuis la fin des années 1980 dans l'industrie pour évaluer l'intérêt des TIC, les fonctions de production ont démontré leur pertinence dans l'analyse de la contribution des systèmes d'information [51, 52]. Si l'on souhaite les appliquer à l'hôpital, les modèles économétriques doivent être adaptés pour intégrer les spécificités du domaine de la santé. Leur emploi fait ressortir des problématiques classiques quant aux mesures de la production, de l'efficience et de la productivité [53-55]. Mais il reste essentiel d'être capable d'estimer les bénéfices des TIC dans le domaine médical.

La première approche économétrique des TIC en santé a été effectuée par les écoles de commerce de trois universités américaines en 2000 (*Texas Tech University College of Business, the University of Illinois at Chicago College of Business, the University of Arizona Karl Eller School of Management*). Nirup Menon et al. ont analysé l'impact des TIC en santé en utilisant des données hospitalières allant de 1976 à 1994 [56]. Les données de l'étude (1 130 observations) furent obtenues du département de la santé de l'état de Washington et excluaient les structures très spécialisées (psychiatrie, rééducation, etc.). Le capital a été subdivisé en trois composantes - capital TIC, capital médical, et capital médical TIC. Le travail a été subdivisé en deux composantes - travail médical et travail TIC. Les auteurs ont utilisé pour leurs calculs une technique de frontière stochastique. Étant donné que cette approche suppose que tous les processus de production sont, par nature, inefficaces, les paramètres de leur modèle ont capturé les inefficacités des processus de production au niveau de l'hôpital [57] de façon réaliste. Les résultats obtenus avec cette méthode ont indiqué que les TIC contribuent positivement à la production de services (*output*) dans le secteur de la santé des États-Unis. Le travail TIC et le travail médical ont affiché une influence positive sur la production ainsi qu'un impact positif sur la moyenne des recettes marginales. Toutefois, il apparaît que le capital médical semble être négativement associé à la production au cours de cette période, ce qui pourrait s'expliquer par une volonté de contenir les coûts en favorisant les soins ambulatoires au détriment des hospitalisations [56].

En 2003, Ko et Osei-Bryson ont exploré l'impact sur la productivité des TIC dans le secteur de la santé en utilisant des régressions de type *regression spline* (RS) avec une approche fondée sur une fonction de production « translog » permettant la possibilité d'étudier les interactions entre les variables prédictives (capital non TIC, travail non TIC, capital TIC) [58]. Ils ont utilisé les mêmes données que Menon et Lee [56]. Le recours à des régressions par directions révélatrices leur a permis de comprendre la relation complexe entre les investissements dans les TIC et la productivité dans le secteur de la santé [59]. Les résultats de cette étude suggèrent que chaque variable d'entrée a une interaction avec d'autres variables d'entrée. Les résultats suggèrent également que, sous certaines conditions, les investissements en capital TIC ont un impact positif sur la productivité, et que cet impact n'est pas uniforme, mais est conditionné par le montant investi dans le capital TIC et les investissements dans le capital non TIC. Ainsi, l'identification d'un niveau optimal de l'investissement dans chaque variable, pourrait conduire à l'accroissement de la productivité au niveau de l'hôpital.

Sur ces bases nous avons procédé, en 2006, à une analyse économétrique sur les données de 17 hôpitaux parisiens parmi les 38 hôpitaux de l'Assistance Publique Hôpitaux de Paris (AP-HP) [16]. L'étude avait plusieurs objectifs : démontrer qu'une fonction de production économétrique classique était adaptable à des hôpitaux français ; calculer et analyser la part de chaque variable d'entrée dans l'explication des productions des hôpitaux ; mesurer et d'évaluer l'impact des investissements TIC sur la production ; et enfin comparer la part des TIC dans les résultats de la production sur deux groupes d'hôpitaux séparés sur la base

du niveau d'intégration de leur système d'information hospitalier (SIH). Les hôpitaux ont été sélectionnés en fonction de leur taille (plus de 350 lits) et de leur activité (soins aigus et courts séjours). Les données utilisées dans cette étude vont de 1998 à 2005. Les 17 hôpitaux ont été divisés en deux groupes selon leur niveau d'intégration TIC. Le groupe 1 est composé de 11 hôpitaux ayant principalement un système d'information ancillaire (administratif et de gestion, laboratoire, radiologie, pharmacie). Le groupe 2 est composé de 6 hôpitaux qui ont installé ou ont commencé à installer des SIH intégrés au cours de la période d'étude. En utilisant une fonction de Cobb-Douglas [45, 50] les liens entre la production et trois différents inputs (capital, quantité de travail, technologies de l'information) ont été évalués en supposant leur élasticité de substitution constante [48]. Les résultats ont montré qu'il est pertinent d'utiliser l'analyse économétrique dans des hôpitaux publics en France. En outre, avec une part de travail dans les hôpitaux supérieure à ce qui est généralement admis dans les secteurs industriels et de services en France, les éléments apportés par la fonction de production de Cobb-Douglas ont souligné l'importance du facteur humain dans l'explication des résultats de production de l'hôpital. Les résultats semblent également confirmer l'impact positif des technologies de l'information sur la production de l'hôpital au cours d'une période de huit années de suivi. Si l'on compare les deux groupes, le calcul tend à révéler que la part des TIC expliquant la production observée est environ 1,7 fois plus importante dans les structures ayant un niveau plus élevé d'intégration. Cette tendance augmente au fil du temps.

En 2007, *PricewaterhouseCoopers* (PwC) et des chercheurs à la *Wharton School* de l'Université de Pennsylvanie ont publié une analyse macroéconomique majeure portant sur les coûts et la qualité des hôpitaux des États-Unis [60]. Le double objectif était de confirmer les relations entre les investissements TIC et d'autres mesures de performance de l'hôpital à l'aide de statistiques avancées et de techniques économétriques. Ils ont voulu aussi établir si de telles relations pouvaient soutenir l'affirmation que les investissements TIC faits par les hôpitaux améliorent effectivement leur performance organisationnelle. Ils ont recueilli les données de près de 6 000 hôpitaux des États-Unis sur une période allant de 1999 à 2004. Après l'exclusion des hôpitaux psychiatriques, des hospices, des petits hôpitaux (< 50 lits), des structures de réadaptation et de soins chroniques, ils se sont retrouvés avec près de 2 000 établissements. L'échantillon a été divisé en deux parties : hôpitaux à but lucratif et à but non lucratif. Pour analyser ensuite l'énorme base de données résultante, ils ont utilisé des outils économétriques standards et des techniques statistiques avancées. Cela a donné naissance à une fonction de coût, plutôt qu'une fonction de production utilisée dans d'autres études. Les résultats montrent que des niveaux élevés d'investissement TIC sont corrélés avec une optimisation des coûts et des performances comptables, bien que parfois légèrement, cette amélioration semble augmenter au fil du temps. Les investissements dans les hôpitaux de soins aigus sont associés à une réduction des dépenses de fonctionnement, mais seulement après que les hôpitaux aient atteint un certain niveau d'investissement. En d'autres termes, les premières étapes d'acquisition des technologies de l'information sont très couteuses jusqu'à ce qu'une «masse critique» soit atteinte ; à ce stade, la relation devient neutre sur une période de temps variable, pour finalement devenir positive. Un autre résultat intéressant est qu'il y a un décalage naturel entre la mise en œuvre de la technologie et la réalisation des bénéfices. Des réductions de coûts se sont produites dans la même année que l'acquisition des TIC, mais généralement, il a fallu deux ans pour atteindre le seuil de rentabilité. Aucun autre effet de l'introduction du décalage n'a été observé. Les hôpitaux à but non lucratif, semblent parvenir à une plus petite réduction des coûts lorsqu'ils implémentent des TIC, que les hôpitaux à but lucratif et atteindre leur seuil de rentabilité à des niveaux plus élevés de capital TIC investis [60]. Les hôpitaux ayant des niveaux plus élevés d'investissements en TIC ont tendance à avoir

un taux de mortalité inférieur (risque ajusté en fonction du *case mix*). L'effet est indépendant des coûts par lit. Le taux de mortalité ne représente certainement pas le « *gold standard* » du niveau de qualité, mais comme N. Beard l'a déclaré : « le taux de mortalité est accepté comme faisant partie d'un ensemble stable de mesures qui forment un indicateur raisonnable de la qualité des soins » [60], de sorte que la constatation que les investissements TIC pourraient correspondre à un plus faible taux de mortalité est une indication prometteuse de l'amélioration possible de la qualité globale [61].

5 Discussion et conclusion

Bien que la littérature comprenne des résultats encourageants, plus de recherche est nécessaire pour comprendre la relation entre investissements TIC et résultats financiers d'un point de vue macro-économique. Tout d'abord, des études devront préciser la durée de la période de décalage existant entre la première année de l'investissement et le seuil de rentabilité. Certaines études rapportent un délai de cinq ans ou plus [20], d'autres ont mesuré deux ans de décalage [60], tandis que certaines rapportent un retour sur investissement positif la même année que celle du déploiement [10]. Selon notre expérience, il faut environ deux à trois ans pour déployer un système d'information hospitalier et encore plus de temps (un ou deux ans) pour qu'il puisse prétendre avoir atteint ses objectifs d'alignement. Les futures études économétriques devraient intégrer ce décalage temporel dans leurs calculs.

Dans les différents travaux économétriques, les variables TIC prises en compte ont été construites d'une manière qui peut être discutée. La plupart des études ont agrégé le capital TIC avec le travail TIC. C'est une approche intéressante, mais la granularité des variables étudiées pourrait être élargie. L'examen de l'influence du travail TIC et du capital TIC pourrait être fait séparément ou en subdivisant le travail, en travail TIC, médical et non médical, afin d'examiner plus en détail l'influence positive ou négative des investissements dans chacune de ces dimensions. Les propriétés mathématiques de la fonction de production permettent ces extensions.

La notion même de capital IT mérite d'être analysée. En effet, dans un hôpital possédant un très haut niveau d'intégration, le capital utilisé pour acheter un nouveau PET-Scan pourrait être assigné aux TIC, compte tenu que l'appareil pourrait être envisagé comme un nouveau dispositif de production d'images en réseau, situation qui serait moins pertinente dans un hôpital avec un faible niveau d'intégration TIC. Cette spécification plus détaillée des variables serait un premier pas vers une meilleure compréhension de la complexité des relations qui peuvent exister entre les variables au moment de l'investissement. En outre, les avantages associés aux TIC semblent être plus élevés lorsqu'ils sont accompagnés d'une refonte cohérente des processus cliniques. Les effets de ces processus de refonte seraient intéressants à comparer entre des structures dont la politique à ce niveau aurait été différente [2, 60, 61].

L'identification correcte des hôpitaux est également un point à développer. Les études ne peuvent agréger des centres hospitaliers à but lucratif et à but non lucratif. A l'intérieur de ces deux groupes, il est important de déterminer quelles catégories de structure poursuivent les mêmes objectifs de soins. Les meilleurs résultats de mesure de rentabilité de projets TIC calculés dans des institutions étrangères envisagées individuellement ne peuvent pas fournir le niveau de confiance nécessaire aux investisseurs pour souscrire localement à un projet TIC ambitieux. Des cohortes bien construites seront indispensables. Sur cette base, l'extension des études, principalement faites aux États-Unis actuellement, devrait être effectuée dans d'autres pays concernés par le même niveau de TIC afin de fournir des données qui prennent en considération les spécificités des systèmes de santé nationaux.

Dans cette étude, nous avons montré que, contrairement à ce que l'on voit encore écrit trop souvent, les études factuelles qui tendent à démontrer la valeur ajoutée des TIC ne sont plus rares. La majorité des contributions montrent des résultats positifs provenant des diverses implémentations de TIC, même si les questions théoriques et méthodologiques demeurent [2, 21, 44, 60, 62, 63]. Les technologies de l'information sont un mélange composite de matériels, de logiciels, de connaissances, de niveau d'intégration, d'appui opérationnel et d'infrastructures. Trop d'études analysent l'impact de projets TIC unique sur une période où il a été démontré que le véritable retour sur investissement ne peut se mesurer qu'en envisageant le projet d'un point de vue global (type portolio) [16, 60]. Ce faisant, ces hôpitaux investissent beaucoup de temps et d'efforts dans des processus de mesure pour tenter de suivre les rendements de leurs investissements. Certains commencent maintenant à s'interroger sur la valeur d'un tel effort. Dans de nombreux cas, les technologies de l'information sont si inextricablement liées avec les processus et les personnels que l'identification des bénéfices d'une technologie spécifique ne rendra compte que des bénéfices marginaux [1]. En outre, les méthodologies financières utilisées par les experts ne se sont pas encore révélées posséder une grande précision et un grand nombre des méthodes utilisées renferment des approximations inhérentes et des postulats économiques qui soulèvent parfois plus de questions que de réponses [2, 5, 7, 10, 42-44, 64].

La productivité est sans doute, et c'est discutable, la plus importante statistique économique. Son corollaire, le retour sur investissement, produit l'étalon sur lequel chaque décision d'investissement futur sera décidé, d'où l'importance de trouver des résultats les plus proche de la réalité de terrain. Pour poursuivre cet objectif, ce travail suggère que de nouvelles approches économiques, dérivées de considérations macroéconomiques, et sur la base d'outils économétriques, pourraient s'avérer plus fiables dans l'évaluation de la valeur réelle du retour sur investissement des TIC et en particulier dans le domaine de la santé [16,56,58,60].

Une autre notion en faveur l'intérêt des outils macro-économiques, réside dans le fait que, lorsque l'on utilise des outils financiers, les experts comparent toujours les investissements potentiels aux situations possibles si l'on ne faisait rien (aucun investissement). C'est une erreur commune que l'on peut qualifier de *méprise entropique*. En effet, le fait de ne pas investir ne garantit aucunement que le système considéré restera tel qu'il est au moment de la prise de décision. Rien n'est moins sûr. Lorsque l'on compare un scénario « *go* » avec un scénario « *no go* », il faudrait calculer l'augmentation de l'entropie de la situation « *no go* » pour avoir une vrai vision de l'écart entre les deux. Et pour aller encore plus loin, la situation d'un hôpital pourrait être pire qu'avant, après avoir fait un investissement, mais meilleure que celle qu'il aurait connu s'il n'avait rien fait [62]. Garder cette dimension à l'esprit lors des discussions sur la pertinence de tel ou tel investissement, pourrait faire économiser du temps et de l'argent, et aussi faire percevoir que les choix ne peuvent se faire uniquement sur des critères de rentabilité positive comptable [60].

Dans les économies avancées, les TIC sont une source prometteuse de croissance de la productivité, mais elles n'auront que peu de contribution directe sur la performance globale d'un hôpital si elles ne sont pas combinées avec des investissements complémentaires dans le capital humain, les pratiques de travail, et certaines restructurations organisationnelles [65, 66]. L'évaluation de l'impact des TIC dans les hôpitaux doit franchir une étape et prendre en compte la totalité des « *inputs* » avec la globalité des « *outputs* », quelle que soit la source de chacun d'eux. Les variables contextuelles sont des facteurs importants à

considérer dans la détermination de l'impact des investissements TIC sur la performance organisationnelle [67, 68], mais peuvent rester très difficiles à quantifier financièrement, altérant potentiellement toutes les analyses qui cherchent l'exhaustivité. Seule une analyse du niveau de celle de l'alignement stratégique peut être pertinente dans les cas concrets, surtout depuis que nous avons identifié que les analyses quantitatives basées sur la fonction de production possèdent le niveau de rigueur nécessaire [69]. Dans un contexte général d'incitations aux économies et à la durabilité, on peut dire que le monde de la santé se déplace progressivement de la médecine factuelle à la médecine fructueuse (ou du moins rentable à l'aide de preuves de rapport coût-efficacité élevé) [70]. L'analyse économétrique ne cherche pas à devenir une « théorie unifiée » de l'analyse des TIC mais les outils économétriques donneront des renseignements utiles pour relever le niveau de la gouvernance des TIC dans les hôpitaux.

Références

[1] Kohn LT, Corrigan JM, Donaldson MS. *To Err Is Human: Building a Safer HealthSystem.* [ed.] Institute of Medicine. Washington DC: National Academy Press, 1999.

[2] Arlotto P, Oakes J. *Return on investment: maximizing the value of healthcare information technology.* Chicago: HIMMS, 2003.

[3] Delone WH, McLean ER. Information Systems Success: The Quest for Dependent Variable. *Information Systems Research* 1992; 3: 60-95.

[4] Grover V, Jeong SR, Segars AH. Information system effectiveness: The construct pace and patterns of application. *Information and Management* 1996; 31 (4): 117-91.

[5] Van der Loo RP, Van Gennip EMSJ, Baker AR. Evaluation of automated information systems in health care: an approach to classifying evaluative studies. *Comput Methods Programs Biomed* 1995; 48 (1): 45-52.

[6] Brennan A, Chick S, Davies R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ* 2006; 15: 1295-310.

[7] Pekka T, Hannu S. The Economic evaluation of Medical Information Systems. *A research paper presented at the 18th Nordic Health Economists' Study Group Meeting.* 22-23 August, 1997: 21.

[8] Gardner RM, Hulse RK, Larsen KG. Assessing the effectiveness of a computerized pharmacy system. *Symp. on Comput. Applications in Medical Care.* 1990, 14: 668-72.

[9] Menachemi N, Saunders C, Chukmaitov A, Brooks R. Hospital quality of care: Does information technology matter? The relationship between information technology adoption and quality of care. *Health Care Manage* Rev 2008; 33 (1): 51-59.

[10] Menachemi N, Brooks RG. *Exploring the return on investment associated with health information technologies.* Florida State University College of Medecine: Center for Patient Safety, 2005.

[11] Stroetmann KA, Jones T, Dobrev A, Stroetmann VN. *eHealth is worth it: the economic benefits of implemented eHealth solutions at ten European sites.* Luxembourg: Office for Official Publication of the European Communities, 2006.

[12] Barlow S, Johnson J, Steck J. The economic effect of implementing an EMR in an outpatient clinical setting. *J Healthc Inf Manag* 2004; 18 (1): 46-51.

[13] Cooper J. Organization, Management, Implementation and Value of EHR Implementation in a Solo Pediatric Practice. *J of Healthcare Info Manag* 2004; 18 (3): 51-55.

[14] Wang SJ, Middleton B, Prosser LA, Bardon CG, Carchidi PJ, Kittler AF, et al. A cost-benefit analysis of electronic medical records in primary care. *Am J Med* 2003; 114, (5) : 397-403.
[15] Featherly K, Garets D, Davis M, Wise P, Becker P. Sharpening the Case for Returns on Investment from Clinical Information Systems. *Electronic Healthcare* 2006; 5 (3): 101-110.
[16] Meyer R, Degoulet P, Omnes L. Impact of Health Care Information Technology on Hospital Productivity Growth: a Survey in 17 Acute University Hospitals. *Medinfo* 2007; 12 (1): 203-7.
[17] Ransom SB, Joshi MS, Nash DB. *The Healthcare quality book: vision, strategy and tools.* Chicago: Health Administration Press, 2004.
[18] Tan J. *e-Healthcare Information Systems: An Introduction for Students and Professionals.* San Francisco: Jossey-Bass, 2005.
[19] [En ligne novembre 2008] http://en.wikipedia.org/wiki/Return_on_investment.
[20] Kaushal R, Jha AK, Franz C, Glaser J, Shetty KD. Return on investment for a computerized physician order entry system. *J Am Med Inform Assoc.* 2006, 13 : 365-7.
[21] Anthes G. [En ligne novembre 2008] http://www.computerworld.com/special_report/000/000/100/special_report_000000180_primary_article.jsp.
[22] McLean RA. Cost-volume-profit and net present value analysis of health information systems. *Top Health Inf Manage* 1998; 19 (1): 39-47.
[23] Kaplan JG. The net present value of investments in health. *Med Interface* 1996; 9 (11): 94-6.
[24] Rendina MC. A net present value analysis of neonatal telecardiology. *Telemed Today* 2000; 8 (2): 23-5.
[25] Boles K, Fleming S. Breakeven under capitation: pure and simple? *Health Care* 1996; 21 (1): 38-47.
[26] Gapenski, LC. Accuracy of investment risk models varies. *Healthc Financ Manage* 1992; 46 (4): 40-44.
[27] Gapenski C. *Healthcare Finance: an introduction to accounting and financial management.* Chicago: Health Administration Press, 2005.
[28] Saltelli A, Ratto M, Andres T, Campolongo F. *Global Sensitivity Analysis. The Primer.* Hoboken: John Wiley & Sons, 2007.
[29] Meltzer MI. Introduction to health economics for physicians. *The Lancet* 2001; 358: 993-8.
[30] Palmer AJ, Annemans L, Roze S, Lamotte M, Lapuerta P, Chen R, *et al.* Cost-effectiveness of early irbesartan treatment versus control or late irbesartan treatment in patients with type 2 diabetes, hypertension, and renal disease. *Diabetes Care* 2004; 27 (8): 1897-903.
[31] Holmes RL, Schroeder RE, Harrington LF. Objective Risk Adjustment Improves Calculated ROI For Capital Projects. *Healthcare Financial Management* 2000, 12.
[32] McGrath R, Ferrier WJ, Mendelow AL. Real Options as Engines ofChoice and Heterogeneity. *Academy of Management Review* 2004; 29 (1): 86-101.
[33] McGrath, R and Nerkar, A. Real Options Reasoning and a New Look at the R&D Investment Strategies of Pharmaceutical Firms." 25: 1-21. *Strategic Management journal* 2004; 25: 1-21.
[34] Williams DR, Walker JA, Hammes PH. Real options reasoning in healthcare: an integrative approach and synopsis. *Journal of Healthcare Management* 2007; 52 (3): 170-86.
[35] [En ligne novembre 2008] http://www.ssco.com/.
[36] [En ligne novembre 2008] http://www.thehackettgroup.com/.

[37] [En ligne novembre 2008] http://www.thepalladiumgroup.com/.
[38] [En ligne novembre 2008] http://www.the-beta-group.com/.
[39] [En ligne novembre 2008] http://www.itpmg.com/.
[40] [En ligne novembre 2008] http://www.forrester.com/rb/research.
[41] [En ligne novembre 2008] http://www.gartner.com/.
[42] Hoffman T. [En ligne novembre 2008] http://www.computerworld.com/managementtopics/roi/story/0,10801,78541,00.html
[43] Ulfelder S. [En ligne novembre 2008] http://www.computerworld.com/special_report/000/000/100/special_report_000000180_primary_article.jsp
[44] Christensen CM, Kaufman SP, Shih WC. Innovation Killers: How Financial Tools Destroy Your Capacity to Do New Things. *Harvard Business Review* 2008; 86 (1): 98-105.
[45] Greene W. *Econometric Analysis*. Upper Saddle river, New Jersey: Pearson Education Inc, 2003.
[46] Gujarati D. *Basic Econometrics*. Columbus : McGraw-Hill Higher Education, 2003.
[47] Cobb CW, Douglas PH. A Theory of Production. *American Economic Review* 1928; 18 (1): 139-65.
[48] Solow RM. A Contribution to the Theory of Economic Growth. The Quarterly *Journal of Economics* 1956; LXX: 65-94.
[49] Hitt LM, Brynjolfsson E. Productivity, business profit-ability, and consumer surplus: three different measures of information technology value. *MIS Quarterly* 1996; 20 (2): 121-142.
[50] Stadelmann D. *Fonction de production – illustration des propriétés pour l'analyse économique*. Fribourg : Séminaire de l'université de Fribourg, 2003-2004.
[51] Weill P. The relationship between investment in information technology and firm performance: a study of the valve manufacturing sector. Information *Systems Research* 1992; 3 (4): 307-333.
[52] Brynjolfsson, E. The productivity paradox of IT. *Communication of the ACM.* December 1993; 36, 12.
[53] Koc C. The productivity of health care and health production functions. *Health Econ* 2004; 13 (8): 739-47.
[54] Hollingsworth B, Dawson PJ, Maniadakis N. Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Manag Sci* 1999; 2 (3): 161-72.
[55] Grosskopf S, Valdmanis V. Measuring Hospital performance: a non-parametric approach. *Journal of Health Economics* 1987; 6: 87-109.
[56] Menon N, Lee B, Eldenburg L. Productivity of Information Systems in the Healthcare Industry. *Information Systems Research* 2000; 11 (1): 83-92.
[57] Lovell CAK. *Production frontiers and production efficiency in The Measurement of Productive Efficiency: Techniques and Applications.* New-York: Oxford University Press, 1993.
[58] Osei-Brison KM, Ko M. Exploring the relationship between information technology investments and firm performance using regression splines analysis. *Information & management* 2004; 42: 1-13.
[59] Eubank R. *Spline Smoothing and Nonparametric Regression*, Marcel. New York : Marcel Dekker Inc, 1987.
[60] Beard N, Elo K, Hitt LM, Housman MG, Mansfield G. Information technology and hospital performance: An econometric analysis of costs and quality.
[En ligne novembre 2008] PricewaterhouseCoopers : http://www.pwc.com/healthcare, 2007.

[61] Holland M. Perspective on the value of HIT investment. Health Industry Insights. [En ligne novembre 2008] http://www.healthindustry-insights.com/HII/getdoc.jsp?containerId=HI206330.
[62] Menachemi N, Burkhardt J, Shewchuk R, Burke D, Brooks RG. Hospital Information Technology and Positive Financial Performance: A Different Approach to Finding a ROI. *Journal of Healthcare Management* 2006; 51, 1.
[63] Schulman J, Kuperman GJ, Kharbanda A, Kaushal R. Discovering How to Think about a Hospital Patient Information System by Struggling to Evaluate It: A Committee's Journal. *J Am Med Inform Assoc* 2006; 14: 537-541.
[64] Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004; 8, 36.
[65] Brynjolfsson, E. The IT Productivity GAP. Optimize. 2003; 21.
[66] Bleom J, VanDoorn M, Mittal P. *Making IT governance work in a Sarbanes-Oxley world.* Oboken: Willey & Sons Inc, 2006. pp. 46-90, 153-157.
[67] Li R, Ye LR. Information technology and firm performance: linking with environmental, strategic and managerial contexts. *Information and Management* 1999; 35: 43-51.
[68] Teo TSH, Wong PK, Chia EH. Information Technology (IT) investment and the role of a firm: an exploratory study. *International Journal of Information Management* 2000; 20: 269-286.
[69] Kumbhakar SC, Tsionas EG. Estimation of stochastic frontier production functions with input-oriented technical efficiency. *Journal of Econometrics*. 2006, 133: 71-96.
[70] Maynard A, Kanavos P. Health economics: an evolving paradigm. *Health Econ* 2000; 9: 183-190.

Adresses de correspondance

Dr Rodolphe Meyer, Service d'informatique médicale, Hôpitaux Universitaires de Genève
24, rue Micheli-Du-Crest , CH-1211 Genève 14, Suisse
E-mail : rodolphe.meyer@sim.hcuge.ch

Pr Patrice Degoulet, Département d'Informatique Hospitalière – HEGP
20 rue Leblanc , 75015 Paris , France
E-mail : patrice.degoulet@egp.aphp.fr

Informatique et Santé, 2009 (17) : 15-22
© Springer – Verlag France

Systèmes d'information cliniques : le défi de l'intégration de systèmes commerciaux

Patrice Trolliard, David Bandon, Antoine Geissbühler, Christian Lovis

Service d'informatique médicale, Hôpitaux Universitaires de Genève, Genève, Suisse

Abstract

This paper reports the integration complexity of a commercial departmental information system within a global clinical information system.

This integration requires multiple interfaces to implement in order to reach a fully computerized process management (billing, scheduling, communication with medical equipments, clinical information exchange with the electronic patient record, product management, audit trail, user authentication.). This integration implementation takes benefit of a component-oriented architecture delivering services and messages, mixing IHE approach and a notification-based middleware. A case illustration is documented in cardiology with the integration of a cardiology information system (CIS).

Keywords

Clinical information systems; Interoperability; Architecture; Commercial system; Integration; Interface; IHE; HL7; DICOM

1 Introduction

Les systèmes d'information cliniques (SIC) sont, au mieux, des systèmes généralistes qui couvrent l'ensemble des besoins standards d'un hôpital : documentation clinique structurée, prescription informatisée et, pour les plus avancés, gestion de flux et processus de prise en charge. Très souvent, les besoins spécifiques à tel ou tel service ne sont que partiellement pris en compte. Des solutions originales et spécifiques existent souvent. Elles sont généralement dédiées à une certaine activité et ont été développées, pour la plupart, par des spécialistes du domaine considéré. Hautement adaptées aux besoins de ce domaine, ces applications sont en revanche démunies pour les autres besoins d'un SIC. Il y a alors souvent d'épiques débats sur l'intérêt d'acquérir ces solutions, débats alimentés par le désir des cliniciens d'utiliser un système qui épouse leurs besoins, et par les contraintes des spécialistes de l'informatique clinique ou hospitalière qui visent à uniformiser les plates-formes et à assurer l'interopérabilité des systèmes.

La notion d'interopérabilité est souvent comprise comme étant la capacité d'un hôpital à communiquer avec un réseau de soin communautaire, par exemple des médecins installés, des laboratoires d'analyses, des pharmacies, des assurances ou d'autres hôpitaux. Ceci est un élément important, mais trop limité. Le premier défi de l'interopérabilité est à l'intérieur même de l'institution, à savoir la capacité du système d'information clinique à communiquer avec l'ensemble des autres éléments du système d'information hospitalier. Ceci comprend bien évidemment toutes les autres parties directement liées au système

d'information clinique, classiquement le système de gestion administrative des patients (ADT), le système de gestions des laboratoires (LIS), de la radiologie (RIS-PACS), de la cardiologie (CIS) et des autres services cliniques. Cependant, de très nombreux autres systèmes doivent pouvoir être rendus interopérables dans une vision stratégique du système d'information : la gestion des ressources humaines, les menus, la gestion des stocks, la facturation, etc. La prescription d'un examen cardiologique et d'un médicament peuvent illustrer ce point. Prescrire un examen de cardiologie devrait pouvoir être connecté avec la disponibilité des ressources humaines, médecin cardiologue, technicien et infirmier, et des ressources médico-techniques, à savoir du plateau de cardiologie, de manière à ce que cette prescription puisse interagir immédiatement avec la planification de l'usage des ressources de cardiologie. De même, la prescription d'un médicament devrait être assujettie à la disponibilité de cette substance dans l'institution. L'administration du médicament devrait automatiquement se répercuter sur la gestion des stocks disponibles et permettre la planification des achats. Dans les deux cas, ces actions devraient se répercuter de manière automatique sur les coûts induits par la prise en charge du patient, y incluant la charge en ressources humaines que ces actes ont nécessitée.

Les deux exemples simples ci-dessus illustrent la complexité des interactions entre les divers éléments du système d'information hospitalier et du système d'information clinique. Une telle complexité ne peut pas être résolue sans une utilisation stricte de nombreux standards d'interopérabilité pour tous les canaux concernés. Ceci ne peut pas se résoudre avec un seul standard et une solution monolithique, mais requiert de nombreux liens entre des systèmes souvent très hétérogènes. De plus, les standards d'échanges du marché ne couvrent pas tous les besoins attendus par des solutions innovantes adaptées à l'expression des besoins des utilisateurs. Pour cette raison, les possibilités d'intégration entre systèmes sont plus grandes dans le cadre de développements internes même si l'usage des standards est favorisé. Car c'est au prix de grands efforts qu'un fournisseur accepte de déroger aux standards ou de développer une nouvelle interface pour améliorer les fonctionnalités de son progiciel. Par exemple l'usage d'une authentification unique et forte avec contrôle par carte à puce (*Single Sign-on*) par des progiciels est un challenge lors d'une nouvelle acquisition.

Il faut noter une complexité supplémentaire. En effet, il est nécessaire que les systèmes soient fortement communicants, tout en maintenant de très hauts niveaux de sécurité et de protection des données.

Enfin les progiciels pouvant fonctionner de manière autonome embarquent leurs propres données patient, séjour, structure et autres. Par conséquent, faire communiquer un système d'information clinique avec ceux-ci demandent le plus souvent à répliquer les données du SIC.

Ce rapport s'attache à illustrer la complexité des interfaces requises pour une intégration fine d'un système d'information dédié à la cardiologie (CIS) et dont l'acquisition a été faite après revue du marché.

2 Contexte

Les Hôpitaux Universitaires de Genève (HUG) regroupent les hôpitaux publics du Canton de Genève répartis sur quatre grands campus et sur de nombreux sites ambulatoires. Ils couvrent tous les types de soins hospitaliers, primaires, secondaires et tertiaires ainsi qu'ambulatoires. Les HUG ont un peu plus de 2 000 lits pour plus de 45 000 hospitalisations et 850 000 consultations par an.

Le système d'information clinique des HUG est une solution basée sur l'intégration de produits existants et le développement de solutions propres, en l'absence de solutions commerciales satisfaisantes. Ceci est le cas de la plate-forme de prescription. Les développements sont faits dans l'environnement Java, sur une architecture fondée sur des composants métiers, un middleware orienté messages, et l'utilisation massive de standards d'interopérabilité tels que HL7 et DICOM. Le dossier patient intégré (DPI), client opérationnel du SIC, est utilisé quotidiennement par plusieurs milliers d'utilisateurs. Chaque jour, ce sont plus de 65 000 images, plus de 2 500 documents, plus de 15 000 prescriptions qui sont enregistrés dans le système.

3 Intégration

On distingue plusieurs piliers de l'intégration :

3.1 Les flux entrants

3.1.1 Les droits d'accès

La gestion des droits d'accès se doit d'être centralisée, il est donc important que toute application tierce utilise un système unifié de gestion des droits d'accès. Dans le cadre de la mise en place du *single sign-on* dans le SIC des HUG, il a été défini trois niveaux d'interopérabilité des droits :

- Login/logout : il y a un droit à utiliser cette application.
- Profil : le profil (médecin, infirmière, etc.) est fourni par le système centralisé.
- Droits fins : en plus du profil, les droits fins associés à chaque profil sont fournis.

Ces informations sont mises à disposition par des services Web intégrant des mécanismes comme LDAP et des flux http/xml.

3.2 Les données administratives des patients

Il s'agit ici essentiellement de flux entrants de type HL7 et permettant de transférer les informations d'identification des patients et de leur trajectoire. Ceux-ci devront être conformes aux recommandations IHE (*Integrating the Healthcare Entreprise*) en matière de transmission des données démographiques. C'est-à-dire supporter les transactions '*Patient Registration*' et '*Patient Update*'. Soit l'envoi par protocole LLP (*Low Layer Protocol*) des messages HL7 de type ADT suivants : ADT^A01, ADT^A02, ADT^A03, ADT^A04, ADT^A05, ADT^A08, ADT^A11, ADT^A38, ADT^A40.

3.3 Les interfaces d'importation de l'agenda institutionnel

Il s'agit ici de permettre de recevoir l'ensemble de la planification, rendez-vous créés, modifiés et supprimés et devra être conforme au standard HL7 en matière de transmission des données de planification des examens. C'est-à-dire supporter la transaction '*Appointment notification*'. Soit l'envoi par protocole LLP des messages HL7 de type SIU suivants SIU^S12, SIU^S13, SIU^S15.

3.4 Les interfaces d'importation des données cliniques

Ce mécanisme sera basé sur l'échange de messages HL7 de type ORU par le protocole LLP.

Ces messages seront transmis par le DPI et contiendront par patient des valeurs biologiques ou observées. Notamment :

- CK Max. durant l'hospitalisation.
- Créatinine à l'admission.
- Bilan Lipidique durant les 24h qui suivent l'admission.
- HDL Cholestérol durant les 24h consécutives aux douleurs thoraciques.
- Allergies.
- Antécédents.
- Etc.

3.5 Les flux sortants

3.6 Création d'un patient provisoire dans le SIH

Dans certaines circonstances particulières, notamment lors d'une urgence, un utilisateur peut être amené à prendre en charge un patient alors que son identité n'a pas été préalablement créée dans le SIH de l'établissement. Dans ce cas de figure, il est nécessaire de pouvoir créer une identité provisoire dans le SIH. Le mécanisme de communication avec le système d'admission sera basé sur un service échangeant un flux HPPT/XML.

3.7 Traçabilité des fichiers de log

Le logiciel devra transmettre à destination du serveur de trace de l'institution les événements importants réalisés dans l'application. L'envoi d'une structure XML via le protocole *syslog* sera implémenté.

3.8 Interface d'exportation des données

Données non structurées

Le logiciel devra exporter les rapports d'examen au format PDF (*Portable Document Format*) après validation définitive de l'utilisateur. L'interface doit générer un message HL7 de type ORU et y associer le fichier PDF. Cependant les informations minimales à transmettre en plus du rapport d'examen sont :

- Identifiant permanent patient (IPP) .
- Identité du patient .
- Date de naissance du patient .
- Numéro de séjour (EdS) .
- Numéro de l'examen (identifiant unique) .
- Nom de l'examen .
- Date et heure de l'examen .
- Auteur du rapport .
- Médecin valideur du rapport .

Il devra intégrer un mécanisme permettant d'émettre à nouveau un rapport modifié après sa diffusion dans le DPI. Le numéro d'examen transmis devra être identique à celui de la première diffusion.

Données structurées

Le logiciel exportera une liste exhaustive de données quantitatives et qualitatives vers le SIC dans un format XML qui devra être précisé. Les données XML seront transmises au fur et à mesure par protocole http/xml suivant un délai après validation du dossier à définir.

À chaque changement majeur de version du logiciel, cette interface pourra être complétée par de nouvelles données pertinentes.

3.9 Module facturation et interface d'exportation

Le logiciel devra générer quotidiennement des messages HL7 de type DFT décrivant les éléments de facturation des examens réalisés contenant notamment :

- Détail du patient (IPP , identité, date de naissance, sexe, adresse) .
- N° d'épisode de soins .
- Détail des prestations réalisées (identifiant unique d'examen, date, médecin prestataire, matricule du médecin, n° unique de prestation, code facturation HUG, libellé de la prestation, attribut facturable ou non) .
- Détail des matériels utilisés (identifiant unique d'examen, date, médecin prestataire, matricule du médecin, n° unique de matériel, code facturation institutionnel, libellé du matériel, attribut facturable ou non) .

Pour cela, chaque prestation/matériel décrit dans le logiciel comportera les codes facturation institutionnels correspondants ainsi qu'un attribut déterminant si la prestation est facturable. La mise à jour de ces tables pourra être réalisée par l'administrateur du logiciel sans l'intervention de l'éditeur. Le protocole de transfert de ces messages HL7 pourra être effectué par FTP (*file transfer protocol*) ou par LLP .

3.10 Gestion du matériel et interface d'exportation dans le système de gestion de stock

Le matériel utilisé en salle devra être connu dans le logiciel par l'intermédiaire d'un lecteur code-barres, éventuellement d'une saisie manuelle. Le matériel sera :

- décrémenté du stock lors de son utilisation ;
- incrémenté du stock lors de sa livraison ;

Le logiciel devra offrir une vision complète sur le suivi et la traçabilité du stock. Les mouvements de stock devront être exportables. Le protocole et le format d'exportation devront être http/xml.

3.11 Interface avec les équipements - DICOM Worklist

Le logiciel devra répondre aux sollicitations des équipements DICOM (Digital Imaging and COmmunications in Medicine) et retourner le détail des examens en attente de réalisation (patient présent) dans la salle abritant cet équipement. En particulier :

- Identité du patient .
- Date de naissance du patient .
- Sexe du patient .
- IPP du patient .
- Accession Number .
- Identité du cardiologue .
- Procédure .
- Study Instance UID .

Il devra être conforme aux recommandations IHE en matière de réponse aux équipements. C'est-à-dire supporter la transaction *'Query Modality Worklist'*.

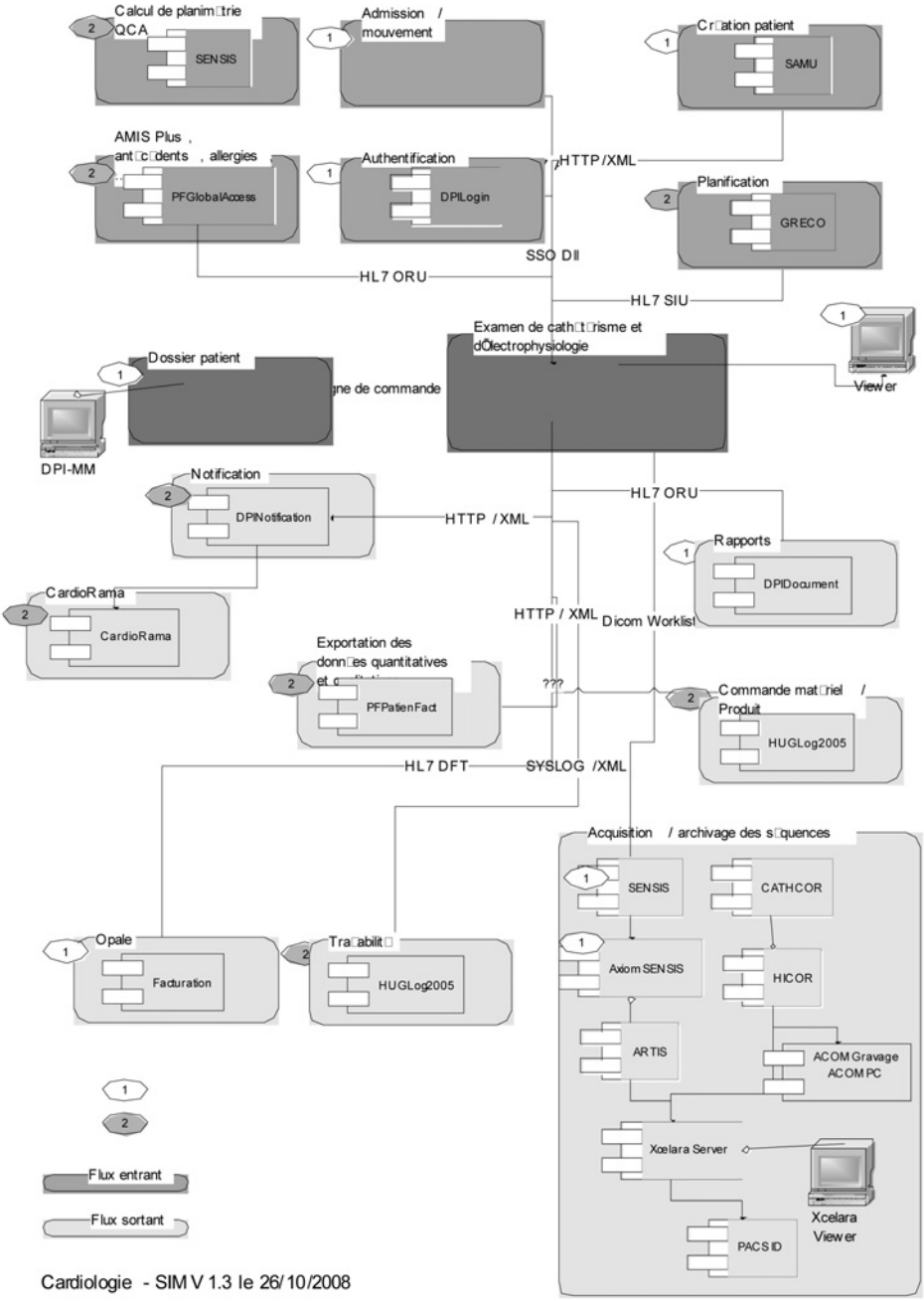


Figure 1 - Schéma d'intégration du progiciel de gestion médico-technique de la cardiologie.

3.12 Interface avec le moteur de notification

Afin d'offrir dans les unités de soins un tableau de bord présentant le planning et l'état de la prise en charge des patients dans les salles, il est demandé au logiciel de notifier les changements d'état des examens. À savoir par exemple :

- Transport *aller* demandé .
- Examen en cours .
- Examen terminé .
- Transport *retour* demandé .
- Rapport diffusé .

Ces notifications seront transmises par protocole http et au format XML.

Le graphique de la figure 1 résume l'ensemble des interfaces requises.

4 Conclusion

Cet article fait un survol rapide de la complexité de l'intégration fine d'un logiciel tiers qui a été acquis dans le cadre de l'informatisation spécifique de l'activité en cardiologie. Un prérequis étant une forte intégration et interopérabilité avec le système d'information clinique en place. Ce rapport démontre le nombre important d'interfaces requises. Il souligne également la nécessité d'avoir un système d'information clinique très ouvert basé sur une architecture orientée services (SOA) pour permettre une telle intégration. En effet l'architecture SOA permet de disposer de services ou composants autonomes, non dépendants de services externes et caractérisés par une unité de traitement :

- À couplage faible réduisant les dépendances.
- À instance unique.
- Localisable.
- Interopérable par la standardisation de sa communication.
- Plate-forme indépendant.
- Atomique, c'est-à-dire couvrant un périmètre fonctionnel restreint.

Abréviations

ADT	*Admission, Discharge, Transfer*
CIS	*Cardiology Information System*
DFT	*Detail Financial Transactions*
DICOM	*Digital Imaging and COmmunications in Medicine*
DPI	*Dossier patient intégré*
FTP	*File Transfer Protocol*
HUG	Hopitaux Universitaires de Genève
EDS	Épisode de soins
HL7	*Health Level 7*
IHE	*Integrating the Healthcare Entreprise*
IPP	Identifiant Permanent du Patient
LDAP	*Lightweight Directory Access Protocol*
LLP	*Low Layer Protocol*
LIS	*Laboratory Information System*
ORU	*Observ result/unsolicited*
PACS	*Picture Archiving and Communication System*

PDF	*Portable Document Format*
RIS	*Radiology Information System*
SIC	Système d'Information Clinique
SIH	Système d'information hospitalier
SIU	*Scheduling Information Unsolicited*
SOA	*Service Oriented Architecture*
XML	*eXtensible Markup Language*

Références

[1] Mildenberrger P., Eichelberg M., Martin E, Introduction to the DICOM Standard, *Eur. Radiology* 2002; 12: 920-927.
[2] Mildenberger P, Wein B, Bursig HP, Eichelberg M. Current developments in DICOM and IHE, *Radiologie* 2005 ; 45(8): 682-9, Article en Allemand .
[3] Boochever SS., HIS/RIS/PACS integration: getting to the gold standard. Radiol. Manage 2004; 26(3): 16-24 .
[4] Bernardini A, Alonzi M, Campioni P, Vecchioli A, Marano P., IHE: integrating the healthcare enterprise, towards complete integration of healthcare information systems. *Rays* 2003; 28(1): 83-93.
[5] Integrating the Healthcare Enterprise (IHE). http://www.ihe.net .

Adresse de correspondance
Patrice Trolliard
Hôpitaux Universitaires de Genève
Service d'informatique médicale, unité d'imagerie numérique
24, rue Micheli-du-Crest
1211 Genève 4, Suisse
E-mail : patrice.trolliard@hcuge.ch

Informatique et Santé, 2009 (17) : 23-34
© Springer – Verlag France

Une procédure d'anonymisation à deux niveaux pour créer un corpus de comptes rendus hospitaliers

Cyril Grouin[1], Arnaud Rosier[2], Olivier Dameron[2], Pierre Zweigenbaum[1]

[1]*LIMSI, CNRS, F-91403 Orsay, France*
[2]*INSERM U936, F-35000 Rennes, France*

Abstract

De-identification is a growing need in medical informatics, and has therefore recently been the subject of renewed interest. De-identification needs to be tuned to the local documents and their specificities, which requires language engineers to work on non-de-identified text. To lower the issues linked to such a situation, we propose a de-identification method which proceeds in two steps. We report experiments on the adaptation of an American de-identifier to French and on the development of a new de-identifier for French patient reports. The latter, evaluated on a set of 23 randomly selected texts, obtains 85 % recall and 91 % precision.

Keywords

Natural Language Processing; Anonymization.

1 Introduction

Dans les travaux qui portent sur l'analyse automatique de textes en langue naturelle, les corpus de textes sont un matériau fondamental. Cette observation s'applique au traitement automatique de textes du domaine médical ou biomédical, qu'il s'agisse d'indexation, de codage automatique [1], de catégorisation de patients[1], de repérage de gènes, de protéines et de leurs interactions [2]. Un corpus de textes est utile pour étudier les problèmes et mettre au point le système d'analyse. Un corpus de textes dans lequel on a ajouté des annotations qui indiquent les résultats attendus [2] permet de plus d'évaluer les résultats d'un programme en cours de développement. Il permet enfin, s'il est suffisamment grand, d'entraîner automatiquement un système fondé sur des mécanismes d'apprentissage.

Le travail sur des textes cliniques pose cependant un problème : ces textes ne peuvent être utilisés en dehors du soin des patients que si toutes les marques permettant d'identifier le patient ont été supprimées. Ils doivent donc être *anonymisés* avant d'être mis entre les mains de chercheurs et de développeurs de méthodes de traitement automatique des langues, ou pour être inclus dans une publication. Cela crée une contrainte forte qui cause une extrême rareté des corpus disponibles dans le domaine clinique, contrairement par exemple aux corpus de résumés d'articles scientifiques qui sont utilisés dans le domaine biomédical (génomique). Zweigenbaum [3] cite cette contrainte comme l'un des facteurs clés dans le différentiel de développement des travaux de traitement automatique des langues en informatique médicale par rapport à ceux effectués en bioinformatique. Ces besoins ont motivé des travaux sur l'anonymisation automatique de comptes rendus cliniques, que nous décrivons plus bas.

---

[1] Voir par exemple la catégorisation du statut d'obésité d'un patient (https://www.i2b2.org/NLP/).

Après une brève revue de travaux proches (section 2) et de notre matériel de départ (section 3), nous présentons ici un travail d'anonymisation en trois points (section 4). D'une part, le travail de mise au point d'une méthode d'anonymisation nécessite lui-même un corpus de développement dont la disponibilité hors d'un service hospitalier pose elle-même le problème de son anonymisation. Nous réduisons cette contrainte en effectuant le plus gros de l'anonymisation à la source, dans le service hospitalier. D'autre part, nous rapportons une expérience de réutilisation d'un anonymiseur réalisée aux États-Unis pour traiter les informations identifiantes [4], dont nous avons converti en français les termes liés à la langue anglaise. Enfin, nous avons développé un anonymiseur directement conçu pour traiter des textes médicaux français, que nous évaluons sur un échantillon de documents de notre corpus. Nous présentons les résultats de son évaluation (section 5), les discutons et concluons sur quelques perspectives (section 6)[2].

2 Contexte et état de l'art

Des travaux sur l'anonymisation automatique de comptes rendus cliniques ont été menés dans différents contextes [5, 6]. Ils ont été récemment dynamisés aux États-Unis avec la problématique des « données de santé protégées » (*Protected Health Information*, PHI). Les États-Unis se sont en effet dotés d'une liste officielle de 18 types d'identifiants dont la suppression dans un document contenant des PHI rend possible la levée des contraintes pesant sur la transmission de ces données et donc de ce document (par exemple à des fins de recherche)[3] : noms et prénoms, toute indication de lieux, dates, âge si supérieur à 90 ans, numéros de téléphone et de télécopie, adresses de messageries électroniques, numéros de Sécurité Sociale, numéro d'enregistrement médical, numéro de complémentaire santé, numéro de compte, numéro de cartes d'identité et de permis de conduire, références sur le véhicule du patient, URLs, adresses IP, numéro de série ou identifiant d'appareil implanté (pacemaker), identifiants biométriques, et n'importe quel autre code ou caractéristique permettant d'identifier un patient tel qu'une photographie du visage, de cicatrices ou de tatouages spécifiques) .

Cette liste a facilité le développement de travaux d'anonymisation aux États-Unis. En revanche il n'existe pas à notre connaissance en France de critères clairs définissant les informations à supprimer d'un texte pour qu'il soit considéré comme anonymisé. En l'absence de ces critères, nous reprenons les principaux parmi ceux employés aux États-Unis, qui sont par ailleurs en cohérence avec des critères retenus après discussion avec la CNIL dans un projet passé [7] : noms, prénoms, lieux, dates et âges.

L'intérêt du partage de données médicales est par ailleurs fortement soutenu par diverses instances publiques, à l'image de l'Institut Canadien de Recherches Médicales (CIHR) qui rend obligatoire, depuis le 1er janvier 2008, la mise à disposition des données utilisées dans le cadre de publications [8]. À cet effet, l'anonymisation des informations médicales personnelles n'en est que plus pertinente.

Une compétition portant sur l'anonymisation de 220 comptes rendus d'hospitalisation a été organisée en 2007 par le projet i2b2[4] [9]. La quasi-totalité des systèmes était fondée sur des méthodes de catégorisation automatique (SVM, arbre de décision, etc.) ou d'étiquetage de séquence (HMM, CRF), travaillant sur des textes généralement pré-analysés à l'aide

2 Ce travail a été partiellement financé par le projet Akenaton (ANR-07-TECSAN-001-06).

3 Office for Civil Rights — HIPAA (http://www.hhs.gov/ocr/hipaa/understanding/summary/) (site visité le 21/1/09).

4 Informatics for Integrating Biology & the Bedside, http://www.i2b2.org/ (site visité le 21/11/08).

d'outils de traitement automatique des langues (étiquetage morpho-syntaxique, etc). Les meilleurs systèmes ont obtenu une F-mesure supérieure à 98 %, avec une précision un peu supérieure et un rappel un peu inférieur[5].

Neamatullah *et al.* [4] remarquent que ces résultats ont été obtenus par apprentissage sur un grand corpus étiqueté à la main au préalable, et qu'il n'est pas évident que cette approche soit transportable à d'autres types de textes. Leur logiciel DE-ID (pour « de-identification » ; voir section 3.2) repère les entités à anonymiser à l'aide de dictionnaires et d'expressions régulières. Évalué sur un corpus de notes infirmières, considéré comme plus difficile que les comptes rendus d'hospitalisation d'i2b2, il obtient un rappel de 96,7 % et une précision de 74,9 %.

Friedlin et McDonald [10] s'attaquent à des sources de textes plus variées provenant de messages HL7 transmis au *Regenstrief Medical Record System*. Leur système, MEDS[6], emploie également des listes et des expressions régulières. Son évaluation finale sur 7 193 comptes rendus d'anatomopathologie a obtenu un rappel de 99,47 % sur les noms propres et de 96,93 % sur les autres éléments (sa précision sur 2 400 documents variés était de 92 %). Les auteurs affirment que les résultats obtenus par leur système sont similaires à ceux des meilleurs systèmes de la campagne i2b2.

Les listes de MEDS sont constituées à partir de l'UMLS pour les noms propres et noms d'hôpitaux. Les expressions régulières sont employées pour les entités numériques. Ses auteurs ont également considéré qu'il était possible d'améliorer l'anonymisation en ayant recourt à l'étiqueteur/lemmatiseur de MetaMap [11], traitement qui se révèle efficace pour traiter les mots qui appartiennent aux listes de noms propres mais qui, dans le contexte où ils apparaissent, ne sont pas des noms propres (notamment dans le cas d'un mot écrit en minuscules et précédé d'un article ou d'un adjectif : *I examined the pat*, où *pat* désigne le patient mais pourrait être pris pour le prénom *Pat*). Enfin, un dernier module étudie le voisinage des noms propres connus de manière à traiter les éventuelles fautes de frappe.

L'anonymiseur le plus connu pour le français [5] n'étant pas de source libre, donc pas adaptable par un tiers, nous nous sommes orientés vers un anonymiseur disponible au téléchargement, en l'occurrence DE-ID. Il était de plus intéressant d'examiner dans quelle mesure un anonymiseur conçu pour l'anglais était adaptable au français. Lorsqu'ensuite nous avons été amenés à concevoir notre propre anonymiseur, nous nous sommes inspirés de DE-ID et de MEDS.

3 Matériel

Nous décrivons ici nos données initiales : un corpus de comptes rendus collecté pour le projet Akenaton et l'anonymiseur DE-ID [4] introduit plus haut.

3.1 Corpus de comptes rendus

L'objectif étant de constituer le corpus le plus large possible, nous avons décidé de cibler un ensemble varié d'unités fonctionnelles (UF) appartenant au centre cardio-pneumologique d'un centre hospitalier régional universitaire. Ces UF comprennent des services d'hospitalisation, des hôpitaux de jour, des services de consultation, une unité d'accueil d'urgence spécialisée ainsi que des services médico-techniques et chirurgicaux,

---

[5] Le *rappel*, ou sensibilité, est la proportion des entités à anonymiser effectivement repérées par le système ; et la *précision*, ou valeur prédictive positive, est la proportion des entités repérées par le système qui devaient effectivement être anonymisées. La *F-mesure* est la moyenne harmonique du rappel et de la précision.

[6] MEDS : *Medical Deidentification System*.

et couvrent donc toute une gamme de prise en charge diagnostique et thérapeutique ainsi que des modes de suivi de patients variés. Trois spécialités étaient concernées : la cardiologie, la chirurgie cardiaque et la pneumologie. Ces spécialités traitent de pathologies thoraciques assez souvent liées sur le plan médical, ce qui signifie une cohérence relative des problèmes médicaux principaux dont il est question dans les comptes rendus créés.

Pour l'ensemble de ces unités, l'intégralité des comptes rendus (hospitalisation et opératoires) et courriers créés entre 2002 et début 2008 a été extraite du système d'information hospitalier, à l'aide d'une requête de la base de données correspondante. Pour chacun des documents, un fichier contenant le texte a été généré, ainsi qu'un fichier associé, au format XML, qui contient sous forme structurée les métadonnées du système d'information hospitalier au sujet du patient : données d'identification et données de provenance du document. Ces métadonnées seront utiles pour l'anonymisation à la source de ces documents (voir la section 4.1), sous le contrôle coordonné de médecins du centre cardio-pneumologique et du département d'informatique médicale.

3.2 L'anonymiseur DE-ID

L'anonymiseur DE-ID [4], réalisé par plusieurs équipes du MIT[7], se compose d'un ensemble de scripts Perl. Les types d'informations anonymisées par DE-ID reposent sur les 18 catégories spécifiques d'informations définies par l'HIPAA.

Les méthodes utilisées par DE-ID reposent sur une combinaison de moyens. En premier lieu, le logiciel mobilise des ressources linguistiques de deux types : des *dictionnaires* (noms communs et expressions médicales) et des *listes d'entités nommées* (noms, prénoms, hôpitaux, villes, abréviations d'États américains, indicatifs téléphoniques, pays, gentilés, etc). Pour un même type d'entité nommée, une distinction est faite entre les entités ambiguës et les entités non ambiguës (à ce titre, le prénom féminin « France » figurera dans la liste des prénoms féminins ambigus, puisqu'il s'agit également d'un nom de pays).

En second lieu, le logiciel exploite des *listes de déclencheurs* qui permettent de détecter des informations à anonymiser dans le voisinage – les quelques mots qui précèdent et qui suivent – de ces déclencheurs. Ces déclencheurs sont typés pour chaque catégorie d'entité nommée : une liste de particules (*De, Mc, Van...*) et de titres (*Mister, Doctor, Professor...*) pour les noms de famille, des éléments précédant un lieu (*lives in, resident of, comes from*), etc.

Enfin, le logiciel utilise un ensemble d'*expressions régulières* reposant sur une précédente identification de termes du voisinage (un mot sera anonymisé comme nom de famille s'il ne figure pas dans le dictionnaire des noms communs et que le mot qui le précède a été identifié comme un prénom de manière sûre).

En dehors de la trace habituelle des anonymisations réalisées, DE-ID attribue un identifiant numérique unique à chaque nom rencontré dans un texte et affiche cet identifiant dans la version anonymisée. Ceci permet ainsi de conserver une information de co-référence – le suivi de la réapparition des différents noms en corpus – sans sacrifier à l'impératif de l'anonymisation des informations personnelles.

Un autre aspect de la conservation des informations contenues dans les textes tout en respectant le devoir de protection des données consiste à associer à chaque document un

[7] *Laboratory for Computational Physiology, Harvard-MIT division of Health Sciences & Technology, Computer Science and Artificial Intelligence Laboratory*

« décalage dans le temps » distinct à chaque document mais qui permet de conserver la vraisemblance des intervalles de temps exprimés dans les différentes dates d'un document.

Les résultats obtenus par DE-ID sur le corpus d'évaluation de Neamatullah *et al.* [4][8] varient selon le type d'entité anonymisé : l'anonymisation des noms de cliniciens obtient ainsi un rappel de 99,5 % et une précision de 72,5 % contre un rappel de 76,1 % et une précision de 71,3 % pour les dates avec années. Globalement, le système DE-ID obtient un rappel de 96,7 % et une précision de 74,9 %. Les auteurs mettent également en avant un taux de faux positifs assez faible (de l'ordre de 19,72 faux positifs pour 100 000 mots), avec un taux de faux positifs étrangement plus élevé pour les dates (7,769 pour 100 000 mots) que pour les noms (1,494 pour 100 000 mots).

4 Méthodes

4.1 Anonymisation à la source

Comment mettre au point des méthodes d'anonymisation poussée sur des textes réels sans pour autant travailler sur des données fortement identifiantes ? Nous n'avons pas trouvé de mention de ce problème dans la littérature sur l'anonymisation de textes cliniques. La solution que nous avons mise en place consiste à effectuer une première passe d'anonymisation reposant sur des méthodes simples, qui peuvent être mises en œuvre à la source lors de la collecte des textes, mais permettant de supprimer la quasi-totalité des informations identifiantes fondamentales : le nom, le prénom et la date de naissance du patient.

Ces informations figurent en effet dans la partie structurée du dossier patient, et ont été extraites dans le fichier de métadonnées associé à chaque document textuel. Ces métadonnées contiennent en tout :

- données patient : nom, nom marital, prénom, date de naissance et sexe du patient ;
- données de provenance du document : identifiant de l'UF, date du compte rendu, nom de l'auteur.

L'algorithme d'anonymisation à la source a été conçu pour réaliser les opérations suivantes, au sein de l'entrepôt de données hospitalier :

- remplacer le texte mentionnant le nom du patient (quelle que soit la casse) par une balise <Nom patient> ;
- remplacer de façon similaire le texte mentionnant le nom marital et le prénom du patient ;
- remplacer la date de naissance par une balise <Date naissance patient>, d'après une identification basée sur trois types de motifs : 01/01/70, 01/01/1970 et 01 janvier 1970.

Pour chaque document, le nombre de remplacements a été comptabilisé et les documents pour lesquels aucune altération du nom ou du nom marital n'était réalisée ont été identifiés et vérifiés manuellement.

Enfin, dans les fichiers de métadonnées, les noms, prénom et date de naissance du patient et le nom du médecin ont été chiffrés selon l'algorithme SHA-256, et les autres informations supprimées. Ces fichiers, fournis avec le corpus, conservent donc une

[8] Ce corpus comprend un ensemble de notes de soins infirmiers anonymisées qui ont été réidentifiées de manière plausible ; ce corpus est utilisé pour le développement et l'évaluation de DE-ID.

information permettant de mettre en évidence les documents parlant d'un même patient, mais sans possibilité de remonter à l'identité du patient. Cette méthode est similaire à celle employée par Quantin *et al.* [12] pour créer des identifiants familiaux chiffrés.

L'ensemble des documents (texte + métadonnées résiduelles codées) constitue le corpus à l'issue de cette première passe d'anonymisation à la source. Chaque texte va ensuite faire l'objet de traitements plus poussés pour prolonger son anonymisation.

Un échantillon de 23 textes a été aléatoirement extrait du corpus à des fins d'évaluation. Chacun de ces textes a été manuellement anonymisé pour servir de référence (*gold standard*) : les expressions de type nom, prénom, date et ville y sont respectivement remplacées par les balises XML <nom />, <prenom />, <date /> et <ville />.

4.2 Conversion partielle en français de l'anonymiseur DE-ID

Devant la qualité des résultats obtenus par le logiciel DE-ID et du fait de sa distribution sous licence GNU autorisant librement sa réutilisation, nous avons essayé de l'adapter au français. Son adaptation à une nouvelle langue passe par plusieurs modifications de difficultés inégales portant sur chacune des trois ressources qu'il utilise.

4.3 Dictionnaires et listes

La qualité des anonymisations dépendant pour une part non négligeable de la richesse des dictionnaires et listes utilisées, une attention particulière doit être accordée aux ressources linguistiques utilisées.

Nous avons d'une part employé des lexiques disponibles sur le site de l'Association des Bibliophiles Universels (ABU), une association de bénévoles qui proposent en libre téléchargement depuis son site Internet[9] des listes d'entités nommées et de dictionnaires. Toutes les listes proposées ne sont pour autant pas d'égale qualité ; si le dictionnaire des noms communs se révèle riche et varié (plus de 251 000 formes différentes), la liste des villes françaises est désaccentuée, tout comme la liste des prénoms qui, de plus, intègre une majorité de prénoms anglo-saxons. Il s'agit cependant de ressources utiles qu'il est possible d'adapter (en tentant une réaccentuation automatique), ou pour lesquelles il est peut être envisagé de contraindre les outils de traitement automatique des langues à s'adapter (en étant insensibles aux caractères accentués).

Nous avons par ailleurs complété cet ensemble de listes par nos propres ressources, soit en reprenant des listes précédemment constituées (liste de 33 380 noms de villes), soit en créant de nouvelles ressources (liste de 2 526 noms d'hôpitaux, cliniques, centres hospitaliers français, constituée à partir d'un annuaire de santé sur Internet[10]).

4.4 Listes de déclencheurs

Adapter les listes de déclencheurs au français s'est révélé être la tâche la plus facile dans le sens où il s'agit d'une traduction des termes avec une éventuelle complétion des listes utilisées[11]. Le logiciel DE-ID reposant sur des listes de déclencheurs distinctes selon que le déclencheur est attendu avant ou après le mot à anonymiser, la contrainte de l'ordre des mots qui diffère en anglais et en français a ainsi pu être aisément contournée.

---

9 http://abu.cnam.fr/(site visité le 21/11/08).

10 http://www.sanitaire-social.com/ (site visité le 14/02/08).

11 En raison de l'ordre des mots différents en français et en anglais, certains déclencheurs présents dans la liste des suffixes en anglais ont été déplacés dans la liste des préfixes en français : les suffixes anglais *Street, Road, Blvd* deviennent les préfixes *Rue, Route, Bld* en français.

4.5 Expressions régulières

Nous avons constaté que modifier les expressions régulières constituait la tâche la plus difficile dans ce travail d'adaptation du logiciel au français. En effet, cette adaptation dépend d'une part de la manière dont le logiciel a été conçu (du point de vue de la manière de coder), et d'autre part de l'objectif visé (en l'occurrence, convertir le logiciel pour une utilisation en français, sur un corpus de comptes rendus médicaux).

Si l'adaptation des expressions régulières pour les entités de type numérique (numéros de Sécurité Sociale, téléphone, dates, etc.) ne pose aucun problème particulier, il en est tout autrement des détections d'entités nommées. La réalisation d'expressions régulières étant fortement contextuelle, il ne nous a guère été possible d'adapter les expressions existantes sans les reprendre à la base, ce qui est très chronophage.

4.6 Conception d'un nouvel anonymiseur : Medina

Devant les difficultés rencontrées pour adapter le logiciel DE-ID au français, nous avons fait le choix de développer notre propre anonymiseur pour le français : Medina (*Medical Information Anonymization*). Nous nous sommes focalisés pour commencer sur les anonymisations de noms, prénoms, dates et villes.

Étant donné l'expérience acquise par les auteurs des logiciels DE-ID et MEDS, nous avons fait le choix de combiner plusieurs méthodes : les expressions régulières pour les entités numériques, les dictionnaires et listes pour les entités nommées, ainsi qu'une seconde anonymisation fondée sur le voisinage des termes déjà anonymisés. Dans la mesure du possible, nous avons essayé de distinguer le programme des ressources utilisées (les ressources sont déportées dans des fichiers distincts). Nous avons également cherché à simplifier la configuration du logiciel en utilisant un fichier de configuration externe au programme.

Nous avons repris les listes et dictionnaires déjà réalisés précédemment lors de la tentative de francisation du logiciel DE-ID. Nous nous sommes par ailleurs inspirés des distinctions faites entre listes ambiguës et non ambiguës de DE-ID pour nettoyer les listes de noms, prénoms et pays. À cet effet, nous avons extrait de ces trois listes tous les termes qui existent dans le dictionnaire de noms communs[12], de manière à réduire le nombre de suranonymisations futures. Les noms ainsi supprimés de ces différentes listes — en particulier pour les prénoms — se révèlent d'autant moins ambigus que les listes utilisées contenaient une majorité de noms anglo-saxons qui semblent assez étranges à porter en France (*Agace, Dragon, Masculine, Travers,* etc).

Ce programme a été réalisé pour un type de corpus particulier et entraîné sur ce corpus. Parmi les limites de cet exercice, nous pouvons mettre en évidence le fait que nous avons utilisé les données médicales provenant d'un seul hôpital. Nous devrions toutefois être en mesure d'obtenir prochainement des comptes rendus opératoires en provenance d'un second hôpital, de telle sorte que nous pourrons tester la généricité du programme sur plusieurs corpus.

4.7 Mode d'évaluation

L'évaluation est réalisée par comparaison des balises attendues avec celles effectivement produites. Il s'agit ici de donner une représentation robuste aux décalages de l'occurrence mais aussi de la position de chaque balise d'anonymisation insérée dans un texte. Pour

---

[12] Nous avons réduit la liste de noms de familles de 3,47 % à 12 994 noms, celles des prénoms l'a été de 5,4 % à 11 746 prénoms, enfin, celle des noms de villes l'a été de 2,8 % à 32 439 noms.

cela, nous relevons dans le fichier de référence chaque balise avec son contexte immédiat gauche et droit de manière à inscrire la balise dans un contexte discriminant. Nous avons limité le contexte aux quatre caractères qui précèdent et suivent la balise car ils suffisent à localiser la séquence dans le texte[13]. Dans la phrase « *Je vous laisse le soin de demander à Monsieur <nom /> de bien vouloir prendre un rendez-vous* », nous relèverons ainsi la séquence « *eur <nom /> de* », ce qui nous permet de localiser la balise <nom /> dans le texte avec son contexte gauche et droit. À l'inverse, prendre plus de caractères contextuels n'a pas permis de mettre en évidence des cas d'ambiguïté. En revanche, si nous réduisons le contexte aux deux caractères qui précèdent et qui suivent, des cas de séquences ambiguës apparaissent[14] — augmentant artificiellement le nombre de balisages corrects — et l'évaluation s'en trouve alors faussée.

Nous relevons également les balises dans le fichier de résultats de l'anonymiseur. Puis nous avons comparé les balises produites avec celles de la référence et avons évalué les résultats produits en termes de rappel (nombre de balisages corrects rapporté au nombre de balisages attendus) et de précision (nombre de balisages corrects rapporté au nombre de balisages produits).

Parmi les limites de ce type d'évaluation, nous pouvons distinguer deux cas. En premier lieu, cette évaluation repose sur une identité stricte entre balises. Dans le cas d'un nom composé, si la référence substitue le nom composé par une balise alors que le système d'anonymisation retourne deux balises (une pour chaque nom), l'évaluation considérera qu'il y a échec du balisage quand bien même l'anonymisation a été produite. La seconde limite se rapporte au typage de l'information anonymisée. Pour une information donnée (typiquement un nom de famille qui relève d'un prénom[15]), la référence typera cette information en qualité de nom alors que l'anonymiseur pourra la typer comme un prénom. Face à cette différence de typage de l'information, l'évaluation considérera qu'il y a échec du balisage alors même que l'anonymisation aura été produite.

5 Résultats

Le corpus constitué comprend 21 749 documents uniques, concernant 11 964 patients individuels. Comme indiqué plus haut, un échantillon de 23 textes a servi à nos évaluations. Nous distinguons dans le tableau 1 l'évaluation de premier niveau (celle réalisée pour pouvoir sortir ces données de leur hôpital d'origine) de celles de second niveau (DE-ID et Medina), ces dernières poursuivant l'anonymisation sur les fichiers semi-anonymisés de premier niveau.

Tableau 1 : Comparaison des balisages des différents outils d'anonymisation.

	Rappel	**Précision**	**Corrects**	**Ramenés**	**Attendus**
Premier niveau	0,91	1,00	73	73	80
DE-ID francisé	0,63	0,25	103	402	163
Medina	0,85	0,91	139	152	163

L'anonymisation de premier niveau obtient une excellente précision et un bon rappel. Notre francisation de DE-ID obtient un rappel moyen et une précision très faible. Le

[13] Trois caractères semblent suffire mais, pour plus de sûreté, nous avons porté à quatre caractères ce contexte.

[14] Les deux caractères renvoient par exemple à la dernière lettre du mot précédent, suivie de l'espace, ce qui se révèle trop imprécis donc ambigu.

[15] Laurent, Martin, etc.

programme se déclenche dans un grand nombre de cas, ce qui cause mécaniquement un bruit important (suranonymisation). Medina obtient de meilleurs résultats que la version francisée de DE-ID avec de bons taux de rappel et précision.

À titre d'illustration du fonctionnement de Medina et des types d'erreurs commises, le tableau 2 présente quelques exemples d'anonymisations réalisées. Les expressions marquées *Entrée* sont le résultat de l'anonymisation de premier niveau, celles marquées *Sortie* sont le résultat de l'application de Medina sur cet état initial.

Tableau 2 : Exemples de résultats de Medina. Note : les informations identifiantes de ces exemples ont été modifiées manuellement avant d'être copiées dans le tableau.

Étape	Énoncé
Anonymisation réussie	
Entrée	*J'ai examiné en consultation Madame <Nom marital patient> Michèle, née le 13.1.1943, âgée de 62 ans, pour le contrôle annuel de son stimulateur double chambre.*
Sortie	*J'ai examiné en consultation Madame <nom /> <prenom />, née le <date />, âgée de <age />, pour le contrôle annuel de son stimulateur double chambre.*
Sur-anonymisation	
Entrée	*(le pace maker est actuellement réglé en VVI à une fréquence de 52/min).*
Sortie	*(le <prenom /> maker est actuellement réglé en VVI à une fréquence de 52/min).*
Sur- (*Pace Maker*) et sous- (*PECRESSE*) anonymisation	
Entrée	*Le Pace Maker avait été contrôlé au mois de janvier par le PR PECRESSE et ne montrait pas de signe d'usure.*
Sortie	*Le <prenom /> <nom /> avait été contrôlé au mois de <date /> par le PR PECRESSE et ne montrait pas de signe d'usure.*
Autres anonymisations réussies, hors évaluation	
Entrée	*un épisode d'IVG justifiant d'un traitement par Lasilix et la réduction de la posologie de Soprol à 1/jour.*
Sortie	*un épisode d'IVG justifiant d'un traitement par <medicament /> et la réduction de la posologie de <medicament /> à 1/jour*

6 Discussion et conclusion

La qualité de l'anonymisation à la source se révèle excellente en précision et son rappel est bon (sous-anonymisation). Nous sommes partis de cette version anonymisée pour constituer la référence ; à ce titre, il ne nous a guère été possible de mettre en évidence des sur-anonymisations. En revanche, nous avons pu pointer les éléments n'ayant pas été anonymisés. En pratique, un examen des balises nom et prénom montre une anonymisation totale des noms de famille et quasi totale des prénoms. Cela confirme dans notre contexte la pertinence de la démarche d'anonymisation initiale qui assure déjà un bon niveau d'anonymisation. De plus, les prénoms ratés au premier niveau comportent des lettres accentuées, signe d'un problème sans doute simple à résoudre.

Deux éléments d'appréciation doivent être pris en compte dans le commentaire des résultats de notre portage partiel de DE-ID au français. En premier lieu, nous n'avons pas été en mesure de produire autant de ressources linguistiques que celles existant en anglais (nous n'avons pas distingué les prénoms masculins et féminins ni même effectué une distinction entre prénoms ambigus et non ambigus ; il en est de même pour les noms de

famille, d'hôpitaux, de lieux, etc). Pour les ressources qui nous faisaient défaut en français, nous avons repris celles existant en anglais s'il s'agit de ressources potentiellement existantes en français (par exemple des listes de noms et prénoms complémentaires). Nous avons abandonné les ressources anglaises le cas échéant. En second lieu, nous n'avons pas été en mesure de modifier efficacement les expressions régulières utilisées en anglais, ce qui se révèle source de nombreuses suranonymisations. Si les premiers points pourraient être améliorés à l'aide de ressources complémentaires, ce dernier point est plus bloquant : il demande de revoir le cœur du repérage d'expressions du programme. Cela justifie à notre avis le choix de reconstruire un nouvel anonymiseur, choix globalement plus efficace.

Pour des raisons de temps, seul un petit échantillon de notre corpus a été anonymisé manuellement. Un échantillon plus grand est en cours de constitution. En tout état de cause, dans la mesure où notre échantillon a été tiré aléatoirement et où les résultats obtenus sont tranchés, nous nous attendons à ce que l'échantillon plus large confirme nos premières mesures.

Avec un investissement jusqu'ici modéré dans son développement, Medina obtient des résultats honorables comme anonymiseur de second niveau. Ici aussi, l'emploi de ressources complémentaires devrait permettre d'augmenter le rappel. Par ailleurs, l'utilisation de composants de traitement automatique des langues (étiqueteur morphosyntaxique, boîte à outils de gestion de transducteurs à états finis, analyseur syntaxique, etc.) pourrait aider à améliorer la précision. Notons cependant que l'usage de tels outils demande généralement une adaptation[16].

Références

[1] Pestian JP, Brew C, Matykiewicz P, *et al.* A shared task involving multi-label classification of clinical free text. In: *Proceedings of BioNLP 2007*. Association for Computational Linguistics, 2007.

[2] Kim JD, Ohta T, Tateisi Y, and Tsujii J. Genia corpus-a semantically annotated corpus for bio-textmining. *Bioinformatics* 2003; 19(Suppl. 1): 180 ;2.

[3] Zweigenbaum P. Natural Language Processing in the medical and biomedical domains: a parallel perspective. In: Rebholz-Schuhmann D, Salakoski T, and Pyysalo S, eds, *Proceedings 3rd International Symposium for Semantic Mining in Biomedicine* (SMBM 2008), Turku. 2008; pp. 3–4. Keynote speech.

[4] Neamatullah I, Douglass MM, Lehman LWH, *et al.* Automated de-identification of free-text medical records. *BMC Medical Informatics and Decision Making* July 2008; 8.

[5] Ruch P, Baud R, Rassinoux A, Bouillon P, and Robert G. Medical document anonymization with a semantic lexicon. *Proc AMIA Symp* 2000: 729 ; 33.

[6] Grouin C. *Chaîne de traitement pour la constitution automatique de corpus : application sur le domaine médical pour le projet corpus CLEF*. DESS d'ingénierie multilingue, Institut National des Langues et Civilisations Orientales, 2002.

[7] Zweigenbaum P, Jacquemart P, Grabar N, Habert B. Building a text corpus for representing the variety of medical language. In: Patel VL, Rogers R, and Haux R, eds, *Medinfo*, 2001; pp. 290; 4.

[8] El Emam K et Kamal-Dankar F. Protecting privacy using k-anonymity. *J Am Med Inform Assoc* September/October 2008; 15(5):627; 37.

[16] Le test d'un système de reconnaissance d'entités nommées « général » (utilisé dans un système de recherche de réponses à des questions en domaine ouvert) non modifié (hormis la conversion de ses balises de sortie) a donné des résultats très faibles, avec un rappel de 0,15.

[9] Uzuner O, Luo Y, and Szolovits P. Evaluating the state-of-the-art in automatic de-identification. *J Am Med Inform Assoc* 2007; 14:550 :63.

[10] Friedlin JF, and McDonald CJ. A software tool for removing patient identifying information from clinical documents. *J Am Med Inform Assoc* September/October 2008; 15(5): 601 ;10.

[11] Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: The MetaMap program. *J Am Med Inform Assoc* 2001; 8(suppl): 17;21.

Adresse de correspondance

Cyril Grouin

LIMSI-CNRS, BP133, F-91403 Orsay Cedex, France

E-mail : grouin@limsi.fr, http://www.limsi.fr/

Informatique et Santé, 2009 (17) : 35-44
© Springer – Verlag France

Systèmes d'informations cliniques: au cœur des informatiques hospitalières

Christian Lovis

Service d'informatique médicale, Unité d'informatique clinique
Hôpitaux Universitaires de Genève
Genève, Suisse

Abstract

This paper focuses on describing the place and functions to be met by a modern clinical information system. Strongly supporting care processes, modern CIS is highly customizable to user needs in order to improve efficiency and safety. Organized around processes and transversal care pathways, it integrates documentation management systems embedded with order entry and decision support. Knowledge representation is deeply implemented to support clinical tasks as well as research, teaching and case-based databases. Improved care processes are tightly connected with the logistics of care, medico-technical and human resources within a unique and shared framework. Optimization of these resources can be done in real-time and help prediction of needs. Clinical information systems are not computerized patient records. The patient record is part of them, amongst many other sources and resources.

Keywords

Clinical information systems; Requirements; Architecture.

1 Introduction

Les systèmes d'information clinique sont en pleine mutation. Historiquement construits par des médecins cherchant à améliorer les outils informationnels à leur disposition, ils sont désormais des systèmes d'information complexes, orientés processus, au cœur des systèmes d'information hospitaliers [1-3]. Les systèmes d'information cliniques sont nés de l'informatisation des dossiers patients, longtemps considérés comme de simples archives informatisées des données médicales du patient. Plusieurs aspects dans l'évolution de ces systèmes peuvent être identifiés, certains étant particulièrement illustratifs :

a) Structuration
   Ces applications se sont initialement centrées sur la capacité à gérer des documents, généralement peu structurés, voire non structurés. Elles se sont développées sur une vision assez statique, avec une organisation en groupes de documents par patient, souvent triés de manière chronologique. Progressivement, le contenu est devenu de plus en plus structuré. Actuellement, les systèmes les plus évolués s'appuient sur des représentations de la connaissance, telles qu'ontologies, classifications et nomenclatures. Ces représentations, lorsqu'elles sont standards, permettent l'interopérabilité sémantique du contenu [4].

b) Contenu
Essentiellement organisé autour des documents médicaux, rapports, lettres, notes, plus ou moins structurés, les systèmes d'information cliniques se sont rapidement étendus vers d'autres types de données. Ceci s'est réalisé en termes de média informationnel, vers les images, les séquences animées, le son, divers signaux. Les systèmes actuels permettent de consolider de nombreux médias dans des environnements intégrés.

c) Intervenants
Le dossier initialement *médical* s'est étendu à l'ensemble des intervenants, tout particulièrement les soignants, pour devenir un dossier centré patient. Actuellement, certains systèmes supportent des visions spécifiques pour chaque type d'intervenants et permettent également au patient d'interagir avec les données le concernant.

d) Communication
Initialement dossier statique, vision du dossier traditionnel sous forme d'archives des données concernant le patient, le système d'information clinique s'est progressivement instauré comme média de communication interactif entre les partenaires de soins, intégrant par exemple la prescription informatisée et l'échange d'information.

e) Processus
Étendus à de multiples intervenants et devenus média de communication interactif, les systèmes d'information cliniques les plus modernes s'attachent désormais à soutenir des processus de prise de charge tout au long du séjour des patients, voire pour toutes les activités liées à leur problème de santé. Ce faisant, ces systèmes deviennent la pierre angulaire de l'opérationnalisation des soins dans des modèles multidisciplinaires et de prise en charge longitudinale.

f) Planification
La gestion opérationnelle des processus telle que supportée par les systèmes d'information clinique modernes permet dès lors la mise en place d'une stratégie effective de gestion et de planification des ressources humaines et médico-techniques et un cadre de management opérationnel et stratégique prédictif. En outre, une meilleure planification des soins et des ressources a un impact sensible sur la qualité de la prise en charge et l'image de l'institution.

Ces évolutions se sont accompagnées d'objectifs plus ambitieux que le simple stockage de l'information. Ces objectifs peuvent être résumés en deux grands piliers : amélioration de la qualité des soins et amélioration de l'efficience de la production des soins.

Les objectifs qualitatifs sont nombreux. Ils incluent la qualité de la décision médicale, comme dans la prescription, mais également de l'application des guides de bonnes pratiques ; l'application de cette décision avec les itinéraires cliniques. Ils incluent aussi la facilitation des prises en charge multidisciplinaires, avec des dossiers communs partagés et une meilleure transmission de l'information. La participation du patient à sa prise en charge, par un meilleur accès à ses données et la capacité à interagir avec son dossier, à y ajouter des informations, s'avère également d'importants facteurs qualitatifs, particulièrement dans les maladies chroniques. Il ne faut pas oublier la recherche clinique et fondamentale qui bénéficient de données cliniques, le raccourcissement du temps entre une découverte et son application grâce à une meilleure communication, l'innovation, pour ne citer que les principaux. Les avantages qualitatifs sont innombrables et bien décrits dans la littérature [5-7].

Un des grands espoirs qui reposent sur les systèmes d'information cliniques réside dans leur capacité à améliorer l'efficience de la production de soins, à savoir une optimisation

des coûts sans sacrifier leur qualité. Ceci est plus difficile à démontrer, mais un nombre croissant de rapports de la littérature démontrent de tels effets, avec toutefois des délais de 5-10 ans entre le début de l'introduction de systèmes d'information cliniques et les bénéfices mesurables. Il s'agit donc d'efforts soutenus et sur le long terme [8-11].

Il ne faut toutefois pas négliger la complexité de la mise en place d'un système d'information clinique. Les échecs sont nombreux et le coût économique et humain de l'implémentation et du déploiement de tels systèmes est très largement sous-estimé. En particulier, il faut souligner le fait que le déploiement d'un système d'information clinique doit être considéré comme un changement majeur dans le mode de fonctionnement de l'institution, surtout lorsqu'il s'accompagne d'une refonte des processus de production de soins. C'est une erreur que de penser qu'un tel déploiement se fait sur la base d'un changement unique et limité dans un temps court. Il s'agit d'un travail continu, de longue haleine et qui nécessite une gestion adéquate du changement. Ce type de modification profonde d'une institution ne peut être le fait d'un service ou d'un département. Il doit s'agir d'une politique globale et institutionnelle, avec des moyens importants et un fort soutien des instances dirigeantes, organisationnelles soignantes et médicales de l'institution. L'effort et la charge pour les utilisateurs sont également souvent sous-estimés, particulièrement en termes de formation et de temps requis à l'usage du système d'information. L'activité devient dépendante du fonctionnement ininterrompu et fluide du système d'information qui devient dès lors un élément critique de l'institution. Ceci ne concerne pas seulement les aspects techniques de fonctionnement, mais également la capacité à offrir un support métier 24/24 h aux utilisateurs [12-26].

2 Enjeux et défis

2.1 Acheter, développer ou intégrer

Les cinq générations des systèmes d'information cliniques selon le Gartner group :

i) Le « Collector »
   Il s'agit ici d'une simple archive qui permet de stocker de l'information.
ii) Le « Documentor »
   Dans ce système, il est possible de documenter de manière structurée directement dans le dossier .
iii) Le « Helper »
   Ce dossier consolide l'information et apporte des aides informationnelles et utilise des systèmes de représentation de l'information.
iv) Le « Colleague »
   Ce système inclut de l'aide à la décision avancée, comme par exemple le guide de bonne pratique ou de l'aide à la prescription.
v) Le « Mentor »
   Ce système permet la gestion de processus complet de prise en charge, comme les itinéraires cliniques, avec la planification des soins et des ressources.

« *To buy or to built* » est sans doute devenu un des thèmes récurrent de discussion concernant les systèmes d'information clinique. La réponse doit être nuancée et va varier essentiellement en fonction de la taille et des objectifs de l'institution de soins.

En Suisse actuellement, la majeure partie des hôpitaux qui ont un système d'information clinique travaillent avec des systèmes commerciaux. Ces systèmes sont, toutefois, pour la majeure partie, de seconde génération. Ils permettent de documenter et d'accéder à la documentation des cas, ils implémentent des systèmes simples de prescription

informatisée, souvent limitée aux médicaments, avec des aides à la prescription de base : information sur les médicaments, groupes d'ordres simples, interactions et alertes initiales. Ils offrent une interopérabilité limitée avec des systèmes externes. Leur support multimédia est généralement restreint à afficher des types de fichiers externes. Ils ont cependant l'avantage d'être disponibles rapidement, d'être relativement robustes et d'avoir des communautés d'utilisateurs.

Certains systèmes sont entièrement développés en interne. Ces systèmes ont un avenir assez limité compte tenu des efforts et des compétences nécessaires à la maintenance. La tendance, dans ces cas, est à la création de spin-off qui permettent la commercialisation de ces systèmes et leur viabilisation économique et technique. Il faut relever que de nombreux systèmes commerciaux, en Suisse et de manière internationale, sont nés ainsi.

Finalement, certains hôpitaux comme les Hôpitaux universitaires de Genève, ont choisi une solution intermédiaire. Il s'agit essentiellement d'une approche intégrative. Cette approche considère qu'un système d'information clinique est trop complexe pour être un système monolithique, qu'il soit développé en interne ou commercial. Dès lors, la solution architecturale s'impose comme étant celle d'un assemblage de composants interopérables, certains provenant de solution industrielle et commerciale, d'autres étant développés en interne. L'effort à fournir est essentiellement lié à la capacité à imposer aux industriels le respect des normes d'interopérabilité pour les composants achetés, et à maintenir des compétences internes en matière d'intégration de composants. L'avantage indiscutable est l'évolutivité du système, son indépendance vis-à-vis d'un constructeur unique et son respect des standards. De plus, un autre avantage est la capacité à développer en interne des composants qui ne seraient pas disponibles sur le marché, voire à les commercialiser.

En résumé, le cadre et le contexte de l'institution et les objectifs attendus du système d'information clinique doivent dicter la stratégie en matière de choix de la solution. Il n'y a pas de réponse claire à ce sujet. Il faut souligner cependant que les systèmes commerciaux sont généralement de génération 3 au mieux et ne permettront pas d'atteindre des objectifs opérationnels d'efficience de manière aussi marquée que les systèmes de cinquième génération. Une étude de l'Union européenne comparant différentes approches et leurs impacts médico-économiques a été réalisée en 2008 dans le cadre du projet EHR*Impact* et dont les résultats préliminaires peuvent être obtenus [27-30].

2.2 Couverture fonctionnelle

Une vision complète de la prise en charge d'un patient requiert une large couverture fonctionnelle. Là encore, les choix et les décisions doivent être nuancés. Si les données d'un patient sont uniques, il est illusoire de penser qu'il existe un seul dossier patient en termes de fonctionnalités pour les utilisateurs. Ainsi, le médecin cardiologue regarde-t-il les données d'une façon radicalement différente de celle du médecin psychiatre. Il en de même pour l'infirmière de soins intensifs ou l'infirmière de dermatologie ou encore l'assistante sociale pour ne citer que ces exemples. La même remarque va s'appliquer aux fonctionnalités disponibles. Il est désormais admis que le radiologue doit disposer d'un système d'information adapté à la manipulation de grands volumes de données requis par les images. Toutefois, le cardiologue est dans la même situation avec les électrocardiogrammes, le neurologue avec les électroencéphalogrammes, le gastro-entérologue avec les endoscopies.

Encore une fois, ce ne sont là que quelques exemples qui illustrent une réalité de plus en plus admise : il n'y a pas de système unique de dossier patient qui soit efficacement utilisable par tous les intervenants. Cette remarque devient d'autant plus pertinente lorsqu'il s'agit de gérer de manière cohérente des dossiers hospitaliers et ambulatoires, des

dossiers médicaux et de soins, des dossiers sociaux, etc. tout en permettant des vues consolidées et pertinentes pour chaque intervenant [31-35].

Ici encore, les approches intégratives par composants apportent des solutions innovantes et puissantes, permettant de composer des vues adaptées au métier des intervenants mais basées sur un ensemble unique de données partagées concernant le patient. Ceci implique toutefois l'usage fort de standards, tant en termes de protocoles de communication entre les divers composants qu'en termes de représentation sémantique des données. Une telle approche permet d'utiliser un système spécifique pour le cardiologue par exemple, ce système étant toutefois assujetti dans un ensemble plus vaste représenté par les autres composants du système d'information clinique. L'interopérabilité devient ici un élément majeur et un défi important. Cette interopérabilité ne peut cependant se résumer à des échanges de données. Elle doit s'étendre à la capacité de faire fonctionner ces différents composants de manière synchrone, particulièrement si on s'attache à gérer des processus qui impliquent de nombreux intervenants et, par là-même, de nombreux composants du système.

2.3 Interopérabilité et standards

La notion d'interopérabilité est souvent comprise comme étant la capacité d'un hôpital à communiquer avec un réseau de soin communautaire, par exemple des médecins installés, des laboratoires d'analyses, des pharmacies, des assurances ou d'autres hôpitaux. Ceci est un élément important, mais trop limité. Le premier défi de l'interopérabilité est à l'intérieur même de l'institution, à savoir la capacité du système d'information clinique à communiquer avec l'ensemble des autres éléments du système d'information hospitalier. Ceci comprend bien évidemment toutes les autres parties directement liées au système d'information clinique, classiquement le système de gestion administrative des patients (ADT), le système de gestions des laboratoires (LIS) et de la radiologie (RIS-PACS). Cependant, de très nombreux autres systèmes doivent pouvoir être rendus interopérables dans une vision stratégique du système d'information : la gestion des ressources humaines, les menus, la gestion des stocks, la facturation, etc. La prescription d'un examen radiologique et d'un médicament peuvent illustrer ce point. Prescrire un examen de radiologie devrait pouvoir être connecté avec la disponibilité des ressources humaines, médecin radiologue et technicien en radiologie, et des ressources médico-techniques, à savoir du plateau de radiologie, de manière à ce que cette prescription puisse interagir immédiatement avec la planification de l'usage des ressources de radiologie. De même, la prescription d'un médicament devrait être assujettie à la disponibilité de cette substance dans l'institution. L'administration du médicament devrait automatiquement se répercuter sur la gestion des stocks disponibles et permettre la planification des achats. Dans les deux cas, ces actions devraient répercuter de manière automatique sur les coûts induits par la prise en charge du patient, y incluant la charge en ressources humaines que ces actes ont nécessités.

Les deux exemples simples ci-dessus illustrent la complexité des interactions entre les divers éléments du système d'information hospitalier et du système d'information clinique. Une telle complexité ne peut pas être résolue sans une utilisation stricte de nombreux standards d'interopérabilité pour tous les canaux concernés. Ceci ne peut pas se résoudre avec un seul standard et une solution monolithique, mais requiert de nombreux liens entre des systèmes souvent très hétérogènes. Il faut noter une complexité supplémentaire. En effet, il est nécessaire que les systèmes soient fortement communicants, tout en maintenant de très hauts niveaux de sécurité et de protection des données.

Actuellement, la communication et l'interopérabilité dans le cadre de réseaux de soins sont d'une complexité sensiblement plus faible techniquement, et les défis et enjeux majeurs se mesurent plus en termes sociétaux, politiques et législatifs (voir les liens TA-Swiss et sur la stratégie eHealth en Suisse) [3, 36-45].

2.4 Connaissance

Tout le problème de la gestion de la connaissance et du savoir ne sera pas abordé dans cet article compte tenu de sa complexité. Toutefois, il est important de relever que la plus-value majeure d'un système d'information clinique réside dans l'information et la connaissance qu'il contient. Ceci ne concerne pas seulement les données cliniques et médicales sur les patients, mais aussi et surtout la connaissance de l'institution, de ses référentiels. Par exemple, la qualité du système d'alertes sur les interactions médicamenteuses aura l'efficacité de sa base de données. Il en est ainsi concernant les protocoles de prises en charges, les ordres groupés, les bases d'allergies, etc. Ces bases de connaissances et ces référentiels sont souvent oubliés dans la mise en place d'un système d'information clinique et requiert de grands efforts, de construction et de maintenance.

Le système lui-même, au fur et à mesure de son enrichissement, devient une base de connaissance, de contenu de savoir et du savoir-faire institutionnel.

2.5 Gouvernance

Le système d'information clinique est un pilier de l'hôpital. Il est un élément clé de transformation et de gestion des processus, il est la pierre angulaire de la capacité de l'institution à mesurer et à planifier son activité en temps réel. Il est considéré comme un des facteurs clés de l'amélioration de la qualité et de l'efficience de la production de soins. À cet égard, il est renversant de constater combien la mise en place d'un tel système est souvent considérée comme un projet annexe et mal représenté dans les instances dirigeante de l'institution. On peut citer Clemenceau qui disait « *la guerre est une chose trop sérieuse pour la laisser aux militaires !* ». De même, le système d'information clinique est une affaire trop sérieuse pour la laisser aux informaticiens. Une nouvelle position devient proéminente dans les institutions américaines, nommée « CMIO – *Chief Medical Information Officer* » [1,2] : *"The new chief medical information officer title indicates its growing importance in health-care. Organizations are likely focusing on the necessary integration of technology and medical applications, which requires significant physician input and leadership"*. Ces positions sont des positions dirigeantes au plus haut niveau de l'institution et permettent d'intégrer les aspects technologiques, médicaux et informatiques dans une vision stratégique de l'institution. La Suisse est bien en retard dans cette vision de la gestion des hôpitaux. La plupart des systèmes d'information clinique sont pilotés par l'informatique ou ne sont pas pilotés du tout, considérés comme de simples applications informatiques. Dès lors, ces systèmes ne sont pas intégrés dans une vision stratégique et opérationnelle de l'institution. Il est indispensable d'avoir un rôle de médecin informaticien responsable du système d'information clinique impliqué de manière décisionnelle dans les aspects stratégiques et opérationnels de la gestion de l'institution [46-48].

1 http://mdsalaries.blogspot.com/2007/01/chief-medical-information-officer-cmio.html

2 http://www.fortherecordmag.com/archives/ftr_09182006p40.shtml

3 Conclusion

Cet article fait un survol rapide et volontairement subjectif de la place des systèmes d'information clinique, au cœur des systèmes d'information hospitalier et comme éléments stratégique central de la gestion et de la gouvernance des hôpitaux. Si la vision défendue est celle de l'auteur et se veut subjective, elle n'est ni isolée, ni universitaire et académique, ni même originale... Elle est celle d'un nombre croissant d'hôpitaux dans le monde, confrontés aux mêmes défis que nos hôpitaux suisses. Elle découle d'un constat simple et rationnel. Les technologies de l'information sont les seules à même de révolutionner en profondeur le fonctionnement des hôpitaux. Cette voie se fait aujourd'hui dans la douleur, sous la pression économique et de l'efficience. Il n'est reste pas moins que cette évolution est indispensable et, surtout, inévitable.

Le choix, le déploiement, l'utilisation et la gouvernance d'un système d'information clinique sont des éléments critiques au fonctionnement d'une institution. Ils nécessitent un leadership fort et inconditionnel, une vision stratégique claire et un investissement important et sur le long terme. La mise en place d'un tel système entraîne des modifications profondes du fonctionnement de l'institution, une révision des processus et, par là-même, une redistribution des enjeux de pouvoir. Il ne peut s'agir d'un projet purement informatique, au risque d'échouer ou de ne pas permettre d'atteindre les objectifs espérés en matière de qualité et d'efficience de production des soins.

Références

[1] ITAC. *Report to the President, Revolutionizing Health Care through Information technology.* President's Information Technology advisory Committee. June 2004.

[2] Lorenzi NM, Ash J, Einbinger J, McPhee W, Einbinger L. *Transforming Health Care through Information.* Lorenzi NM, Ash J, Einbinger J, McPhee W, Einbinger L, editors. New York, LLC: Springer-Verlag, 2004.

[3] Knaup P, Bott O, Kohl C, Lovis C, Garde S. Electronic Patient Records: Moving from Islands and Bridges towards Electronic Health Records for Continuity of Care. *Methods Inf Med.* 2007; 46 Suppl 1: 34-46.

[4] Rector AL, Nowlan WA, Kay S, Goble CA, Howkins TJ. A framework for modelling the electronic medical record. *Methods Inf Med.* 1993; 32(2): 109-19.

[5] Hillman JM, Given RS. Hospital implementation of computerized provider order entry systems: results from the 2003 leapfrog group quality and safety survey. *J Healthc Inf Manag.* 2005; 19(4): 55-65

[6] Dy SM, Garg P, Nyberg D, Dawson PB, Pronovost PJ, Morlock L, *et al.* Critical pathway effectiveness: assessing the impact of patient, hospital care, and pathway characteristics using qualitative comparative analysis. *Health Serv Res.* 2005; 40(2): 499-516.

[7] T raynor K. Customization key to successful CPOE. *Am J Health Syst Pharm.* 2004 Jun 1; 61(11): 1087, 92, 94.

[8] Thouin MF, Hoffman JJ, Ford EW. The effect of information technology investment on firm-level performance in the health care industry. *Health Care Manage Rev.* 2008; 33 (1): 60-8.

[9] Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: combined and uneven development. *Med Decis Making.* 2008; 28(1): 21-32

[10] Khoumbati K, Shah SG, Dwivedi YK, Shah MH. Evaluation of investment for enterprise application integration technology in healthcare organisations: a cost-benefit approach. *Int J Electron Healthc.* 2007; 3(4): 453-67

[11] Grieger DL, Cohen SH, Krusch DA. A pilot study to document the return on investment for implementing an ambulatory electronic health record at an academic medical center. *J Am Coll Surg.* 2007; 205 (1): 89-96.

[12] Aarts J, Berg M. Same systems, different outcomes--comparing the implementation of computerized physician order entry in two Dutch hospitals. *Methods Inf Med.* 2006; 45(1): 53-61.

[13] Mirco A, Campos L, Falcao F, Nunes JS, Aleixo A. Medication errors in an internal medicine department. Evaluation of a computerized prescription system. *Pharm World Sci.* 2005; 27(4): 351-2.

[14] Horsky J, Kuperman GJ, Patel VL. Comprehensive analysis of a medication dosing error related to CPOE. *J Am Med Inform Assoc.* 2005; 12(4): 377-82.

[15] Caudill-Slosberg M, Weeks WB. Case study: identifying potential problems at the human/technical interface in complex clinical systems. *Am J Med Qual.* 2005; 20(6): 353-7.

[16] Finegan BA. Access denied; care impaired: the benefits of having online medical information available at the point-of-care. *Anesth Analg.* 2004; 99(5): 1450-2; table of contents.

[17] Van Der Meijden MJ, Tange HJ, Troost J, Hasman A. Determinants of success of inpatient clinical information systems: a literature review. *J Am Med Inform Assoc.* 2003; 10(3): 235-43.

[18] Stavri PZ, Ash JS. Does failure breed success: narrative analysis of stories about computerized provider order entry. *Int J Med Inform.* 2003; 72(1-3): 9-15.

[19] Gainer A, Pancheri K, Zhang J. Improving the human computer interface design for a physician order entry system. *AMIA Annu Symp Proc.* 2003: 847.

[20] Featherly K, Garets D, Davis M, Wise P, Becker P. Sharpening the case for returns on investment from clinical information systems. *Healthc Q.* 2007; 10(1): 101-10, 4.

[21] Simon SJ. An examination of the financial feasibility of Electronic Medical Records (EMRs): a case study of tangible and intangible benefits. *Int J Electron Healthc.* 2006; 2(2): 185-200.

[22] Frisse ME. Computers and productivity: is it time for a reality check? *Acad Med.* 1998; 73(1): 59-64.

[23] Saathoff A. Human factors considerations relevant to CPOE implementations. *J Healthc Inf Manag.* 2005; 19(3): 71-8.

[24] Despont-Gros C, Mueller H, Lovis C. Evaluating user interactions with clinical information systems: A model based on human-computer interaction models. *J Biomed Inform.* 2005; 38(3): 244-55.

[25] Cegala DJ, Gade C, Lenzmeier Broz S, McClure L. Physicians' and patients' perceptions of patients' communication competence in a primary care medical interview. *Health Commun.* 2004; 16(3): 289-304.

[26] Boyle DK, Kochinda C. Enhancing collaborative communication of nurse and physician leadership in two intensive care units. *J Nurs Adm.* 2004; 34(2): 60-70.

[27] Frank S. Current and emerging business models in the health care information technology industry: a view from wall street. *Stud Health Technol Inform.* 2003; 92: 49-55.

[28] Hodgkins ML. Are you ready for the computer-based patient record? *J Ambulatory Care Manage.* 1995; 18(3): 1-8.

[29] DG INFSO and Media uIfH. EHR Impact. 2008 [updated 2008; cited]; Available from: http://www.ehr-impact.eu/index.htm.

[30] Sicotte C, Denis JL, Lehoux P. The computer based patient record: a strategic issue in process innovation. *J Med Syst.* 1998; 22(6): 431-43.

[31] Tang PC, LaRosa MP, Gorden SM. Use of computer-based records, completeness of documentation, and appropriateness of documented clinical decisions. *J Am Med Inform Assoc.* 1999; 6(3): 245-51.

[32] Moen A, Henry SB, Warren JJ. Representing nursing judgements in the electronic health record. *J Adv Nurs.* 1999; 30(4): 990-7.

[33] Kelly CS. Perinatal computerized patient record and archiving systems: pitfalls and enhancements for implementing a successful computerized medical record [In Process Citation]. *J Perinat Neonatal Nurs.* 1999; 12(4): 1-14.

[34] van Bemmel JH, van Ginneken AM, Stam B, van Mulligen E. Virtual electronic patient records for shared care. *Medinfo. 1998*; 9(Pt 1): 37-41.

[35] Sicotte C, Denis JL, Lehoux P, Champagne F. The computer-based patient record challenges towards timeless and spaceless medical practice. *J Med Syst.* 1998; 22(4): 237-56.

[36] OFSP. Stratégie nationale Suisse en matière de cybersanté ("eHealth"). 2008 [updated 2008; cited 2008 08.2008]; Available from: http://www.bakom.admin.ch/themen/infosociety/01689/index.html.

[37] Smith B, Ceusters W. HL7 RIM: an incoherent standard. *Stud Health Technol Inform.* 2006; 124: 133-8.

[38] Waegemann CP. Closer to reality. Personal health records represent a step in the right direction for interoperability of healthcare IT systems and accessibility of patient data. *Health Manag Technol.* 2005; 26(5): 16, 8.

[39] Ruotsalainen P. A cross-platform model for secure Electronic Health Record communication. *Int J Med Inform.* 2004; 73(3): 291-5.

[40] Geissbuhler A, Spahni S, Assimacopoulos A, Raetzo MA, Gobet G. Design of a patient-centered, multi-institutional healthcare information network using peer-to-peer communication in a highly distributed architecture. Medinfo. 2004 ;11(Pt 2): 1048-52.

[41] Blobel B. Architecture and tools for open, interoperable and portable EHRs. *Stud Health Technol Inform.* 2003; 97: 25-33.

[42] Blobel B. Secure communication and co-operation of distributed Electronic Patient Records. *Stud Health Technol Inform.* 2002; 87: 28-37.

[43] Dolin RH, Alschuler L, Beebe C, Biron PV, Boyer SL, Essin D, et al. The HL7 Clinical Document Architecture. *J Am Med Inform Assoc.* 2001; 8(6): 552-69.

[44] Aymard S, Fieschi D, Volot F, Joubert M, Fieschi M. Towards interoperability of information sources within a hospital Intranet. *Proc AMIA Symp.* 1998: 638-42.

[45] TA-Swiss. publifocus «eHealth»: le dossier électronique du patient. 2008 [updated 2008; cited 2008 08.2008]; Available from: http://www.ta-swiss.ch/f/them_info_pfeh.html.

[46] Runy LA. The changing role of the CMIO. *Hosp Health Netw.* 2008; 82(2): 37-42, 1.

[47] Hagland M. As I.T. and clinical continue to blend, the role of CMIO will grow. Sitting at the nexus of healthcare and I.T., CMIOs are in a unique position to ensure both sides of the aisle are on the same page. *Healthc Inform.* 2008; 25(2): 78, 80.

[48] Leviss J, Kremsdorf R, Mohaideen MF. The CMIO--a new leader for health systems. *J Am Med Inform Assoc.* 2006; 13(5): 573-8.

Adresse de correspondance

Prof Christian Lovis
Médecin responsable de l'unité d'informatique clinique
Hôpitaux Universitaires de Genève
24, rue Micheli-du-Crest
1211 Genève 4, Suisse
E-mail : christian.lovis@hcuge.ch

Chapitre 2

SNOMED, interopérabilité sémantique et multiterminologies

Informatique et Santé, 2009 (17) : 47-56
© Springer – Verlag France

SMTS® : un serveur multiterminologies en santé

M. Joubert[1], B. Dahamna[2], J. Delahousse[3], M. Fieschi[1], SJ. Darmoni[2]

[1] *LERTIM, Faculté de Médecine de Marseille, Université de la Méditerranée, France*
[2] *CISMEF, Centre Hospitalo-Universitaire de Rouen, France*
[3] *MONDECA, Paris, France*

Abstract

In the health sector, there exist practically as many different terminologies as there are fields of application. Hence, and given the growing need for cooperation between health players, a new "interoperable" terminology base is now required featuring a shared semantic frame of reference to enable the various terminologies to interact effectively. Three groups (LERTIM, CISMeF, and MONDECA) have set up a partnership designed to incorporate a dozen or so French-language terminologies into a single computer-based system capable of processing them simultaneously. In the first stage, we designed a model of the terminologies in question: CIM-10, CCAM, SNOMED 3.5, CISP, CIF, DRC, MeSH and CISMeF (an extension of MeSH). Once validated, this general model was merged with the ITM® model (MONDECA) for implementation. The result is an operational multi-terminology server integrating homogeneously nine of the most frequently-used French-language health-field terminologies. The next stage is aimed to incorporate other French-language terminologies – LOINC for laboratory data (as the French translation becomes available), the Orphanet terminology of rare diseases, and the unified drug thesaurus that the French medical publishing company Vidal SA is currently constructing.

Keywords

Terminology ; Terminologies integration ; Information systems.

1 Introduction

Il existe dans le domaine de la santé pratiquement autant de terminologies que de champs d'application. Ce sont, pour les plus utilisées d'entre elles :

- SNOMED pour le codage des données cliniques (même si la SNOMED International [1] que la France a acquise est peu utilisée actuellement, la SNOMED Clinical Terms [2] est maintenant largement acceptée dans le monde anglophone) ;
- CIM-10 (Classification Internationale des Maladies, Organisation mondiale de la santé) [3] et CCAM (Classification Commune des Actes Médicaux) [4] sont réglementairement utilisées à des fins médico-économiques dans tous les établissements français de soins publics ou privés ;
- CISP [5] est une classification conçue pour la médecine de famille et les soins primaires à l'initiative de la WONCA (*World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians)* ;

- MeSH (*Medical Subject Headings*) [6] est le thesaurus d'indexation et de recherche d'information privilégié dans le domaine de la santé dans le monde entier (sa version francophone est maintenue par l'INSERM) ;
- les domaines de la pharmacologie et de la pharmacovigilance utilisent les codes ATC (*Anatomical Therapeutical Chemical*, dont la version française est maintenue par le CNHIM) [7], MedDRA (*Medical Dictionary for Regulatory Activities*) [8] et WHO-ART (*Adverse Reaction Terminology*, Organisation mondiale de la santé) [9];
- LOINC (*Logical Observation Identifiers Names and Codes*) [10] est un standard pour enregistrer électroniquement les résultats de laboratoire (l'Assistance Publique – Hôpitaux de Paris collabore à sa francisation).

Il en existe bien d'autres. Elles sont souvent de conception locale, liées à une institution ou à un service hospitalier. Cela soulève un autre problème : celui de la normalisation et de la standardisation.

Devant ce constat et la nécessité grandissante de permettre la coopération de différents acteurs de la santé, il apparaît nécessaire de rendre ces terminologies « interopérables ». Nous entendons par là, la mise en œuvre d'un référentiel sémantique commun permettant l'interaction efficace, avec un minimum de perte de sens, à la manière de systèmes, informatiques ou autres, qui s'échangent des informations et se « comprennent » parce qu'ils partagent des connaissances. Cette nécessité est criante lorsque l'on parle du dossier médical, partagé entre différents professionnels de santé et le patient, incluant différentes terminologies de santé, dont celles énumérées ci-dessus. Par ailleurs, l'introduction du patient et de sa famille, du citoyen en général, dans le système d'information de santé soulève un problème de capacité de lecture et de compréhension du contenu auquel différents auteurs se sont intéressés [31, 35-38]. L'intéropérabilité sémantique nécessite à nos yeux un modèle commun de représentation des termes, quels que soient leur terminologie ou leur référentiel d'origine, ainsi que les moyens de mettre en relation les termes d'une terminologie avec ses équivalents, directs ou indirects, dans d'autres référentiels. Nous illustrerons notre propos à l'aide d'un exemple.

Depuis les années 1980, le projet UMLS (*Unified Medical Language System*) [11] de la « *National Library of Medicine* » (NLM) des États-Unis intègre les terminologies de santé dans un thesaurus unique appelé Metathesaurus. Plus d'une centaine de terminologies à ce jour, pour plus de 1 400 000 concepts. Malgré tout, le Metathesaurus ne rend pas les terminologies intégrées interopérables au sens où nous l'avons énoncé ci-dessus puisqu'il se contente d'intégrer les différentes terminologies telles qu'elles se présentent sans établir de liens entre les termes de celles-ci autrement que par le rattachement de termes équivalents à un même identifiant du Metathesaurus. À la fin des années 1990, dans le monde anglophone, des chercheurs en science de l'information et en informatique médicale se sont penchés sur le problème de la mise à disposition de serveurs de terminologies dans le domaine de la santé [12, 13]. De ces études sont nés un certain nombre de systèmes, dont le plus célèbre est sans doute en Europe le système GALEN [14]. Depuis, des projets et initiatives ont suivi cette voie :

- SYMBIOmatics (*Synergies in Medical Informatics and Bioinformatics*) dans le domaine des « *Information & Communication Technologies for Health* » [15] ;
- L'objectif du réseau d'excellence intitulé « *Semantic Interoperability and Data Mining in Biomedicine* » (NoE) [16] est de faire de l'Europe le leader scientifique international en informatique médicale et biomédicale ;
- Le projet SemanticHEALTH [17] a pour ambition de fédérer au niveau européen les différentes actions en faveur de l'interopérabilité sémantique des systèmes de santé et de leurs infrastructures ;

- Le CEN (Comité Européen de Normalisation) [18] pour la normalisation en informatique de santé en Europe.

Pour les entreprises, citons Health Language, Inc. [19] qui délivre actuellement une plate-forme offrant des services facilitant l'accès aux terminologies standards et le « mapping » entre elles.

On voit donc, au travers des initiatives décrites, que la France et le français sont pauvrement représentés aux niveaux national et international en ce qui concerne les terminologies dans le domaine de la santé. Le LERTIM, CISMeF et un industriel (MONDECA), tous trois spécialisés dans la représentation et le traitement des terminologies et des connaissances se sont associés pour concevoir et réaliser un « serveur multi-terminologies de santé » francophones. Les équipes LERTIM et CISMeF ont collaboré au projet VUMeF[1] qui visait à l'amélioration de la présence du français dans l'UMLS. Ces trois partenaires ont réalisé l'intégration d'une douzaine de terminologies francophones, dont celles citées ci-dessus (à l'exception de LOINC actuellement), dans un système informatique unique capable de les exploiter simultanément. Ces exploitations sont multiples et correspondent à des services particuliers en rapport avec des tâches utiles dans les systèmes d'information en santé : aide au codage de l'activité des professionnels, communication entre applicatifs utilisant des nomenclatures différentes, intégration dans des entrepôts de données de bases de données hétérogènes, analyse de données dont le codage a été réalisé avec des référentiels ayant évolué dans le temps,...

2 Le serveur multiterminologies de santé (SMTS®)

La conception et la réalisation du SMTS® se sont faites en deux étapes : 1) conception d'un modèle général capable de représenter chacune des terminologies qui soit compatible avec le mode de représentation des terminologies du logiciel ITM® (Intelligent Topic Manager) de MONDECA, et 2) développer pour chaque terminologie un logiciel (appelé « parser ») pour l'intégrer dans le système.

2.1 Modélisation des terminologies

Nous avons conçu un modèle de chacune des terminologies. La confrontation de ces modèles a permis de dégager un modèle général dont chacun des modèles initiaux est un sous-ensemble (d'un point de vue conceptuel). Une telle approche présente l'avantage de combiner respect de la structure originelle de chacune des terminologies, et rassemblement des meta-données inhérentes à toute terminologie. Ainsi la structure de chaque terminologie n'étant rattachée qu'en quelques points généraux à l'ensemble, il devient aisé d'en intégrer les mises à jour spécifiques. On bénéficie de surcroît des mécanismes d'héritage puisque les éléments les plus génériques sont conservés, notamment les notions récurrentes de : "descripteur", "relation hiérarchique (BroaderTerm-NarrowerTerm)", "relation voir aussi (RelatedTerm-RelatedTerm)", "libellé (Label)" et "synonyme (UsedFor)". Les Figures 1 et 2 présentent les modèles de la CIM-10 et de la CCAM respectivement. Ces modèles sont décrits en UML (Unified Modeling Language), le langage standard de modélisation orienté-objet [20]. On y voit, dans le cas de la CIM-10 par exemple (figure 1), qu'un code a un libellé, appartient à un chapitre, ce chapitre est composé de catégories, elles-mêmes composées de sous-catégories. Il existe entre des codes de la CIM-10 des relations possibles de manifestation (ex : M00.0 Arthrite et polyarthrite à staphylocoques, M00.2 Arthrite et polyarthrite à streptocoques) et de

[1] Vocabulaire Unifié Médical Francophone. Projet RNTS 2004-2005.

causalité (ex : A15.1 Tuberculose pulmonaire, A17.0 Méningite tuberculeuse). Il existe aussi une relation possible d'exclusion indiquée dans la pratique du codage (ex : A15.4 Tuberculose des ganglions intrathoraciques, à l'exclusion de A15.7 Primo-infection tuberculeuse de l'appareil respiratoire). De manière analogue (figure 2), un acte fait l'objet d'une procédure unique, les codes des actes de la CCAM sont organisés en sous-paragraphes et paragraphes, puis en sous-chapitres et chapitres. Un acte peut faire l'objet d'un geste complémentaire (ex : EDPA005 Mise à plat d'un anévrisme iliaque avec remplacement prothétique aorto-iliaque ou aorto-fémoral unilatéral, par laparotomie, EDEA003 Réimplantation ou pontage de l'artère iliaque interne, au cours d'une intervention sur l'aorte) et être associé médicalement parlant à d'autres actes (ex : HHFA011 Appendicectomie par laparotomie, HMFA007 Cholécystectomie par laparotomie).

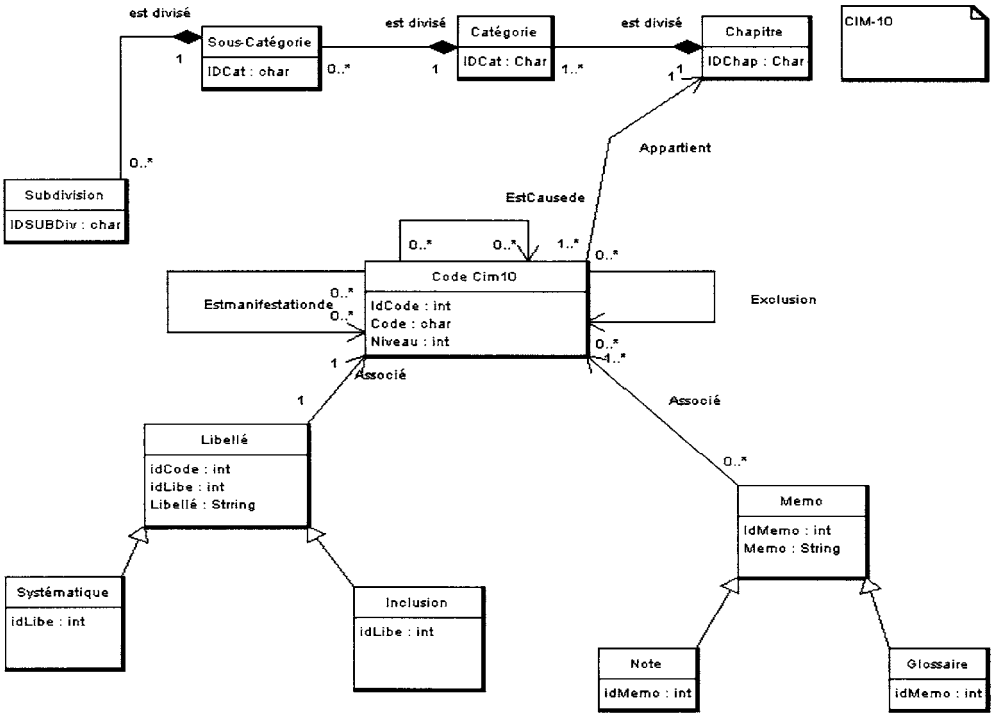


Figure 1 - Modèle UML de la CIM-10

Le modèle général validé a été confronté à celui d'ITM® afin de l'implémenter. Ceci ayant été fait, nous avons réalisé pour chaque terminologie le « parser » adapté afin de l'intégrer dans la base de données des terminologies d'ITM®. Ce modèle n'est pas publié ici pour des raisons de confidentialité industrielle qui unissent les différents partenaires.

2.2 Les résultats

Ce travail a abouti, en décembre 2007, à un serveur multi-terminologies opérationnel. Il intègre de manière homogène neuf terminologies de santé francophones parmi les plus utilisées, c'est-à-dire : CCAM, CIF (Classification Internationale du Fonctionnement, du

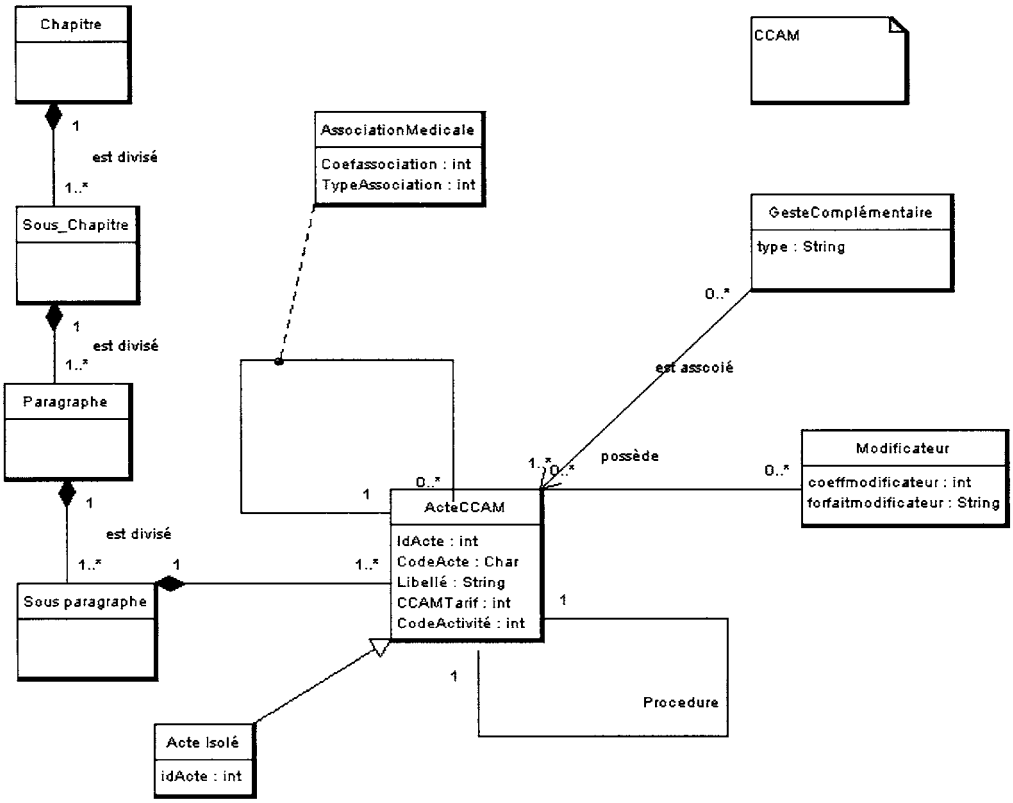


Figure 2 - Modèle UML de la CCAM.

handicap et de la santé – OMS) [21], CIM-10, CISP2, DRC (Dictionnaire des Résultats de Consultation publié par la SFMG) [22], Medline (le langage d'interrogation de MEDLINE) [23], MeSH, CISMeF (une extension de MeSH) [24] et SNOMED.

Bien que ces terminologies ne soient pas encore complètement interopérables à ce jour, au sens énoncé ci-dessus, elles « cohabitent » grâce aux éléments de connaissances que chacune d'elles amène et que nous exploitons.

Par exemple, la figure 3 montre une image d'écran relative au terme « infarctus post-opératoire du myocarde » choisi dans l'axe des maladies cardiovasculaires de la SNOMED et mis en relation avec un équivalent dans la CIM-10. En cliquant sur le lien proposé par le système, nous pouvons consulter les informations concernant l'« infarctus aigu du myocarde » dans la CIM-10 (I21.9). Entre autres, nous y trouvons tous les termes de la SNOMED mis en relation avec ce terme de la CIM-10, comme le montre la figure 4.

De manière analogue, l'« infarctus aigu du myocarde » de la CISP est mis en relation dans la CIM-10 avec « infarctus du myocarde » (I21), « infarctus du myocarde à répétition » (I22) et « certaines complications récentes d'un infarctus aigu du myocarde » (I23), comme le montre la figure 5.

EST DECRIT PAR, IS DESCRIBE WITH	
est Decrit par, is described by	decrit, describe
infarctus	infarctus post-opératoire du myocarde
myocarde, SAI	infarctus post-opératoire du myocarde
état postopératoire	infarctus post-opératoire du myocarde

SNOMED CIM10, SNOMED CIM10	
SNOMED/CIM10, SNOMED/CIM10	linked to SNOMED, lié à SNOMED
infarctus post-opératoire du myocarde	CIM10 I21.9

Figure 3 - Mise en relation d'« infarctus post-opératoire du myocarde » entre la SNOMED et la CIM-10.

Comme évoqué précédemment, des parseurs spécifiques à chaque terminologie mettent en adéquation les données et le modèle général. Ils traduisent les différents ensembles de données de leur format initial vers un format uniforme. Le langage adopté en ce sens est le principal langage d'ontologies Web : OWL (*Web Ontology Language*). Outre son caractère standard, il offre notamment l'avantage d'être suffisamment ouvert pour couvrir l'ensemble de nos besoins en termes de représentation des connaissances. Le modèle général a ainsi été décrit en OWL, de sorte que les différents parseurs terminologiques s'appuient directement sur celui-ci lorsqu'ils restructurent l'information.

SNOMED CIM10, SNOMED CIM10
SNOMED/CIM10, SNOMED/CIM10
rupture du coeur
infarctus aigu du myocarde
infarctus post-opératoire du myocarde
syndrome de nécrose du myocarde dû à un médicament
rupture d'une artère coronaire
infarctus du myocarde

Figure 4 - Les termes de la SNOMED mis en relation avec « infarctus aigu du myocarde » dans la CIM-10.

Cette dernière étape, que l'on nomme sérialisation, consiste à instancier les différentes classes du modèle. Les individus ainsi obtenus contiennent bien l'information des terminologies d'origine, et sont contraints de respecter le modèle commun. Ces différents fichiers d'instances, un par terminologie médicale, ont ensuite servi à alimenter la plate-forme ITM.

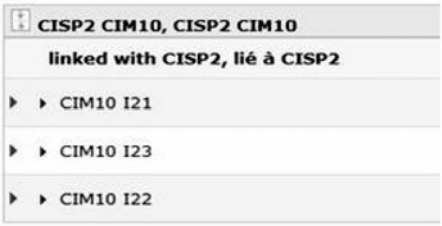


Figure 5 - Mise en relation du terme « infarctus aigu du myocarde » de la CISP avec des termes de la CIM-10.

Actuellement, le SMTS intègre en tout 198 « classes nommées » (ou classes primitives), et plusieurs centaines de milliers d'individus. La figure 6 présente la description OWL de la classe « Descripteur MeSH » comme une sous-classe de la classe « Desciptor ».

```
<owl:Class rdf:about="#MeSH"
itm:abstract="true">
<rdfs:label xml:lang="fr">Descripteur MeSH</rdfs:label>
<rdfs:label xml:lang="en">MeSH descriptor</rdfs:label>
<rdfs:subClassOf>
<owl:Class rdf:about="http://www.mondeca.com/system/publishing#Descriptor"/>
</rdfs:subClassOf>
</owl:Class>
```

Figure 6 - Représentation OWL de la classe « Descripteur MeSH » dans le SMTS.

3 Discussion

Le SMTS n'est pas une simple intégration statique de différentes terminologies dans un même modèle. Il propose une interopérabilité de ces terminologies enrichie par l'exploitation de leurs modèles de construction et de leurs structures. En effet, si nous reprenons l'exemple des figures 3 et 4, nous voyons que le terme « infarctus post-opératoire du myocarde », qui n'a pas d'équivalent direct dans la CIM-10, est mis en relation avec le terme « infarctus aigu du myocarde, SAI » de la CIM-10. Ceci offre, par exemple, la possibilité de proposer automatiquement un codage CIM-10 médico-économique au vu d'une information clinique codée en SNOMED.

C'est ce type de relations que nous comptons développer encore pour rapprocher la SNOMED et la CCAM, par exemple, mais aussi tous les rapprochements possibles entre les terminologies intégrées dans la base de terminologies du SMTS. L'intégration de nouvelles terminologies francophones doit se faire dans les prochains mois, à commencer par LOINC au fur et à mesure de sa traduction en français, mais aussi la terminologie Orphanet des maladies rares [25], et le thesaurus unifié en cours de construction par le Vidal [26]. Ce dernier doit unifier l'ensemble des terminologies utilisées par le Vidal dans sa gestion des informations sur les médicaments : les codes ATC, les indications et contre-indications, les interactions, etc.

Le projet InterSTIS[2] soutenu par l'appel à propositions TecSan de 2007 par l'Agence Nationale pour la Recherche (Projet ANR-07-TECSAN-010) doit nous permettre d'établir

[2] Ce projet est mené en partenariat avec Vidal SA, LERTIM, CISMeF, Mondeca, Mémodata, DSPIM (CHU de St Etienne), LIMSI (CNRS Orsay), LabSTIC (Faculté de Médecine de Nice) et la fondation Health on the Net (HON, Genève)

définitivement les relations souhaitées entre les termes des terminologies de santé intégrées dans le STMS en 2010. L'objectif du projet est de rendre interopérables les terminologies de santé francophones au sein d'un même système informatique. Il est dans la continuité du SMTS présenté ici et ses résultats en seront l'aboutissement. Pour ce faire, nous exploiterons toutes les sources de connaissances génériques possibles dans le domaine biomédical, dont l'UMLS et la SNOMED Clinical Terms. Le SMTS est actuellement à l'état de prototype et n'est donc pas accessible librement sur Internet.

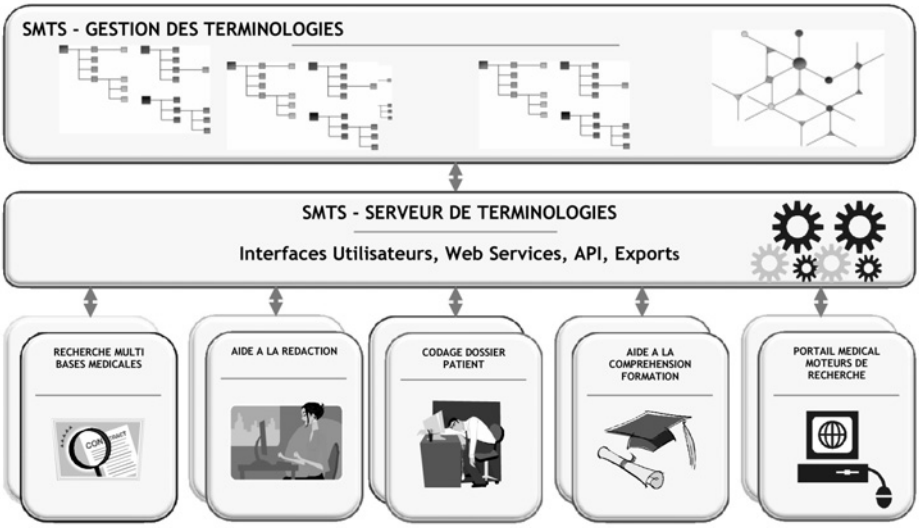


Figure 7 - Architecture et services du SMTS.

Les recherches réalisées pour rendre le contenu des textes médicaux plus facilement compréhensible par les patients et leurs familles sont d'un grand intérêt. Citons les réflexions de A.T. McCray [27] à propos de la capacité de lecture et de compréhension de documents relatifs à la santé des personnes concernées. Ces personnes ont besoin d'interagir dans une grande variété d'environnements incluant les cabinets médicaux, les cliniques et les hôpitaux. Elles sont aussi confrontées à une quantité importante d'informations relatives à la santé, comme les instructions relatives aux traitements, la formation à l'appareillage, les prescriptions, et l'assurance maladie sous toutes ses formes. Progressivement, ils seront de plus en plus concernés par leur propre prise en charge de leur état de santé. Face à cette situation, plusieurs projets se sont intéressés à une meilleure compréhension du langage médical par les citoyens. On peut citer : le « *Terminology Server* » de la NLM [28] dont l'objectif est de réduire l'écart de vocabulaire entre celui des utilisateurs des systèmes de santé et celui des systèmes d'information cliniques et des systèmes de recherche ; le projet WRAPIN[3] [29] dont l'objectif est de mettre à disposition des citoyens des sites Internet de santé de qualité accréditée.

Ces actions visent à l'intermédiation sémantique qui rejoint le problème abordé avec le SMTS. Elles visent à bénéficier de l'usage des terminologies standards, sans les faire dévier de leurs objectifs premiers, à des utilisations particulières. C'est sans doute dans cette voie que s'orienteront les travaux autour du SMTS pour l'adapter à l'usage qu'en font ses utilisateurs. Dans le cadre du projet InterSTIS, le laboratoire LabSTIC en expérimente

[3] Projet européen IST-2001-33260.

l'utilisation à des fins d'indexation et de recherche de documents pédagogiques, la Fondation HON intègre ses fonctionnalités dans le moteur de recherche WRAPIN.

La mise en place du SMTS doit déboucher sur la définition de Web services qui exploiteront cette plate-forme d'intermédiation sémantique entre nomenclatures pour favoriser l'activité des utilisateurs et l'exploitation des données recueillies à des fins éventuellement différentes de celle de la finalité qui a présidé à leur recueil. L'architecture du SMTS et des exemples de services auxquels il est destiné sont présentés dans le schéma de la figure 7. Outre ces applications dédiées à des utilisateurs professionnels de la santé, le SMTS offre les possibilités d'une plate-forme de développement pour des projets internes : indexation de documents [30] et recherche d'information [31] multi-terminologiques.

Remerciements

Les auteurs remercient toutes les personnes qui les ont aidés dans leur réalisation : les élèves ingénieurs de l'INSA de Rouen et les ingénieurs de MONDECA qui y ont activement participé. Ils remercient aussi le Dr Philippe Massari pour ses conseils avisés en matière de terminologie de santé.

Références

[1] Lussier YA, Rothwell DJ, Côté RA. The SNOMED model: a knowledge source for the controlled terminology of the computerized patient record. *Methods Inf Med* 1998; 37(2): 161-4.

[2] Stearns MQ, Price C, Spackman KA, Wang AY. SNOMED clinical terms: overview of the development process and project status. *Proc AMIA Symp* 2001: 662-6.

[3] Organisation Mondiale pour la Santé. *Classification statistique internationale des maladies et des problèmes de santé connexes*. Dixième révision.

[4] Assurance Maladie. CCAM. http://www.ameli.fr/professionnels-de-sante/gestionnaires-de-centres-de-sante/exercer-au-quotidien/codage-actes-medicaux-c.c.a.m./index.php

[5] Groupe De Travail CISP Pour L'Organisation Mondiale Des Collèges Nationaux, Académies et Associations Académiques. *CISP, Classification Internationale Des Soins Primaires*. Lacassagne, 1992.

[6] INSERM et National Library of Medicine. Le MeSH bilingue français-anglais. http://ist.inserm.fr/basismesh/mesh.html

[7] World Health Organization. Collaborating Centre for Drug Statistics Methodology. http://www.whocc.no/atcddd/

[8] Northrop Grumman. MedDRA and the Maintenance and Support Services Organization. http://www.meddramsso.com/MSSOWeb/index.htm

[9] WHO Adverse Reactions Terminology. http://www.umc-products.com/graphics/3149.pdf

[10] Forrey AW, McDonald CJ, DeMoor G *et al.* Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. *Clin Chem* 1996; 42(1): 81-90.

[11] Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med* 1993; 32(4): 281-91.

[12] Rector AL, Bechhofer S, Goble CA, Horrocks I, Nowlan WA, Solomon WD. The GRAIL concept modelling language for medical terminology. *Artifical Intelligence in Medicine* 1997, 9: 139-71.

[13] Chute CG, Elkin PL, Sheretz DD, Tuttle MS. Desiderata for a clinical terminology server. *Proc. AMIA Symp* 1999: 42-46.

[14] Rector AL, Solomon WD, Nowlan WA, Rush TW, Zanstra PE, Claassen WM. A Terminology Server for medical language and medical information systems. *Methods Inf Med.* 1995; 34(1-2): 147-57.

[15] Synergies in Medical Informatics and Bioinformatics. http://www.symbiomatics.org/

[16] Semantic Interoperability and Data Mining in Biomedicine. http://www.semanticmining.org

[17] SemanticHEALTH. http://www.semantichealth.org

[18] Comité Européen de Normalisation. http://www.centc251.org/

[19] Health Language, Inc. http://www.healthlanguage.com/

[20] UML, le langage de modélisation objet unifié. http://uml.free.fr/index-cours.html

[21] World Health Organization. Classification internationale du fonctionnement, du handicap et de la santé (CIF).

[22] Société Française de Médecine Générale. http://omg.sfmg.org/index.php

[23] US National Library of Medicine. http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed

[24] Douyère M, Soualmia LF, Névéol A, Rogozan A, Dahamna B, Leroy JP, Thirion B, Darmoni SJ. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. *Health Info Libr J* 2004; 21(4): 253-61.

[25] Orphanet. http://www.orpha.net/consor/cgi-bin/index.php?lng=FR

[26] Vidal SA. http://www.vidal.fr/societe/vidal

[27] McCray A. Promoting Health Literacy. *J Am Med Inform Assoc* 2005; 12(2): 152-63.

[28] Cognitive Science Branch. Terminology Server. http://www.lhncbc.nlm.nih.gov/cgsb_site/servlet/Turbine/template/research%2Cmedterm%2CTermServer.vm

[29] Joubert M, Gaudinat A, Boyer C, Fieschi M, HON Foundation Council members. WRAPIN: a tool for patient empowerment within EHR. *Stud Health Technol Inform* 2007; 129: 147-51.

[30] Pereira S, Névéol A, Kerdelhué G, Dahamna B, Joubert M, Darmoni SJ. Using multi-terminology indexing for the assignement of MeSH descriptors to health resources in a French online catalogue. *Proc. AMIA 2008*: 586-90.

[31] Dibad AD, Sakji S, Prieur E, Pereira S, Joubert M, Darmoni SJ. Recherche d'information multi-terminologique en contexte. *JFIM* 2009. À paraître.

Adresse de correspondance

Michel Joubert
LERTIM
Faculté de Médecine
27, bld Jean Moulin
13305 Marseille
E-mail : mjoubert@ap-hm.fr
http://lertim.org

Informatique et Santé, 2009 (17) : 57-68
© Springer – Verlag France

F-MTI : outil d'indexation multiterminologique : application à l'indexation automatique de la SNOMED International

Suzanne Pereira[1, 2, 3], Philippe Massari[1], Antoine Buemi[4], Badisse Dahamna[1], Elisabeth Serrot[3], Michel Joubert[2], Stéfan J. Darmoni[1]

[1]*CISMeF, CHU de Rouen, & TIBS, LITIS EA 4108, Université de Rouen, France*
[2]*LERTIM, Université de la Méditerranée, Marseille, France*
[3]*Vidal, Issy-les-Moulineaux, France*
[4]*AP-HP, Paris, France*

Abstract

Background: SNOMED is becoming a major health terminology to index electronic medical records. Most of developed countries have chosen SNOMED CT. France has chosen SNOMED International which is already translated in French. We developed the F–MTI tool, a generic automatic indexing tool able to index documentation in several health terminologies written in French (CCAM, TUV) or translated in French (MeSH, ICD-10, SNOMED International).

Objective: The objective of the paper is to evaluate the discharge summaries indexing in SNOMED performed by F-MTI and compare it to the one performed by Snocode (a commercial Canadian product).

Methods: We compared F–MTI and Snocode on a corpus of 100 discharge summaries from the Rouen University Hospital.

Results: The results showed a Hooper's measure of 32.9% comparing the two sets of SNOMED codes. With the help of a SNOMED-ICD10 mapping we could compare in terms of diagnosis these two sets with a manual coding. We obtained a precision of 6.1% for Snocode and 4.4% for F-MTI and respectively a recall of 24.7% and 27.0%.

Conclusion: Snocode and F-MTI indexing are as close as two manual indexing can be, in terms of precision and recall. They also provided close results in terms of diagnosis.

Keywords

Abstracting and Indexing/methods; Algorithms; Information Storage and Retrieval/methods; Evaluation Study, France; Natural Language Processing; Systematised Nomenclature of Medicine; Medical records; International classification of diseases.

1 Introduction

Représenter les données du dossier médical informatisé à l'aide de terminologies standardisées permet de faciliter les pratiques médicales, de rechercher plus facilement de l'information et de mieux gérer les coûts [1, 2].

La recherche en informatique médicale a montré que la SNOMED (nomenclature systématique de médecine humaine et vétérinaire) est la terminologie la plus adaptée à

l'indexation des informations du dossier médical [1, 2]. En 2006, la France s'approprie les droits de la version en français de la SNOMED International (SNOMED 3.5 VF) avec l'objectif d'en faire l'outil d'indexation des dossiers médicaux.

Ce choix peut paraître anachronique alors que plusieurs pays choisissent une version plus récente, la SNOMED CT (*Clinical Terms*) [3], mais se justifie par la non-disponibilité d'une traduction française de cette version[1]. La SNOMED 3.5 VF bénéficie d'une validation partielle de la traduction grâce aux travaux du consortium VUMeF[2] (2003-2006) financé par le programme RNTS. La SNOMED 3.5 et la SNOMED CT sont très proches sémantiquement (la première est incluse à 91 % dans la seconde), et une compatibilité à rebours existe entre les deux versions : la convergence avec la SNOMED CT, lorsque celle-ci sera disponible en français, est par conséquent facilement envisageable.

La SNOMED [4] est, à l'origine, une extension à l'ensemble de la médecine du concept développé par le College of American Pathologists avec la SNOP (Nomenclature Systématique d'anatomie Pathologique, 1965) [5], À ce jour, la version 3.5 contient plus de150 000 concepts répartis en 11 axes de description. Un axe regroupe les termes d'un sous-domaine (par exemple, l'axe D correspond aux diagnostics). Tous les axes peuvent être combinés pour former un concept SNOMED. La SNOMED 3.5 est traduite en 11 langues, la version anglaise étant incluse dans l'UMLS (*Unified Medical Language System* [6]).

Quel que soit le système proposé, les médecins sont réticents à pratiquer une indexation précise et exhaustive des dossiers médicaux, essentiellement du fait du grand nombre de codes à rechercher et de la complexité des référentiels proposés. Des systèmes d'aide à l'indexation et au codage sont nécessaires.

Certaines solutions existent pour la SNOMED CT utilisant les variations lexicales anglaises de l'UMLS [7, 8]. Pour la SNOMED RT (*Reference Terminology* [9]) et 3.5, elles utilisent les transcodages vers la SNOMED 3.5 [10, 11]. Certains produits commerciaux sont aussi disponibles : Language Engine-LE[3], CoPath Plus[4], 3M. DialeCT. Encoder[5] et Snocode[6].

À notre connaissance, le seul outil existant pour la SNOMED 3.5 VF est Snocode. C'est un outil d'indexation automatique interactif ou autonome destiné à l'indexation des textes cliniques proposant des codes SNOMED 3.5, ainsi que des termes provenant d'autres terminologies comme la Classification Internationale des Maladies 9e et 10e révisions (CIM9 et CIM10 [12]), la CCAM (Classification Commune des Actes Médicaux)[13] et la SNOMED CT. L'outil Snocode est disponible en plusieurs langues, incluant l'anglais et le français. En France, depuis 2000, Snocode est intégré au système de gestion des services d'anatomie pathologique des hôpitaux de Colmar et de Mulhouse.

Nos travaux sur l'indexation du dossier médical français débutent en 2005. Une étude préliminaire compare les codes CIM10 produits par un système combinant un extracteur MeSH (*Medical Subject Heading* [14]) et un transcodage entre le MeSH et la CIM10 à celle produite par Snocode combiné avec le même transcodage sur le même corpus [15].

---

[1] Une traduction partielle en français de la SNOMED CT est en cours au Canada (URL : http://www.infoway-inforoute.ca/).

[2] Vocabulaire unifié médical français.

[3] http://www.healthlanguage.com

[4] http://www.misyshealthcare.com

[5] http://www.3m.com

[6] http://www.medsight-info.com

Les résultats montrent que l'utilisation de la SNOMED avec Snocode à la place du MeSH aboutit à une précision de 25 % comparé à un codage descriptif manuel (vs. 4 % pour notre système) et un rappel de 46 % comparé à un codage descriptif (vs. 68 % pour notre système).

Le développement d'un système appelé F-MTI (*French Multi-Terminology Indexer*[7]) suit ces premiers travaux. F-MTI est un outil d'indexation automatique générique capable d'indexer des documents à l'aide de plusieurs terminologies. Nous avons déjà implémenté cinq terminologies françaises ou traduites en français : SNOMED 3.5, CIM10, MeSH, CCAM et TUV[8]. Il indexe directement et indirectement un document en utilisant les transcodages. Nous avons évalué cet outil lors de l'indexation de comptes rendus à l'aide de la CIM10. F-MTI aboutit à une précision de 3,7 % et un rappel de 32,9 % comparé à une indexation manuelle CIM10 [16].

Le but de ce travail est de confirmer nos objectifs précédents d'indexation multi-terminologique et de montrer que F-MTI peut aussi indexer un document en SNOMED. Dans ce travail, nous comparons les deux propositions d'indexation SNOMED de F-MTI et de Snocode produites à partir d'un échantillon de 100 comptes rendus d'hospitalisation. L'indexation manuelle de ces documents est considérée comme la référence.

2 Matériel et méthodes

2.1 Snocode

Snocode 3 incorpore les technologies propriétaires de stockage d'information et d'indexation en langage naturel de la société Medsight. Il utilise des techniques de transcodage et de synonymie pour comparer des séquences de mots d'un texte clinique aux termes SNOMED restructurés afin de faciliter et de rendre plus rapide les comparaisons. La stratégie de base consiste à ne retenir que les correspondances exactes et les plus longues. En même temps, une analyse syntaxique est utilisée pour les transformations singulier et pluriel. Les résultats peuvent être paramétrés. Dans notre étude, nous avons choisi les options : codes SNOMED et codes CIM10. Snocode est utilisé en mode autonome (sans action humaine). Le mode « une phrase à la fois » est sélectionné. Pour la comparaison au système F-MTI, tous les codes pour chaque phrase sont groupés et les redondances sont éliminées.

2.2 F-MTI

La première version de F-MTI est créée en 2006. Elle indexe un document à l'aide des terminologies implémentées : MeSH, SNOMED, CIM10, CCAM et TUV. Puis toutes les terminologies sont projetées vers les terminologies désirées par l'utilisateur (ici nous sélectionnons la CIM10 et la SNOMED) par le biais des transcodages. La plupart de ces transcodages proviennent du méta-thesaurus de l'UMLS. Certains sont développés par l'équipe CISMeF (transcodage de la CCAM vers le MeSH) ou par la société Vidal (transcodage TUV vers la CIM10). Cette méthode permet de retrouver plus de termes en identifiant plus de formes textuelles pour les termes des différentes terminologies. En effet, chaque terminologie a ses propres termes, synonymes et variations pour un même concept, les concepts équivalents dans différentes terminologies étant reliés par des liens de transcodage. Cette approche multiterminologique a été définie pour la première fois par la

[7] En rappel du projet MTI de la US National Library of Medicine [17].

[8] TUV :Terminologie Unifiée du Vidal décrivant les propriétés thérapeutiques des médicaments.

US National Library of Medicine avec MTI (MeSH *Terminology Indexer* [17]). À notre connaissance, F-MTI est le premier outil multiterminologique développé pour une autre langue que l'anglais. Cette réalisation est étroitement liée à la présence du français dans 10 terminologies de santé incluses dans l'UMLS (sans parler d'autres terminologies qui ne le sont toujours pas, comme la CCAM et le TUV).

Nous utilisons l'algorithme du sac de mots implémenté dans F-MTI:

- Le document est, dans un premier temps, converti en texte et les phrases sont extraites avec la rubrique correspondante grâce à un transducteur NOOJ [19].
- Ensuite, chaque phrase est normalisée (les accents sont éliminés, minuscules...) et découpée en mots. Les mots vides et non pertinents (car non inclus dans un terme) sont éliminés. Chaque mot est ensuite désuffixé (*stemming* en anglais). F-MTI est aussi capable de prendre en compte certains éléments du contexte pour une phrase. Le stème correspondant à la rubrique « antécédents » est ajouté pour les phrases appartenant à rubrique correspondante dans le compte rendu médical.
- Le sac de mots ainsi obtenu est mis en correspondance indépendamment de l'ordre des mots avec les termes des terminologies MeSH, SNOMED, CIM10, CCAM et TUV traités de la même manière. Les règles pour la mise en correspondance sont : tous les termes correspondant à au moins un stème du sac sont retrouvés et les correspondances les plus importantes sont privilégiées par rapport aux plus petites.
- Puis, les termes d'indexation candidats sont restreints aux termes SNOMED 3.5 les plus proches sémantiquement en utilisant les transcodages entre concepts. Les transcodages sont : MeSH->SNOMED 3.5 extrait de l'UMLS2007AA et une version du transcodage CIM10 vers SNOMED incluant le transcodage de l'UMLS2007AA et un transcodage français non validé inclus dans la version en français de la SNOMED. Une étude a débuté pour le valider [20].
- Finalement, tous les termes SNOMED et CIM10 extraits du document sont groupés. Les redondances sont éliminées, et les termes plus précis sont gardés.

2.3 Comparaison des deux systèmes : F-MTI et Snocode

Initialement, l'objectif de cette étude est de comparer les deux systèmes à l'aide d'une indexation manuelle SNOMED réalisée par un expert sur un échantillon de comptes rendus d'hospitalisation. Cette indexation descriptive (hors problèmes de codage PMSI) aurait été la première expérience d'indexation manuelle à l'aide de la SNOMED en France.

Toutefois, cette indexation manuelle en SNOMED s'avère trop consommatrice en temps : un jour de travail est nécessaire à l'expert pour indexer un compte rendu de 3 pages. Ce temps était partagé entre la reformulation du texte original, l'élimination des ambiguïtés, la reconnaissance des acronymes inhabituels, l'analyse du contexte, la reconnaissance des concepts et l'analyse des synonymes SNOMED. Lorsque le texte est correctement structuré et formulé sans ambiguïtés et non-dits, le temps d'indexation est beaucoup plus court. La nécessité de disposer de documents structurés avec une rigueur descriptive est ainsi réaffirmée !

Après ce constat, la projection des codes SNOMED vers une terminologie moins complexe pouvant être indexée manuellement semble être une bonne façon de comparer ces indexations, en termes de diagnostics tout au moins. La terminologie CIM10, utilisée dans les hôpitaux en France dans un but médico-économique, est retenue pour notre étude. Cette méthode indirecte reproduit la nécessité de coder les comptes rendus en CIM10 (figure 1) pour le PMSI en France [21].

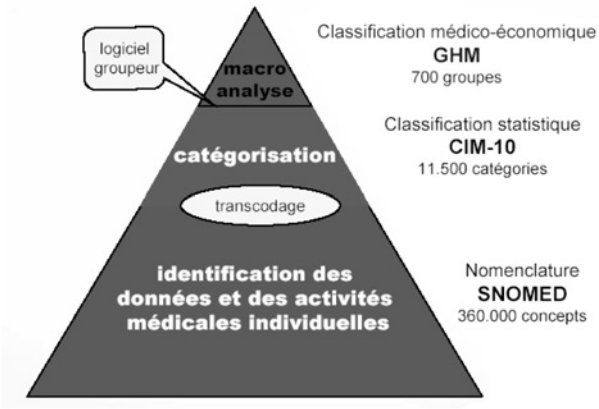


Figure 1 - Les niveaux de l'information médicale.

Les deux ensembles de codes SNOMED produits par F-MTI et Snocode sont comparés sans aucune référence avec de simples mesures. Puis les deux ensembles résultant de la projection des codes SNOMED en codes CIM10 sont comparés en utilisant une indexation manuelle descriptive à l'aide de la CIM10 réalisée par un expert.

Le processus de projection de la SNOMED vers la CIM10 est différent dans les deux systèmes. Snocode utilise un transcodage réalisé par le SFINM (Secrétariat Francophone International de Nomenclature Médicale) et F-MTI utilise une combinaison de celui-ci avec celui inclus dans l'UMLS2007AA. Nous disposons en interne de celui de F-MTI, c'est donc celui-ci qui est utilisé préférentiellement dans la suite de ce travail.

C'est pourquoi nous présentons deux évaluations : une pour les différents transcodages et l'autre en utilisant seulement le transcodage de F-MTI.

2.4 Corpus choisi

Un échantillon de 100 comptes rendus rédigés par des médecins du CHU de Rouen en 2005 est utilisé pour cette évaluation. Cinquante comptes rendus proviennent du service de cardiologie et cinquante de celui de pneumologie. Ces choix découlent du fait qu'il s'agit des domaines de prédilection de notre expert en indexation médicale (PM). Les dossiers sont extraits du système de dossier médical électronique du CHU de Rouen (1 080 384 patients et 182 808 comptes rendus en 2005). Un compte rendu médical décrit l'histoire de la maladie, les actes effectués, les examens et la prescription de médicaments. Après avoir lu chaque compte rendu, l'expert est invité à les indexer de manière descriptive à l'aide de la CIM10 pour former notre indexation de référence.

2.5 Mesures d'évaluation

Nous utilisons la mesure de Hooper pour comparer les deux ensembles de codes SNOMED. Cela mesure la consistance de l'indexation entre deux indexeurs. Nous avons considéré F-MTI et Snocode comme des indexeurs différents [22]. La mesure de Hooper :

$$H = (100*C)/F + S + C) \quad (1)$$

C = nombre de termes extraits par Snocode (S) et F-MTI (F); F = nombre de termes extraits par F mais pas par S; S = nombre de termes extraits par S mais pas par F.

Nous calculons aussi la couverture de chaque indexation l'une par rapport à l'autre.

Puis, nous mesurons la précision, le rappel et la F-mesure afin de comparer les deux indexations à la référence. La précision et le rappel sont des mesures usuelles en sciences de l'information. La précision est le ratio du nombre de codes pertinents extraits (codes proposés par le système et par la référence) sur le nombre total de codes extraits (par le système). Le rappel est le ratio du nombre de codes pertinents extraits sur le nombre total de codes de la référence.

La F-mesure est la moyenne pondérée harmonique de la précision et du rappel [23]. La formule de la F-mesure est :

$$F = 1/(\alpha*(1(1/P) + \alpha*(1/R)) \quad (2)$$

P = précision ; R = rappel ; dans notre cas $\alpha = 1$

3 Résultats

Les résultats de cette étude sont présentés dans les tableaux 1, 2 et 3.

Tableau 1 - Comparaison des outils Snocode et F-MTI en ne considérant que les codes SNOMED.

	Nombre de codes par CRH
Snocode	54.9
F-MTI	100.3
Pourcentage de codes extraits par F-MTI et recouvrant les codes de Snocode	29,9 %
Pourcentage de codes extraits par Snocode et recouvrant les codes de F-MTI	51,5 %
Mesure de Hooper	31,3 %

Le tableau 1 montre que Snocode extrait moitié moins de codes que F-MTI (54,9 versus 100,3). La moitié des codes extraits par Snocode sont aussi extraits par F-MTI.

Les tableaux 2 et 3 présentent les résultats de la comparaison des deux outils après transcodage vers la CIM10.

Tableau 2 - Comparaison des outils Snocode et F-MTI en ne considérant que les codes CIM10 produits par différentes projections des codes SNOMED vers les codes CIM10.

	Nombre de codes CIM 10par CRH		
Indexation manuelle	4,2		
Snocode	6,5		
F-MTI	26,5		
Pourcentage de codes extraits par F-MTI et recouvrant les codes de Snocode			4,4 %
Pourcentage de codes extraits par Snocode et recouvrant les codes de F-MTI			9,9 %
Mesure de Hooper			17,4 %
	Précision %	Rappel %	F-measure %
Snocode	15,0	22,2	17,9
F-MTI	2,5	30,7	4,6

Le changement du type de transcodage produit des résultats différents. Le nombre moyen de codes extraits par compte rendu passe de 7 à 17 codes extraits par Snocode (vs. F-MTI 26,5 codes et 4,2 pour l'indexeur humain).

F-MTI trouve plus de codes que Snocode et que l'expert, ce qui donne une précision très faible 2,5 %. Snocode produit une meilleure précision 15 % et 6,1 % avec le même transcodage. Les scores se rapprochent beaucoup lorsque l'on utilise le même transcodage. F-MTI produit un meilleur rappel (30,7 % vs. 22,2 %) et une plus faible précision (4,4 % vs. 6,1 %) par rapport à Snocode. La F-mesure de Snocode est supérieure à celle de F-MTI : 17,9 et 4,6. Ces résultats sont beaucoup plus proches (tableau 3) entre les deux outils quand on utilise les mêmes outils de transcodage.

Tableau 3 - Comparaison des outils Snocode et F-MTI considérant les codes CIM10 produits par la même projection des codes SNOMED vers les codes CIM10

	Nombre de codes CIM 10 par CRH		
Indexation manuelle	4,2		
Snocode	17,1		
F-MTI	26,5		
Pourcentage de codes extraits par F-MTI et recouvrant les codes de Snocode			4,4 %
Pourcentage de codes extraits par Snocode et recouvrant les codes de F-MTI			6,1 %
Mesure de Hooper			32,9 %
	Précision %	Rappel %	F-measure %
Snocode	6,1	24,7	12,2
F-MTI	4,4	27,0	7,6

4 Discussion

4.1 Comparaison entre les propositions d'indexation de Snocode et de F-MTI

Il n'est pas surprenant que le nombre de codes générés par les deux systèmes varie grandement (moyenne de 55 codes SNOMED pour Snocode contre 100 pour F-MTI ; moyenne de 17 codes CIM10 pour Snocode vs. 26 pour F-MTI). Ces variations sont dues au fait que Snocode se fonde uniquement sur la nomenclature SNOMED alors que F-MTI se fonde sur quatre autres terminologies pour générer des codes SNOMED.

Dans le tableau 1, la mesure de Hooper montre que les deux outils produisent des indexations aussi différentes que peuvent l'être deux indexations humaines (31,3 %). À titre de comparaison à la NLM, les indexeurs humains génèrent une mesure de Hooper de 39 % pour l'indexation manuelle des articles à inclure dans MEDLINE avec le thesaurus MeSH [24].

D'après les tableaux 2 et 3, nous pouvons suggérer que les principales différences de résultats entre les deux outils sont liées aux différences de transcodage SNOMED-CIM10 utilisés. L'application du même transcodage que celui utilisé par F-MTI aboutit à une diminution de 8,9 % de la précision et une augmentation du rappel de 2,5 % pour Snocode.

La projection des codes SNOMED vers la CIM10 montre que, comparé à une indexation manuelle, Snocode produit une meilleure précision (+ 2 %) et un plus faible rappel (- 6 %) que F-MTI en termes d'extraction de maladies. Les résultats peuvent être considérés comme assez proches alors que nous comparons un système mono-terminologique de plus

de 20 ans d'expérience (Snocode) et un système multi-terminologique de seulement 6 ans et qui peut encore beaucoup évoluer (F-MTI). Sachant que Snocode est un outil déjà commercialisé et en place dans certains hôpitaux, nous pouvons considérer que les résultats obtenus par F-MTI sont relativement satisfaisants.

4.2 Analyse des erreurs

L'analyse de l'indexation produite par F-MTI met en évidence quelques difficultés ou erreurs :

- L'extraction de termes non pertinents pour l'indexation, par exemple les termes de l'axe G de la SNOMED contenant les qualificatifs et termes de relations qui n'ont aucun sens lorsqu'ils ne sont pas reliés aux autres termes SNOMED.
- F-MTI, tout comme Snocode, ne permet pas de post-coordonner des termes appartenant à différents axes de la SNOMED lors de leur indexation. Il n'existe pas de règles d'indexation à ce sujet. Il est donc nécessaire d'implémenter des règles afin d'indexer correctement les comptes rendus médicaux. Cette étape devra être réalisée en post-coordination.
- Certains termes sont incorrectement retrouvés car l'extraction par la méthode du sac de mots ne permet pas de respecter l'ordre des mots. Des améliorations doivent être apportées dont l'implémentation de l'analyse sémantique des phrases.
- Les transcodages relient généralement des concepts de sens strictement équivalents mais parfois des degrés de précision différents peuvent être rencontrés. Les transcodages doivent donc être validés afin d'éliminer les transcodages inadéquats et ainsi faire diminuer le bruit généré par F-MTI.
- Les redondances entre termes extraits : les diagnostics et leurs symptômes ou différentes formes du même diagnostic ou bien encore la manifestation et la maladie initiale. Les relations « symptôme de » et « diagnostic de » sont présentes dans la SNOMED CT qui est reliée par des relations de synonymie à la SNOMED 3.5 dans l'UMLS (car reliés aux mêmes concepts UMLS - voir plus haut). Un travail réalisé par un doctorant de l'équipe CISMeF [25] permet de transposer les relations «symptôme de» et «diagnostic de» de la SNOMED CT à la SNOMED 3.5. Une future version de F-MTI doit intégrer ces règles et ces relations.
- L'analyse du contexte : antécédents, autre membre de la famille touchée, négations etc. Des améliorations au niveau de l'analyse du contexte, avec par exemple des transducteurs, sont en cours d'implémentation.
- F-MTI ne peut raisonner comme un médecin et, par exemple, associer des informations provenant de différentes parties du texte. Un système de règles peut être utile ici.
- La formulation des textes soumis à l'indexation : il existe un manque de précision des comptes rendus avec, par exemple, des diagnostics non décrits. Les médecins doivent être invités à une plus grande rigueur descriptive dans leurs comptes rendus.

4.3 Évaluation

Cette approche d'évaluation consistant à employer un transcodage vers d'autres terminologies moins complexes produit un biais important. Mais, nous partons de l'hypothèse que, si un code CIM10 est produit par Snocode et par F-MTI, il doit être issu d'un (ou plusieurs) codes SNOMED 3.5 équivalents, proches ou sémantiquement reliés. Cette évaluation permet donc de comparer en termes de diagnostics, la proximité (et non l'équivalence) des indexations produites par les deux outils.

Une évaluation manuelle par un expert des codes SNOMED 3.5 produits par les deux outils (sans biais provoqué par un transcodage quelconque) est en cours.

Cette approche d'évaluation peut facilement s'appliquer à d'autres évaluations où l'indexation manuelle est difficile, par exemple la SNOMED CT qui est beaucoup plus complexe que la SNOMED 3.5 et qui possède des liens d'équivalences avec la CIM10 dans l'UMLS.

Un expert n'indexe manuellement pas plus de 5 codes par compte rendu. En revanche, un outil automatique indexe dix fois plus de codes. Ce qui conduit à s'interroger sur la pertinence d'un codage exhaustif et sur la sélection des informations importantes selon l'utilisation qui en est faite. Dans sa pratique clinique, le médecin ne souhaite consulter que les éléments importants comme les maladies en cours et les derniers traitements. Dans le cadre médico-économique, les termes d'indexation sont souvent limités aux codes classants en vue de la tarification. En revanche, dans un contexte de recherche d'information, d'analyse de données ou d'alertes, une extraction complète des concepts présents dans le compte-rendu et décrits par la terminologie utilisée est préférable.

Dans cet esprit, il est prévu de réaliser une évaluation secondaire qualitative des codes extraits par F-MTI en assignant à chaque code, à dire d'expert, une étiquette associée à sa pertinence en matière de recherche d'information dans le dossier patient (« pertinent », « non pertinent » et « peu pertinent »).

4.4 Indexation SNOMED : une tâche complexe

La nomenclature SNOMED 3.5 VF contient sept fois plus de termes que la CIM10 et son utilisation est plus complexe du fait de la multiaxialité. Vu le peu de temps qui peut être consacré à l'indexation, il est certain que cette dernière ne peut se concevoir sans une assistance informatique ou une restriction très sévère des termes utilisés. Ces observations se transposent à l'indexation en SNOMED CT, celle-ci renfermant environ 370 000 concepts et 1 000 000 de termes (presque trois fois plus que la SNOMED 3.5) et plus de 1 300 000 relations (dans sa version 2007).

4.5 Perspectives

Plusieurs améliorations sont prévues pour F-MTI : management des négations, implémentation de dictionnaires et transducteurs. Avec l'intégration de F-MTI dans BIBLIS, un outil d'annotation [26], F-MTI sera capable de souligner manuellement les parties du texte que l'on désire indexer automatiquement (comme dans Snocode). BIBLIS sera aussi capable d'enregistrer les liens entre les parties du texte original et les termes.

F-MTI est un outil générique capable d'indexer n'importe quelle terminologie en français à partir du moment où sa structure est compatible avec celle de sa base de données. Celle-ci est créée sur un modèle générique inspiré du méta-thesaurus de l'UMLS. La SNOMED CT pourra facilement être intégrée lorsqu'elle sera traduite en français.

En pratique, nous envisageons d'exploiter les propositions d'indexation automatique produites par F-MTI dans un outil d'aide à l'indexation de documents. Le médecin, indexant son compte rendu d'hospitalisation, pourra être aidé par les propositions déjà effectuées par l'outil. Ce système pourrait être aussi intégré sans interaction avec le médecin mais ceci demande encore que l'outil soit amélioré.

L'outil indexe un compte rendu en 1,8 s en moyenne. Nous envisageons d'expérimenter l'outil en pratique semi-automatique, avec analyse du temps nécessaire à l'indexation

manuelle complémentaire et du gain qualitatif.

Le développement de l'extracteur multi-terminologie F-MTI fait partie d'un projet plus important dans l'équipe TIBS du laboratoire LITIS portant sur l'univers multi-terminologique. Trois thèses ont commencé en 2007 et 2008 : deux s'intéressent à la recherche d'information multi-terminologique en contexte [18][9] pour des portails de santé de l'Internet, comme le catalogue CISMeF [14] et pour des dossiers électroniques de patient. La troisième thèse s'intéresse à l'interopérabilité inter- et intra-terminologies en santé[10].

5 Conclusion

Nous avons développé un indexeur multiterminologique automatique effectif pour l'indexation en SNOMED 3.5, CIM10, CCAM, TUV et MeSH. La comparaison entre F-MTI et Snocode montre que les outils Snocode et F-MTI produisent des résultats proches. Ceci est encourageant pour notre projet. Après certaines améliorations, nous espérons voir F-MTI intégrer un système de dossier patient électronique.

Remerciements

Ces travaux ont été financés par la société Vidal (http://www.vidal.fr) et le projet PSIP FP7-ICT-1-5.2-Risk Assessment en Patient Safety.

Références

[1] Campbell JR, Carpenter P, Sneiderman C, Cohn S, Chute C, Warren J. Phase II Evaluation of Clinical Coding Schemes: Completeness, Taxonomy, Mapping, Definitions and clarity. *JAMIA* 1997 ; 238-251.

[2] Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. *AMIA Annu Symp Proc* 2003 ; 699-703.

[3] College of American Pathologists. SNOMED Clinical Terms Guide. January 2006.

[4] Côté RA, Rothwell DJ, Patolay JL, Beckett RS, and Brochu L, eds. The Systematized Nomenclature of Human and Veterinary Medicine: SNOMED International. *College of American Pathologists*, Northfield, 1993.

[5] Côté RA. From SNOP to SNOMED - A Challenge for the Medical Record Librarian. *Bulletin of the Canadian Association of Medical Record Librarians*; vol.5, n°1, December 1972.

[6] McCray AT and Nelson SJ. The semantics of the UMLS knowledge sources. *Methods Inf Med* 1995 ; 34(1/2).

[7] Brigl B, Mieth M, Haux R, Glück E. The LBI-method for automated indexing of diagnoses by using SNOMED. Part 2. Evaluation. *Int J Biomed Comput* 1995 ; 38 (2): 101-8.

[8] Long W. Extracting diagnoses from discharge summaries. *AMIA Annu Symp Proc.* 2005:470-4.

[9] Cette thèse est en partie financée par le projet européen PSIP (*Patient Safety Through Intelligent Procedures in Medication*), 7e PCRD.

[10] Ce projet est en partie financé par le projet InterSTIS (programme ANR Tecsan 2007).

[9] Spackman, KA, Campbell KE, Côté RA. SNOMED RT: a reference terminology for health care. *AMIA Annu Symp Proc* 1997 ; 640-4.

[10] Moore GW, Berman JJ. Automatic SNOMED coding. *Proc Annu Symp Comput Appl Med Care* 1994; 225-9.

[11] Sager N, Lyman M, Nhàn NT, Tick LJ. Automatic indexing into SNOMED III: a preliminary investigation. *Proc Annu Symp Comput Appl Med Care* 1994 ; 230-4.

[12] International Statistical Classification of Diseases and Related Health problems: World health Organisation (WHO), 1992.

[13] Hanser S., Zaiss A., Schulz S. Comparison of ICHI and CCAM basic coding system. *Stud Health Technol Inform* 2006; 124: 795-800.

[14] Douyère M, Soualmia LF, Névéol A, Rogozan A, Dahamna B, Leroy JP, Thirion B, Darmoni SJ. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. *Health Info Libr J* 2004 ; 21(4): 253-61.

[15] Pereira S, Névéol A, Massari P, Joubert M, Darmoni SJ. Construction of a semi-automated ICD-10 coding help system to optimize medical and economic coding. *MIE*2006; 845-850

[16] Pereira S, Massari P, Joubert M, Serrot E, Darmoni SJ. Exploring Multi-terminology Indexing of Discharge Summaries. *MIE*2008 poster.

[17] Aronson AR, Mork JG, Gay CW, Humphrey SM, Rogers WJ. The NLM Indexing Initiative's Medical Text Indexer. *MedInfo* 2004.

[18] Sakji S. Recherche multi-terminologique de l'information de santé sur l'Internet. CORIA (cinquième édition de la Conférence en Recherche d'Information et Applications), 2008.

[19] Silberztein M. NOOJ's Dictionnary. LTC2005.

[20] Buemi A, Boisvert MA, Côté RA. Transcodage SNOMED - CIM-10 : proposition pour une meilleure indexation des dossiers médicaux. *JFIM*2002.

[21] Geoffroy-Perez B, Imbernon E, Gilg Soit Ilg A, Goldberg M. Comparison of the French DRG based information system (PMSI) with the National Mesothelioma Surveillance Program database. *Rev Epidemiol Sante Publique* 2006 ; 54(6): 475-83.

[22] Hooper R. Indexer consistency tests—Origin, measurements, results and utilization. Bethesda. *IBM* 1965.

[23] Rijsbergen CJ. Information Retireval. *Butterworths*, London, 1979.

[24] Funk ME, Reid CA, McGoogan LS. Indexing consistency in MEDLINE. *Bull Med Libr Assoc* 1983; 2 (71): 176-83.

[25] Merabti T, Pereira S, Lecroq T, Joubert M, Darmoni SJ. Inheritance of SNOMED CT Relations between concepts to two Health Terminologies (SNOMED International and ICD-10). *KR-Med* 2008.

[26] Patriarche R, Gedzelman S, Diallo G, Bernhard D, Bassolet C, Ferriol S, Girard A, Mouries M, Palmer P, Simonet A, Simonet M. A Tool for Textual and Conceptual Annotations of Documents. *eChallenges* 2005.

Adresse de correspondance

Suzanne Pereira

Équipe CISMeF, Bibliothèque Médicale

1 rue de Germont 76031 Rouen cedex

E-mail : Suzanne.pereira@chu-rouen.fr

Informatique et Santé, 2009 (17) : 69-78
© Springer – Verlag France

Vers une version française de la SNOMED CT

Paul Fabry[1], Renald Lemieux[1], Andrew Grant[3]

[1]CESS - Centre d'expertise en santé de Sherbrooke, Sherbrooke, Canada
[2]CRED – Collaboration en Recherche pour l'Efficacité en Diagnostic, Sherbrooke, Canada

Abstract

In Canada, SNOMED CT® was identified by Canada Health Infoway as the best choice of terminology for the clinical content of the EHR. Canada is a bilingual country, therefore a French version of SNOMED CT is needed in order to fully support the EHR interoperability goals. A first phase of translation of SNOMED CT® is actually carried on by the Centre d'expertise en santé de Sherbrooke. In this paper the authors present the methodology and tools used for this project, the results obtained, and discuss the lessons learned with this project so far. A French version of the SNOMED CT® would be a key enabler for implementing terminology standards in all countries using French in health care services.

Keywords

Medical informatics; Terminology; SNOMED CT; Translation; French.

1 Introduction

La SNOMED CT (*Systematized NOmenclature of MEDicine – Clinical terms*) est considérée comme la terminologie la plus complète dans le domaine de la santé. Maintenant sous la responsabilité de l'IHTSDO (*International Health Terminology Standard Development Organisation*), la SNOMED CT est devenue la terminologie clinique de référence pour de nombreux pays [1].

Le Canada poursuit actuellement un vaste effort d'informatisation de son système de santé, incluant la mise en place d'un dossier santé électronique (DSÉ) accessible pour tous les Canadiens dans les prochaines années. Cet effort est piloté par Inforoute Santé du Canada (ISC), qui est une organisation à but non lucratif financée par le gouvernement canadien, dont le but est de favoriser l'élaboration et l'adoption de systèmes d'information et de normes dans le domaine de la santé [2].

Ainsi, le Groupe de Travail sur les Normes d'Inforoute Santé du Canada a identifié la SNOMED CT comme étant le meilleur choix de terminologie pour l'échange des données cliniques, et a recommandé son utilisation pour plusieurs sous-domaines d'information du DSÉ.

Le Canada a deux langues officielles, l'anglais et le français, et l'adoption de la SNOMED CT pour l'informatisation de son système de santé nécessite de disposer d'une version en français. Dans ce but, Inforoute Santé du Canada a lancé un appel d'offres en août 2007 pour la traduction d'une première tranche de la SNOMED CT pour les années 2008 /2009.

Le Centre d'expertise en santé de Sherbrooke (CESS) est un organisme à but non lucratif faisant la liaison entre la faculté de médecine et des sciences de la santé (FMSS), le centre hospitalier universitaire de Sherbrooke (CHUS) et le Centre de santé et de services sociaux – Institut universitaire de gériatrie de Sherbrooke (CSSS-IUGS). Le CESS se donne comme mission de soutenir le développement du savoir pratique en santé, d'optimiser la capacité d'exporter ce savoir et de proposer des stratégies d'amélioration du réseau de la santé et des services sociaux par la mise en application concrète des connaissances en santé [3].

Reconnaissant l'importance du DSÉ et la nécessité d'avoir une terminologie clinique bilingue de bonne qualité au Canada, le CESS a soumis une proposition pour la réalisation de cette traduction, qui a été acceptée par ISC. La revue d'un projet de traduction de la SNOMED CT nécessite de bien comprendre les principales caractéristiques de cette terminologie.

1.1 SNOMED CT

La SNOMED CT résulte de la fusion de la terminologie de référence SNOMED, élaborée par le Dr Roger Côté, professeur à la FMSS de l'Université de Sherbrooke, et le College of American Pathologists (CAP), et de la version 3 des Clinical Terms (CTV3), anciennement appelés Read Codes, développée par le National Health Service (NHS) du Royaume-Uni. Depuis 2005, la SNOMED CT est sous la supervision de l'IHTSDO.

La SNOMED CT comprend trois composants principaux : des concepts, des termes et des relations (figure 1).

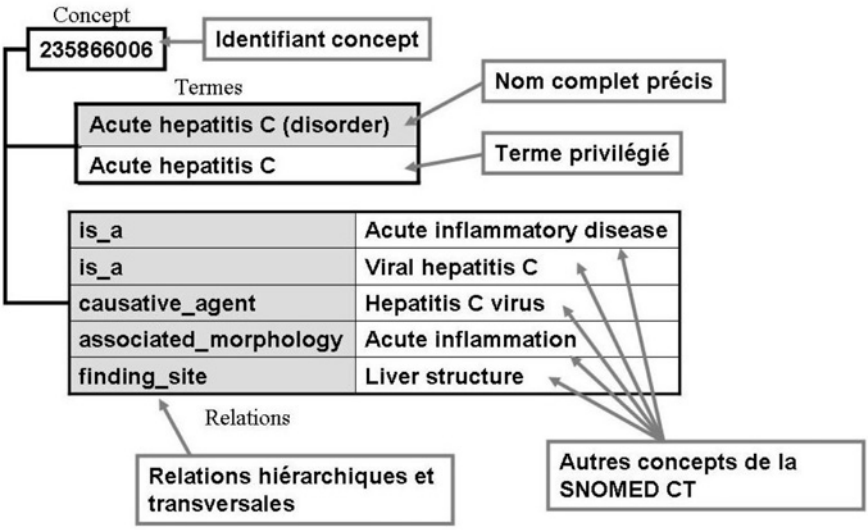


Figure 1 - Composants de la SNOMED CT.

1.1.1 Concepts

Dans la SNOMED CT, un concept est une « unité de sens » clinique et il est représenté par un identifiant numérique unique qui ne change jamais. Les concepts sont définis par les termes qui leur sont associés et par les relations qu'ils ont avec les autres concepts.

1.1.2 Termes

Chaque concept est associé à au moins deux termes distincts :

- Le Nom complet précis (*Fully Specified Name* en anglais), qui est destiné à identifier de manière unique un concept et à clarifier sa signification sans lui donner nécessairement le terme le plus communément utilisé.
- Le Terme privilégié (*Preferred term* en anglais), qui est le terme couramment utilisé par les cliniciens pour nommer le concept.

Il existe aussi des termes synonymes associés à certains concepts. Contrairement aux Noms complets précis, les termes privilégiés et les synonymes ne sont pas nécessairement uniques. Parfois, le Terme privilégié pour un concept peut également être un synonyme ou le Terme privilégié d'un concept différent.

1.1.3 Relations

Les relations sont de deux types :

- Des relations hiérarchiques de généralisation – spécialisation (*is_a*) qui organisent tous les concepts en une grande taxonomie. Cette taxonomie comprend un concept « racine », le concept [SNOMED CT], qui subsume tous les autres. Ce concept a 19 concepts enfants formant les 19 hiérarchies de la SNOMED CT : [body structure], [procedure], [disorder], etc.
- Des relations transversales plus variées qui unissent des concepts de hiérarchies différentes. Par exemple, la relation *finding_site* peut être utilisée pour relier un concept de la hiérarchie *disorder* avec un autre de la hiérarchie *body structure*.

À partir des relations de hiérarchisation, il est possible de définir le niveau d'un concept par rapport au concept « racine ». Le concept [SNOMED CT] représente le niveau 1 de la hiérarchie, ses 19 concepts enfants le niveau 2, leurs concepts enfants le niveau 3 et ainsi de suite.

Les autres caractéristiques importantes de la SNOMED CT sont d'une part sa taille et, d'autre part, la grande variété des domaines couverts par ses concepts. La dernière version a été publiée en juillet 2008 et comprend plus de 283 000 concepts actifs et plus de 700 000 termes actifs. En plus des domaines classiques des terminologies médicales comme les maladies, les structures du corps humain, les organismes vivants, elle couvre aussi le domaine vétérinaire, l'organisation de la société et des systèmes de santé, les substances pharmacologiques, ainsi que des notions de base en géographie, ethnographie, physique, chimie, etc.

2 Matériel et méthodes

La méthode mise en place pour réaliser cette traduction repose sur un processus de traduction manuelle coordonnée entre le CESS et ISC, et basé sur une plate-forme Internet. Ce processus est guidé par une série de principes établis en prenant compte des expériences de traductions antérieures, des recommandations de l'IHTSDO et des besoins identifiés par Inforoute santé du Canada.

2.1 Choix des concepts et des termes à traduire

Le choix des concepts à traduire combine deux approches distinctes : une approche systématique et une approche par priorité.

L'approche systématique consiste à traduire les concepts de la SNOMED CT par ordre hiérarchique décroissant, c'est-à-dire que l'on commence par le concept de niveau 1, puis

tous les concepts de niveaux 2, puis tous les concepts de niveaux 3, etc. Pour chaque niveau, les concepts de toutes les hiérarchies sont traduits.

L'approche par priorité consiste à traduire les concepts qui ont été identifiés dans le cadre d'autres projets d'informatisation de santé. Par exemple, les projets de DSÉ en cours au Québec, ainsi que des projets de surveillance de la santé publique en chantier au Québec, au Nouveau-Brunswick, en Ontario et au Manitoba ont établi les différents groupes de concepts SNOMED CT dont ils avaient besoin.

À l'opposé, certains concepts ne sont pas traduits, notamment les concepts du domaine vétérinaire. De plus, la SNOMED CT contient des termes qui sont spécifiques à un contexte national particulier et qui ne s'appliquent pas à la réalité canadienne. Par exemple, le concept [Patient NHS number] fait référence à un numéro d'identification de patient propre au système de santé du Royaume-Uni, et n'a pas d'utilité au Canada. Ces concepts ne seront pas traduits.

Un concept est associé à plusieurs termes, mais seul le Terme privilégié est traduit pour un concept donné, car celui-ci représente le terme considéré comme le plus adéquat du point de vue clinique. De plus, l'IHTSDO a émis la recommandation que le Nom complet précis demeure en anglais.

Les synonymes anglais ne seront pas traduits, et les synonymes français ne seront pas rajoutés afin de ne pas augmenter le travail de validation et d'assurer une fonction de normalisation à la terminologie.

2.2 Principes de base de la traduction

Le principe fondamental dirigeant la traduction est le choix d'une approche onomasiologique, c'est-à-dire qui se base sur le sens du concept plutôt que sur le terme associé au concept. En effet, la traduction d'un terme nécessite de connaître le contexte dans lequel ce terme est utilisé. Dans le cas de la SNOMED CT, le contexte est défini uniquement par la position du concept dans la hiérarchie et par ses relations avec les autres concepts, et il est impératif de prendre cela en compte dans la traduction.

La langue française comprend de nombreuses variations en fonction du pays ou de la région où elle est utilisée. Cependant, il a été décidé qu'il n'y aurait qu'une seule version française pour l'ensemble du Canada. De plus, la traduction doit privilégier les termes français en usage au Canada par rapport à des termes en français « international ». Par exemple, des termes comme *acétaminophène* ou *fibrose kystique* sont préférés à *paracétamol* ou *mucoviscidose*.

De plus, les termes français candidats à la traduction doivent répondre aux exigences suivantes :

- Non-ambiguïté : au sein d'une hiérarchie donnée, un Terme privilégié ne doit correspondre qu'à un concept et un seul.
- Justesse linguistique : les conventions orthographiques et les règles de syntaxe du français doivent être respectées.
- Transparence : le terme doit refléter les caractéristiques du concept qu'il représente.
- Acceptabilité : le terme doit tenir compte du langage en usage dans les milieux cliniques.
- Uniformité et cohérence : des termes de construction semblable doivent être utilisés pour des concepts de sens similaires.

Enfin, à côté de ces principes directeurs, une série de règles éditoriales et terminologiques ont été définies pour répondre à des questions très précises concernant par exemple la traduction de certaines expressions, le respect du genre ou du nombre, l'utilisation des signes de ponctuation, de caractères particuliers, etc.

2.3 Processus de traduction

La traduction s'appuie sur la compréhension des concepts de la SNOMED CT à travers l'analyse de leurs relations. Par conséquent, c'est un processus de traduction manuelle qui a été privilégié par rapport à une traduction automatique ou semi-automatique. En revanche, les traducteurs ont accès à des sources terminologiques bilingues pour faciliter leur travail, comme la Classification canadienne des interventions (CCI) [4] et les banques terminologiques médicales de l'Office québécois de la langue française (OQLF) [5].

Le processus comprend quatre étapes consécutives et deux flux de travail distincts (fig. 2).

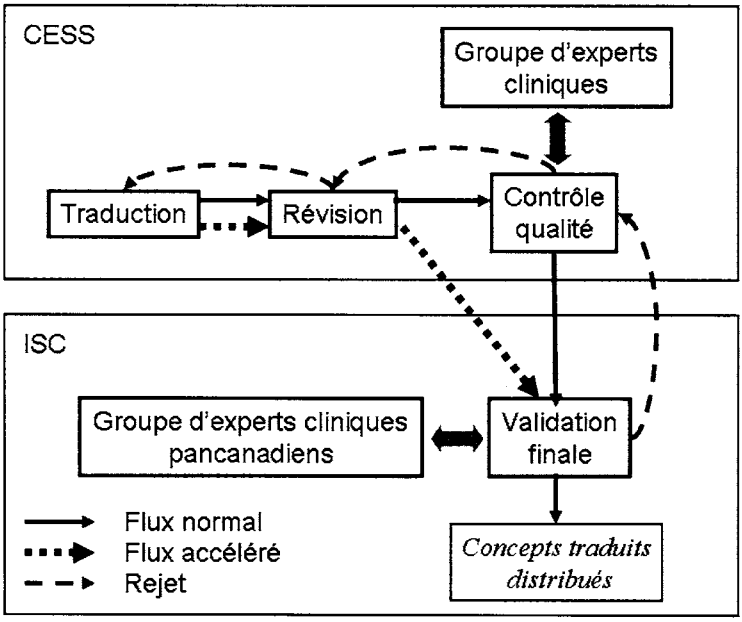


Figure 2 - Étapes du processus de traduction.

Les étapes de traduction sont les suivantes :

- La traduction : il s'agit de la première traduction proprement dite, où un traducteur propose un terme français pour le terme privilégié du concept à traduire.
- La révision : le terme français proposé est ensuite revu par un deuxième traducteur, qui peut le modifier ou le renvoyer au traducteur s'il n'est pas satisfait de la traduction, ou alors l'envoyer à l'étape suivante.
- Le contrôle de la qualité : le rôle du contrôle de la qualité est de s'assurer que le terme français révisé respecte bien le sens du concept, suit les règles éditoriales et correspond à une certaine réalité clinique. Le contrôle de la qualité peut modifier le terme, le renvoyer au réviseur, ou alors l'envoyer à l'étape suivante pour validation

finale. Il peut aussi demander un avis complémentaire auprès d'un groupe d'experts cliniques.

- La validation finale : la validation finale est faite par ISC pour s'assurer que les concepts sont prêts à être publiés et distribués au Canada. En cas de doute, les personnes responsables de la validation finale peuvent faire appel à un autre groupe d'experts cliniques pancanadiens pour s'assurer que le terme répond bien aux critères d'acceptabilité pancanadiens.

Les trois premières étapes sont sous la supervision du CESS alors que la dernière est faite par Inforoute santé du Canada.

Deux flux de travail sont distingués dans le processus. En plus du flux de travail normal où le concept passe toutes les étapes précédemment citées, il existe un flux de travail accéléré où le concept passe par le traducteur, le réviseur, puis va directement à la validation finale. Le choix entre les deux flux de travail est déterminé en fonction de la nature du concept à traduire. Quand il s'agit d'un concept du domaine général, dont le sens est évident, ou alors d'un concept qui ne doit pas être traduit, alors c'est le processus accéléré qui est choisi.

Ce processus a nécessité la constitution par le CESS d'une équipe de traducteurs avec une formation complémentaire en santé pour les étapes de traduction et de révision, et d'une équipe de médecins avec une bonne connaissance de la SNOMED CT pour le contrôle de la qualité et pour le groupe d'experts cliniques.

Enfin, un comité éditorial a été mis en place afin d'encadrer le travail de traduction. Le comité éditorial est composé de membres du contrôle de la qualité, du groupe d'experts cliniques et d'Inforoute santé du Canada. Son rôle est de développer un ensemble de règles éditoriales et terminologiques pour faciliter et normaliser le travail des traducteurs et d'assurer un contrôle de la qualité global. Ces règles sont diffusées par le biais d'un guide linguistique fréquemment mis à jour en fonction des problèmes soulevés par les membres de l'équipe.

L'ensemble de ce processus de traduction est rendu possible par l'utilisation d'une plate-forme de traduction informatique.

2.4 Plate-forme de traduction

La plate-forme utilisée est le logiciel HealthTerm, développé par la société danoise Carecom [6]. Cette plate-forme a été conçue pour la traduction de la SNOMED CT au Danemark et est aussi utilisée en Suède.

Il s'agit d'une application Web, spécialisée dans la gestion des flux de traduction. Son fonctionnement tourne autour de trois modules principaux :

- Un module de gestion de lots : il permet de sélectionner les concepts à traduire et de les regrouper en lots. Des rôles sont attribués à chaque lot, correspondant aux différentes étapes du processus de traduction, c'est-à-dire traduction, révision, contrôle de la qualité et validation finale. Ce module permet d'assigner pour chaque lot le rôle que joue un utilisateur donné, avec certaines contraintes. Par exemple, un utilisateur donné pourra avoir un rôle de traduction sur un lot, et de révision sur un autre, mais ne pourra pas avoir deux rôles différents sur le même lot.
- Un module de traduction : à partir de ce module, les utilisateurs ont accès aux concepts à traduire dans les lots qui leur ont été attribués, ils peuvent entrer ou modifier le terme français proposé en fonction de leur rôle, et ils peuvent rajouter un commentaire afin de justifier leur traduction et de donner des références. Ce

module permet aussi de voir les relations des concepts à traduire avec les autres concepts de la SNOMED CT et propose une fonction de recherche de concordance : pour chaque Terme privilégié anglais à traduire, une liste de termes similaires au terme français correspondant est proposée. Cette liste est obtenue à partir des concepts de la SNOMED CT déjà traduits et des autres sources terminologiques disponibles pour le projet, c'est-à-dire la CCI et le dictionnaire de l'OQLF (Office québécois de la langue française).

- Un module de recherche : ce module propose divers outils pour naviguer dans la hiérarchie de la SNOMED CT ou pour rechercher des concepts à partir d'un terme ou d'un identifiant.

À côté de ces modules, l'application propose des outils d'audit et de statistique ainsi qu'une base de documents contenant le guide linguistique et les autres références à utiliser pour la traduction. L'application fonctionne en temps réel, ainsi un concept traduit est directement accessible à la personne assignée à l'étape suivante.

En plus de cette plate-forme, les membres de l'équipe utilisent le navigateur de la SNOMED CT CliniClue [7] et la base de données terminologique du bureau de la traduction Termium Plus [8].

3 Résultats

Le projet a débuté en mai 2008 par une phase pilote de 4 mois pour éprouver la procédure de traduction. Pour la phase pilote, le CESS a constitué une équipe de 4 traducteurs pour les étapes de traduction et de révision, 2 médecins pour le contrôle de la qualité et 2 médecins pour le groupe d'experts cliniques.

À la fin de la phase pilote, cette équipe a été étendue à 9 traducteurs et 3 médecins au contrôle de la qualité, avec pour objectif la traduction mensuelle de 4 000 concepts.

Au 21 novembre, près de 3 500 concepts ont été acceptés par ISC, 5 300 concepts sont en attentes de validation et 13 500 concepts sont en cours de traduction. La traduction réalisée couvre la totalité des concepts des niveaux 1, 2, 3 et 4 ainsi qu'un lot de concepts prioritaires pour le DSÉ du Québec. Il est prévu de traduire près de 35 000 concepts.

4 Discussion

Notre projet s'est inspiré des expériences danoise et suédoise de traduction de la SNOMED CT, ainsi que du travail réalisé par le groupe de travail sur la traduction de l'IHTSDO [9]. Cependant, ce projet possède des particularités liées à l'utilisation de la langue française d'une part, et aux choix des concepts à traduire en fonction de priorités cliniques d'autre part.

Concernant les principes régissant la traduction, l'approche par concept est l'élément clé. Il ne s'agit pas de traduire des termes anglais, mais d'associer des termes français à des concepts dont le sens est décrit non seulement par des termes anglais, mais aussi par leurs relations avec les autres concepts. Par exemple, le terme anglais *swab* se retrouve dans de nombreux concepts de la SNOMED CT et sa signification change suivant la hiérarchie dans laquelle se trouve le concept auquel il est associé :

- Dans la hiérarchie Objet physique (*physical object*), *swab* indique un objet utilisé pour faire des prélèvements : *écouvillon.*
- Dans la hiérarchie Intervention (*procedure*), *swab* indique l'intervention utilisant cet objet : *prélèvement par écouvillonnage.*

- Dans la hiérarchie Échantillon (*specimen*), *swab* indique le produit de l'intervention : *échantillon (prélevé) par écouvillonnage*. On retrouve parfois dans cette hiérarchie l'expression *swab sample* qui exprime exactement la même chose, *sample* étant implicite dans les autres termes formés avec *swab*.
- Dans la hiérarchie Constatation clinique (*clinical finding*), *swab* indique le résultat de la procédure : *écouvillonnage positif* (ou négatif).

Cette modification du sens des termes anglais en fonction de la hiérarchie ou des relations du concept met en évidence l'avantage d'un processus de traduction manuelle par rapport à des outils de traduction automatique, basés sur des termes. Cependant, une traduction manuelle a ses propres limitations.

Il s'agit d'un processus lent et onéreux, nécessitant de trouver une main-d'œuvre très qualifiée, possédant à la fois une expérience en traduction, une connaissance du domaine de la santé, et une certaine souplesse intellectuelle pour comprendre les particularités de la SNOMED CT et utiliser une plate-forme internet pour le travail. La formation de l'équipe a demandé beaucoup de temps et d'énergie au début du projet, mais les traducteurs arrivés plus tard ont pu bénéficier des leçons apprises et ont eu des temps d'apprentissage beaucoup plus courts.

Le processus utilisé est très différent d'une traduction classique. En effet, la traduction doit prendre en compte les contraintes liées à la structure logique de la SNOMED CT et aux habitudes de langage des cliniciens. Les exigences de justesse linguistique, de transparence, de non-ambiguïté des termes, qui sont considérées comme les bases d'une bonne traduction, entrent souvent en conflit les unes avec les autres lorsqu'il s'agit de traduire la SNOMED CT. Ainsi, la préférence marquée pour les éponymes et les abréviations dans le langage clinique s'oppose à la transparence ou à la justesse linguistique d'un terme, et la nécessité de représenter le sens du concept et de prendre en compte les autres concepts existants entre souvent en conflit avec tout le reste. La priorité est le respect du sens du concept et cela impose très souvent de faire des compromis sur la qualité de la langue française employée, ce qui est contre nature pour des traducteurs professionnels.

SNOMED CT n'est pas une terminologie parfaite. Des travaux précédents ont montré l'existence d'erreurs sur le plan conceptuel [10, 11], mais il en existe aussi sur le plan linguistique (syntaxe, orthographe). Cependant, il ne faut pas tenter d'interpréter la SNOMED CT, dans le but de corriger une erreur et de proposer un terme français de meilleure qualité. Le risque est d'aller à l'encontre de l'interopérabilité, avec un terme français qui n'aurait plus le même sens que le concept anglais. Par conséquent, des concepts ambigus se retrouvent traduits par des termes ambigus en français. En revanche, les ambiguïtés ou les incohérences trouvées par les membres de l'équipe sont documentées pour être retournées à l'IHTSDO dans le but d'améliorer la qualité de la SNOMED CT.

Le projet exclut l'ajout de synonyme lors de la traduction, afin de ne pas multiplier la charge de travail au niveau du contrôle de la qualité et de la validation finale. Cependant, les synonymes sont couramment utilisés dans le langage médical et le choix d'un terme par rapport à un autre a un côté arbitraire qui peut retarder l'adoption par les utilisateurs. Aussi nous avons décidé de rajouter les synonymes les plus importants sous forme de commentaires, afin de les conserver lors de la traduction et d'être en mesure de les extraire facilement s'il s'avère nécessaire de les rajouter. De plus, l'ajout de synonyme pourra être fait au moment de l'intégration de la SNOMED CT dans un système informatique.

L'utilisation d'une plate-forme internet avec gestion de flux de travail est nécessaire pour un tel travail. Elle permet de s'affranchir de contraintes de lieux et de temps pour le travail et permet d'avoir un bassin de recrutement plus large pour constituer l'équipe. Cependant,

cela entraîne une certaine rigidité dans le processus de traduction. Une fois qu'un utilisateur a envoyé un concept à l'étape suivante, il n'a plus de possibilité pour le modifier. De plus, le fait d'avoir une équipe qui n'est pas réunie physiquement peut nuire à l'uniformisation du travail. Par conséquent, si les traducteurs travaillent à domicile, l'équipe du contrôle de la qualité est maintenant réunie dans un même lieu.

Le projet utilise des sources bilingues anglaises françaises pour faciliter le travail de traduction. Malheureusement, la version française de la SNOMED International, prévue initialement dans le projet, n'est toujours pas disponible. Des travaux évaluent que près de 91 % des 150 000 termes de la SNOMED International pourraient être directement réutilisables pour la traduction d'autant de concept de la SNOMED CT [12]. La non-disponibilité de cette source nous oblige à traduire manuellement ces concepts, ce qui a entraîné une forte réduction de nos objectifs concernant le nombre de concepts traduits à la fin de cette première phase, et nous prive de termes cliniques de grande qualité.

Enfin, la qualité de la traduction est difficile à évaluer. Jusqu'alors les révisions des termes français faites par le groupe d'expert clinique d'Inforoute santé du Canada ont mis en évidence la pertinence clinique et le respect du sens des concepts, mais la qualité finale de la traduction ne pourra être établie qu'une fois ces termes utilisés en pratique.

5 Conclusion

La SNOMED CT devient une terminologie incontournable au niveau international dans le domaine de la santé. Parallèlement à la traduction, d'autres travaux sont effectués concernant son intégration dans la norme d'échange de message HL7 et son interopérabilité avec d'autres terminologies comme LOINC ou la CIM-10.

Ce travail constitue à notre connaissance la première traduction de la SNOMED CT en français. Ce travail représente un énorme défi, devant la complexité et le nombre de concepts de la SNOMED CT, et le peu d'expérience antérieure concernant sa traduction ou son utilisation. Il représente aussi une très grande opportunité pour apporter au monde francophone cette terminologie.

Références

[1] IHTSDO: International Health Terminology Standards Development Organisation. http://www.ihtsdo.org/.

[2] Inforoute Santé du Canada. http://www.infoway-inforoute.ca/.

[3] Centre d'expertise en santé de Sherbrooke. http://www.expertise-sante.com/accueil.htm.

[4] Classification canadienne des interventions en santé. Ottawa; 2006.

[5] Office québécois de la langue française. http://www.olf.gouv.qc.ca/.

[6] CareCom. http://www.carecom.dk.

[7] CliniClue® - Clinical terminology server. http://www.cliniclue.com/.

[8] TERMIUM® - La base de données terminologiques et linguistiques du gouvernement du Canada. http://www.termiumplus.gc.ca/site/accueil_home_f.html.

[9] Høy A. *Guidelines for Translation of SNOMED CT®*. IHTSDO. Copenhagen, 2008.

[10] Bodenreider O, Smith B, Kumar A, Burgun A. *Investigating subsumption in SNOMED CT: an exploration into large description logic-based biomedical terminologies.* Artif.Intell.Med. 2007 Mar;39(3):183-195.

[11] Schulz S, Suntisrivaraporn B, Baader F, Boeker M. *SNOMED reaching its adolescence: Ontologists' and logicians' health check.* Int J Med Inform. 2008 Sep 10.

[12] Inheritance of SNOMED CT Relations between Concepts to two Health Terminologies (SNOMED International and ICD-10). KR-MED 2008; 2008.

Adresse de correspondance

Paul Fabry

CESS - Projet de traduction SNOMED CT

375, rue Argyll

Sherbrooke, Qc, J1J 3H5 Canada

E-mail : pfabry.cess@ssss.gouv.qc.ca

Informatique et Santé, 2009 (17) : 79-88
© Springer – Verlag France

Projection des relations SNOMED CT entre les termes de deux terminologies (CIM10 et SNOMED 3.5)

Tayeb Merabti[a, b], Hocine Abdoune[b], Thierry Lecroq[a], Michel Joubert[b], Stéfan J. Darmoni[a]

[a] *CISMeF, Département de recherche, Hôpital Universitaire de Rouen,France &* TIBS, *LITIS EA 4108, Institut de recherche Biomédicale, Université de Rouen, France*
[b] *LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille, France*

Abstract

Background: The situation of medical coding and medical economics is quite specific in France. Two specific health terminologies are used: ICD10 and CCAM. Another will be used in the near future SNOMED Int.

Objective: The objective of this study is to define relations between SNOMED International preferred terms and ICD-10 preferred terms inherited from the 136 SNOMED Clinical Terms relations.

Methods: We investigated the UMLS knowledge source (UMLS2007AB) to map the SNOMED CT relations to the two terminologies SNOMED Int. and ICD-10; we propose an interoperability wildcard to achieve this objective.

Results: 2,437,839 couples of preferred terms (PT) are in relation in SNOMED CT. Each SNOMED CT PT has at least one relation to another SNOMED CT PT; there is 74% of SNOMED International PT represented in 183,726 couples of SNOMED International PT linked and there is only 46% of ICD-10 PT represented in all the 5890 couples of ICD-10 PT linked. In the contrary, 70% of ICD-10 PT and 13% of SNOMED International PT are represented in the 33,097 couples of ICD10 PT and SNOMED International PT linked by at least one SNOMED CT relation.

Conclusion: This work should lead to an optimization of multi-terminology indexing tools, multi-terminology information retrieval and navigation among a multi-terminology server.

Keywords

International classification of diseases; SNOMED; Terminology; Vocabulary, Controlled; UMLS.

1 Introduction

La situation du codage médical est relativement spécifique en France. Dans de nombreux pays, CIM9 ou CIM10 (Classification Internationale des Maladies) [1] et CPT (*Current Procedural Terminology*) [2] ou bien même SNOMED CT (*Systematized NOmenclature of MEDicine Clinical Terms*) [3] sont utilisées. En France, deux terminologies spécifiques sont utilisées : CIM10 et CCAM (Classification Communes des Actes Médicaux) [3], SNOMED International (*Systematized Nomenclature Of MEDicine* – Nomenclature systématique de médecine) [4] ont été récemment acquis par l'État français pour décrire le dossier électronique du patient et notamment le futur DMP [5] (Dossier Médical

Personnel). Ainsi, se pose le problème de l'interopérabilité entre les différentes terminologies existantes dans le domaine médical. Il devient nécessaire de proposer différentes méthodes et plates-formes permettant le passage d'une terminologie vers une autre. Cette interopérabilité est complexe car la plupart des terminologies médicales existantes diffèrent entre elles, que ce soit dans leurs représentations (base de données terminologiques, classifications, logique de description...) ou bien dans leur contexte d'utilisation (intention) (domaine clinique pour SNOMED, base de données bibliographiques pour MeSH (*Medical Subject Headings*), médico-économique pour CIM10 ou CCAM...).

Le projet UMLS « *Unified Medical Language System* » [6], développé par « *US National Library of Medicine* », est le parfait exemple d'une plate-forme proposée pour permettre l'interopérabilité entre terminologies médicales existantes [7]. Actuellement, l'UMLS est considéré comme la plus large base de données terminologiques avec plus de 130 terminologies dans sa version 2007. Le SMTS (Serveur Multi-Terminologique de santé) [8] est un autre exemple de plate-forme développée pour permettre une interopérabilité sémantique entre terminologies francophones. Le SMTS utilise une technologie ITM (*Intelligent Topic Manager*) [9] et une plate-forme développée par la société MONDECA[1] pour la représentation des connaissances en utilisant les Topics-maps[2].

D'autre part, différentes méthodes de projection (« *mapping* » en anglais) entre terminologies ont été proposées où l'UMLS est utilisé comme source de connaissances [10]. Ces méthodes permettent en général la projection des termes d'une terminologie spécifique vers une ou plusieurs autres terminologies. Par exemple Cimino *et al.* [11] proposent une méthodologie permettant un passage de la terminologie ICD9-CM vers le MeSH en utilisant l'UMLS, Bodenreider *et al.* [12] proposent une autre méthode afin d'exploiter le réseau sémantique de l'UMLS pour des projections entre terminologies.

L'objectif de ce travail est de proposer une méthode permettant la projection de façon automatique des relations existantes entre les termes de SNOMED CT vers des couples de termes préférentiels CIM10 et SNOMED International en utilisant l'UMLS comme source de données pivot. Un terme préférentiel est un terme choisi pour représenter un concept au sein d'une terminologie. Par exemple, dans SNOMED International, le terme « achondroplasie » est le terme préférentiel pour représenter la classe des termes « achondroplasie de Parrot et Marie », « achondrodystrophie fœtale », « maladie de Parrot », « nanisme achondroplasique ».

Ce travail contribue au projet InterSTIS (INTERopérabilité Sémantique des Terminologies dans les systèmes d'Information de Santé français), un projet financé par l'ANR (Agence Nationale de la Recherche) et son programme TecSan (Technologie pour la Santé)[3].

2 Préliminaires

SNOMED pourrait devenir dans un avenir proche l'une des terminologies majeures dans le domaine de la santé pour l'indexation et la description pragmatique des documents médicaux. Plusieurs pays anglo-saxons ont choisi SNOMED CT, la France ayant fait le choix de SNOMED International parce qu'elle est déjà traduite en français. La France pourrait opter pour SNOMED CT lorsque celle-ci sera traduite en français.

---

[1] http://www.mondeca.com/

[2] http://www.topicmaps.org/

[3] Projet ANR-07-TECSAN-010

2.1 SNOMED CT

La SNOMED CT est une terminologie clinique qui fournit un contenu clinique et une expressivité pour la documentation et l'indexation cliniques. Elle peut être utilisée à des fins d'indexations, de recherches et d'analyses des documents cliniques. SNOMED CT est le résultat de fusion de SNOMED RT (*Reference Terminology*) [13] développée par le collège des pathologistes américains et « *clinical terms* » version 3 développée par la « *National Health Service* » anglaise. Plus de 310 000 concepts actifs sont présents dans SNOMED CT organisés suivant plusieurs niveaux hiérarchiques (tableau 1).

Tableau 1 - Les concepts de plus haut niveau dans SNOMED CT.

Clinical Finding	Specimen	Social context
Procedure	Qualifier value	Situation with explicit content
Observable entity	Record artifact	Staging and scales
Body structure	Physical object	Linkage concept
Organism	Physical force	Special concept
Substance	Events	
Pharmaceutical/biologic product	Environments/geographical locations	

Quatre types de relations existent dans SNOMED CT :

Les ***caractéristiques définies*** : la relation « ISA » avec les attributs définis sont considérées comme les « caractéristiques définies ». Les attributs définis relient deux concepts et établissent le type de relation entre eux. Ensemble avec la relation « ISA », elles sont considérées comme des « caractéristiques définies » car elles représentent d'une façon formelle la définition d'un concept en le liant avec d'autres concepts. La définition logique d'un concept dans la SNOMED CT inclue un ou plusieurs « concepts hiérarchiques » (modélisés avec la relation « ISA ») et un ensemble d'attributs définis pour expliciter la sémantique du concept et aider à le différencier des autres définitions avec les mêmes concepts hiérarchiques. Par exemple, le concept « fracture du métatarse» est définie par :

- ISA « fracture du pied ».
- Finding_Site_of (Localisation) « Structure osseuse de la cheville ».
- Associated Morphology (Morphologiquement associé à) « Fracture ».

Les ***caractéristiques qualifiées*** : ces relations sont utilisées pour créer des concepts plus complexes (post-cordination[4]) comme par exemple les relations « *severity* » et « *laterality* ». Par exemple le concept « bronchite aiguë » peut être post-cordonné en utilisant le concept « bronchite » et l'attribut qualifié « Clinical_Course » avec la valeur « aiguë ».

Les relations ***historiques*** : relient des concepts inactifs à des concepts actifs.

Des relations ***supplémentaires*** (« additional » en anglais) : sont des relations qui ne sont pas définies mais retenues pour être compatibles avec SNOMED RT. Les quatre types de relations définies précédemment sont représentés dans les 136 relations.

2.2 SNOMED International

La SNOMED International a été développée par le « Secrétariat francophone des nomenclatures internationales médicales » du Québec défini à l'origine pour les pathologistes. Elle est multi-axiale. La France a choisi SNOMED International en 2006

[4] Décrire un concept médical par la combinaison de deux ou plusieurs concepts.

dans sa version 3.5. Cette version est incluse dans l'UMLS et plus de 90 % des termes de SNOMED International sont présents dans SNOMED CT.

2.3 CIM10

L'appellation exacte de la Classification Internationale des Maladies est « Classification statistique internationale des maladies et des problèmes de santé connexes ». La désignation usuelle abrégée de « Classification internationale des maladies » est à l'origine du sigle couramment utilisé pour la désigner : « CIM ». CIM permet le codage des maladies, des traumatismes et de l'ensemble des motifs de recours aux services de santé. CIM10 est une classification mono-axiale, elle a été publiée en 1993 par l'Organisation Mondiale de Santé (OMS) [1]. Elle est utilisée à travers le monde historiquement pour enregistrer les causes de morbidité et de mortalité, mais très majoritairement dans un but médico-économique (intention épidémiologique).

2.4 UMLS

L'UMLS intègre plus de 2 millions de noms pour quelque 900 000 concepts issus des 130 terminologies, ces concepts sont liés entre eux en utilisant plus de 12 millions de relations.

UMLS est composé de deux composantes principales qui sont :

- le Métathésaurus, qui constitue la base unifiée des concepts médicaux. Il comprend des synonymes, des variations lexicales et des concepts associés afin de dresser la liste de tout le vocabulaire des expressions médicales disponibles. Chaque concept dans le Métathésaurus permet de regrouper différentes expressions des différentes sources qui expriment le même concept. Par exemple les termes « Fibrillation auriculaire », « FIBRILLATION AURICULAIRE », « FA », « AFib » qui proviennent de différentes terminologies sont rattachées à un même concept ;
- le réseau sémantique, qui spécifie les types sémantiques qui permettent de catégoriser tous les concepts médicaux définis dans le Métathésaurus. Il contient également les relations sémantiques existantes entre chaque type sémantique. Il existe 135 types sémantiques dans l'UMLS avec 54 relations sémantiques.

3 Méthodes

La première étape de notre étude a consisté à extraire tous les concepts UMLS liés par au moins une relation SNOMED CT. Par exemple, les deux concepts UMLS C0004099, C0004096 sont rattachés aux deux termes préférentiels SNOMED CT « asthme à l'effort » et « asthme » respectivement qui sont liés suivant la relation SNOMED CT « ISA ». Le tableau 2 représente les 10 relations SNOMED CT les plus représentées dans l'UMLS.

La deuxième étape de notre travail a consisté à projeter les relations SNOMED CT vers les termes préférentiels de CIM10 et de SNOMED International, ainsi l'idée consiste à projeter les couples de concept UMLS trouvés dans l'étape précédente vers les deux terminologies CIM10 et SNOMED International et à ne filtrer que les termes préférentiels. Au final, les couples de concepts UMLS sont remplacés par des couples de codes des deux terminologies liées par des relations SNOMED CT (figure 1). Formellement, le schéma d'interopérabilité proposé est défini ainsi :

Supposons que nous avons quatre termes SNOMED CT A, B, C et D.

- S'il existe deux termes CIM10 A' et B' telle que A' et B' correspondent aux termes A et B respectivement de SNOMED CT.

- S'il existe deux termes SNOMED International C' et D' tels que C' et D' correspondent aux termes C et D respectivement de SNOMED CT.
- Le fait qu'une relation SNOMED CT R1 existe entre A et B implique que cette relation sera projetée d'une manière automatique entre les termes A' et B' de la CIM10. De la même manière, l'existence d'une relation SNOMED CT R2 entre les termes C et D impliquera automatiquement une projection de cette relation entre les termes C' et D' de SNOMED International.

L'extensibilité de schéma permet non seulement une interopérabilité intra-terminologies entre termes d'une même terminologie mais elle permet une projection de relations inter-terminologies. Alors, s'il existe une autre relation SNOMED CT R3 entre B et C, la relation R3 sera projetée pour relier les deux termes B' et C' qui sont deux termes CIM10 et SNOMED International respectivement.

Tableau 2 - Les 10 relations SNOMED CT les plus représentées dans l'UMLS

Top 10 des relations SNOMED CT	**Nombre de couples en relation**
ISA	496 784
Finding _Site_of	86 358
Associated morphology	80 036
Method _of	54107
Part of	47810
Compenent_of	9135
Directe_procedure_site_of	34002
causative_agent_of	23628
indirect_procedure_site_of	8925
finding_method_of	8419

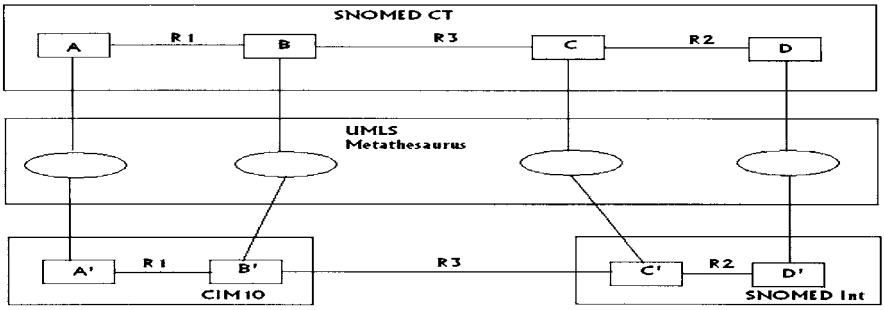


Figure 1- Schéma d'interopérabilité liant termes CIM10 et SNOMED International par des relations SNOMED CT.

4 Résultats

Un total de 1 051 085 termes de SNOMED CT est inclus dans l'UMLS, avec un nombre de 308 893 termes préférentiels. 2 437 839 couples de termes préférentiels sont en relation avec au moins une relation SNOMED CT, par conséquent chaque terme préférentiel SNOMED CT est lié à au moins un autre terme préférentiel SNOMED CT. De plus, 2 867 568 couples de termes non préférentiels sont en relation via au moins une relation SNOMED CT.

Tableau 3 - Le pourcentage des termes préférentiels de SNOMED International et de CIM10 partageant au moins un concept UMLS avec un terme préférentiel SNOMED CT

Terminologies	**Nombre de termes préférentiels**	**Nombre de termes préférentiels dans SNOMED CT**
SNOMED International	107 900	97 080 (91 %)
CIM10	9 308	7 933 (85 %)

Le tableau 3 indique le nombre de termes préférentiels SNOMED International et CIM10 qui partagent au moins un concept UMLS avec un terme préférentiel SNOMED CT. Pour ces deux terminologies (CIM10 et SNOMED International), 91 % de SNOMED International et 87 % de CIM10 ont un équivalent dans SNOMED CT.

Tableau 4 - Les 10 premières relations SNOMED CT projetées entre les termes de SNOMED International.

Relation SNOMED CT	**Nombre de couples de termes préférentiels de SNOMED International**
ISA	93 221
Finding Site of	24 661
Associated Morphology of	18 760
Direct_Procedure_Site_Of	11 077
Method_Of	7362
Causative_agent_Of	6062
Indirect_procedure_site_of	2883
Component_of	2678
Direct_morphology_of	2421
Active_ingredient_of	2097
Interprets	2090
Procedure_site_of	1508
Part_of	1401

Au sein de SNOMED CT, les 136 relations sont orientées, donnant lieu à 68 couples de relation : par exemple « la relation ISA » est représentée par deux relations symétriques dans SNOMED CT, « ISA » et « inverse ISA ». Nous considérons dans cette étude uniquement les 68 relations directes en représentant uniquement les relations à

caractéristiques définies (N = 50). En plus de ces relations, nous représentons aussi la relation supplémentaire « Part_Of » (Partie de), très utile en anatomie.

4.1 Projection des relations SNOMED CT vers SNOMED International

Un total de 183 726 couples de termes préférentiels SNOMED International est en relation avec au moins une relation SNOMED CT. Le nombre de couples de termes SNOMED International en relation est distribué d'une façon non uniforme suivant le type de relation, ainsi, nous avons 93 221 couples de termes préférentiels SNOMED International qui sont en relation via la relation SNOMED CT « ISA » et seulement 1 401 couples de termes préférentiels SNOMED International qui sont en relation suivant la relation SNOMED CT « Part_Of ».

Le tableau 4 montre les 10 premières relations SNOMED CT (en nombre) projetées vers les couples de termes SNOMED International.

Nous avons 74 % des termes préférentiels de SNOMED International qui sont représentés dans les 183 726 couples de termes liés par au moins une relation SNOMED CT.

4.2 Projection des relations SNOMED CT vers CIM10

Comme présenté dans la section précédente pour SNOMED International, un total de 5 890 couples de termes préférentiels CIM10 est en relation avec au moins une relation SNOMED CT. De même pour CIM10, le nombre de couples de termes CIM10 en relations est distribué d'une façon non uniforme suivant le type de relation. Ainsi, nous avons 5 019 couples de termes préférentiels CIM10 qui sont en relation via la relation SNOMED CT « ISA ». En revanche, nous n'avons aucun terme en relation via les deux relations « Method_of » et « Part_of ».

Le tableau 5 montre les principales relations SNOMED CT projetées vers les couples de termes CIM10. Nous avons 48% des termes préférentiels de CIM10 représentés dans les 5 890 couples de termes préférentiels qui sont liés par au moins une relation SNOMED CT.

Tableau 5 - Les principales relations SNOMED CT projetées entre les termes CIM10.

Relations SNOMED CT	**Nombre de couples de termes préférentiels de CIM10**
ISA	5019
Associated Morphology Of	834
Definitional manifestation	109
Associated with	74
Due to	49

4.3 Projection des relations SNOMED CT vers des couples de termes CIM10 et SNOMED International

Comme décrit dans la partie « méthodes » de cet article, l'extensibilité de notre schéma permet de projeter les relations SNOMED CT entre deux couples de terminologies différentes. Ainsi en appliquant ce schéma sur nos deux terminologies de références CIM10 et SNOMED International un total de 33 097 couples de termes CIM10 et SNOMED International, sont en relation via au moins une relation SNOMED CT. Le tableau 6 expose toutes les principales relations SNOMED CT projetées entre au moins un

terme CIM10 et un terme SNOMED International, 6 242 des termes préférentiels de CIM10 (67%) et 14 276 des termes préférentiels de SNOMED International (13%) sont liés via au moins une relation SNOMED CT.

Tableau 6 - Les principales relations SNOMED CT projetées entre termes SNOMED International et CIM10.

Relation SNOMED CT	Nombre de couples CIM10 et SNOMED International en relation
ISA	17 780
Associated_Morphology_of	6518
Finding_Site_of	5243
Causative agent_of	1258
Associated with	218
Direct_morphology_of	204

L'exemple de la figure 2 démontre bien la création d'une relation SNOMED CT inter-terminologique entre un terme CIM10 et deux termes SNOMED International. La figure montre l'exemple du terme CIM10 « achondroplasie » qui est lié d'une part avec le terme SNOMED International « dysplasie congénitale » suivant la relation SNOMED CT « associated morphology » (Morphologiquement associé à) et d'autre part avec le terme SNOMED International « Os » via la relation SNOMED CT « Finding Site of » (Localisation).

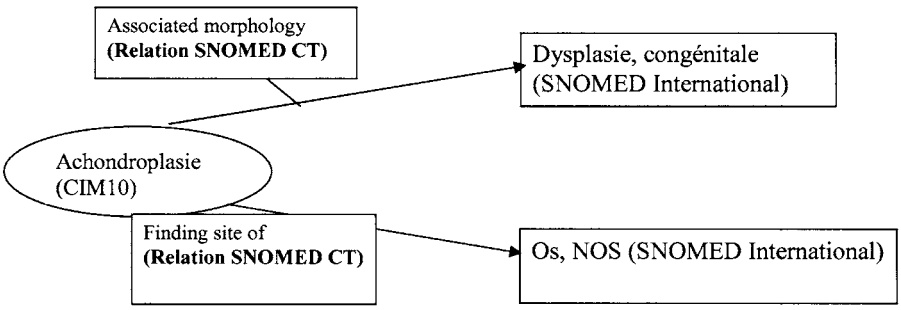


Figure 2 - Exemple d'application d'une projection de relations SNOMED CT entre deux termes SNOMED International et un terme CIM10.

5 Discussion

Nous nous sommes limités dans cette étude à deux terminologies (CIM10 et SNOMED International) qui sont des terminologies très importantes surtout dans l'indexation des dossiers médicaux du patient d'un côté et le fait que ces deux terminologies font partie actuellement du serveur multi-terminologique francophone de santé (SMTS) d'un autre côté.

Les résultats de cette étude devraient permettre, entre autres, d'optimiser l'indexation multi-terminologique (semi) automatique en utilisant la projection des relations SNOMED

CT vers les deux terminologies CIM10 et SNOMED International comme un poids supplémentaire dans le processus d'indexation. Le sujet de la prochaine étude sera l'impact de cette méthodologie sur l'extracteur multiterminologique [14] développé par l'équipe CISMeF [15].

Nos résultats actuels vont aussi être exploités dans le cadre de la recherche d'informations multiterminologique [16] afin d'étendre ou bien de restreindre une requête. Par exemple, si un utilisateur formule dans CISMeF la requête suivante « Achondroplasie », une proposition d'extension ou de limitation de sa requête lui sera faite :

« **Localisation (Finding_Site_Of)** Os ».
L'utilisateur pourra ainsi choisir d'étendre sa recherche avec « Achondroplasie **ou** Os » ou bien la limiter avec « Achondroplasie **et** Os ».

La principale perspective sera d'appliquer cette même méthodologie pour d'autres terminologies puis de l'appliquer à la matrice $n \times n$ des terminologies de santé disponibles en français dans le SMTS (n = 10 terminologies). Une évaluation est en cours selon ce même principe appliqué à la terminologie MeSH, qui est utilisée dans CISMeF, dans un contexte intraterminologique uniquement. Cette étude nous a permis d'établir des nouvelles relations entre termes MeSH qui n'existaient pas à l'origine. Par exemple, le terme MeSH « **syndrome d'Adams Strockes** » sera lié suivant une relation « **Localisation** » (Finding_Site_Of) avec le terme MeSH « **Cœur** ». Cette étude va permettre aussi d'améliorer la navigation dans le serveur multi-terminologique de santé en intégrant ces relations entre les terminologies déjà intégrées dans le serveur.

6. Conclusion

Cette étude a permis de définir des relations entre termes préférentiels CIM10 et SNOMED International. La méthode mise en œuvre doit être étendue à l'ensemble des terminologies de santé francophones. Cela permettra notamment d'améliorer les outils d'indexation multiterminologiques [14], les outils de recherche d'information multiterminologiques [16] et la navigation au sein du serveur multiterminologique [8].

Remerciements

Ce travail a été en partie financé par le projet INTERSTIS (Projet ANR-07-TECSAN-010-02)

Références

[1] CIM10 : URL : http://www.who.int/classifications/icd/en/ [Accès Oct 2008].
[2] CPT : URL : http://www.ama-assn.org/go/cpt [Accès Oct 2008].
[3] College of American Pathologists. *SNOMED Clinical Terms – User Guide*, July Release 2007, 2007.
[4] CCAM : URL : http://www.ccam.sante.fr [Accès Oct 2008].
[5] Côté RA, Rothwell DJ, Patolay JL, Beckett RS, and Brochu L, eds. *The systematised Nomenclature of Human and Veterinary Medecine*: SNOMED International. College of American Pathologists, Nothfield, 1993.
[6] Fieschi M. Les données du patient partagées : la culture du partage et de la qualité des informations pour améliorer la qualité des soins. *Rapport au ministre de la Santé de la Famille et des personnes handicapées*, 2003.
[7] Lindberg C. The Unified Medical Language System (UMLS) of the National Library of Medicine. *J Am Med Rec Assoc* 1990; 61(5): 40-2.

[8] Bodenreider O. The Unified Medical Language System (UMLS): Integrating biomedical terminology. *Nucleic Acids Res* 2004 ; 32(Database issue): D267-70.
[9] Joubert M, Dahamna B, Delahousse J, Fieschi M, Darmoni SJ. SMTS® : Un Serveur Multi-Terminologies de Santé. *JFIM* 2009.
[10] ITM : URL : http://www.mondeca.com/index.php/fr/intelligent_topic_manager [Accès Nov 2008].
[11] Fung KW, Bodenreider O. Utilizing the UMLS for semantic mapping between terminologies. *AMIA Annu Symp Proc* 2006 ; 266-270.
[12] Cimino J, Johonson S, Peng P, Aguirre A. From ICD9-CM to MeSH using the UMLS: a how-to guide. *Proc Annu Symp Comput Appl Med Care 1993*; 730-4.
[13] Bodenreider O, Nelson SJ. Beyond synonymy: Exploiting the UMLS Semantics in Mapping Vocabularies. *Proc AMIA Annu Symp* 1998; 815-9.
[14] Spackman KA, Campbell KE, Côté RA. SNOMED RT: A reference Terminology for Health Care. *Proc AMIA Annu Fall Symp* 1993;640-4.
[15] Pereira S, Neveol A, Kerdelhué G, Serrot E, Joubert M, Darmoni SJ. Using multi-terminology indexing for the assignement of MeSH descriptors to health resources in a French online catalogue. *Proc AMIA Annu Symp* 2008; 586.
[16] Darmoni SJ, Leroy JP, Baudic F, Douyère M, Piot J, Thirion B. CISMeF: a structured health resource guide. *Methods Inf Med* 2000; 39(1): 30-5.
[17] Letord C, Sakji S, Pereira S, Dahamna B, Kergoulay I, Darmoni SJ. Recherche d'information multi-terminologique : application à un portail d'information sur le médicament en Europe. *Ingénierie et Recherche Biomédicale/ Biomedical Engineering and Research*, Volume In Press, 2008.

Adresse de correspondance

Tayeb Merabti
Équipe CISMeF - Bibliothèque médicale
Centre universitaire hospitalier
1 Rue Gaumont, 76031 Rouen Cedex,
E-mail : tayeb.merabti@chu-rouen.fr

Chapitre 3

Représentation des connaissances et recherche d'information

Informatique et Santé, 2009 (17) : 91-100
© Springer – Verlag France

Définition d'une structure catégorielle pour les effets indésirables liés aux médicaments

Cédric Bousquet [1,2], Béatrice Trombert [1], Philippe Gasperina [1], Lucienne Clavel [1] et Jean-Marie Rodrigues [1,3]

[1]*Université de Saint-Étienne, Département de Santé Publique et Information médicale, France*
[2]*INSERM UMR-S 872, Eq 20, Paris, France*
[3]*Centre Collaborateur de l'OMS pour la Famille des Classifications Internationales en Français*

Abstract

The CEN Categorial structure is defined as a minimal set of health care domain constraints to represent a biomedical terminology in a precise healthcare domain. We present in this paper the first proposal of a categorial structure within pharmacovigilance. This study is based on the analysis and definition of 639 WHO-ART terms and the construction of an ontological model in an iterative way. Starting from a basic model based on previous experience, we added new semantic categories, changed the relations until the schema allowed to represent every term in a consistent way. We plan to investigate how this lite ontological model can help for better definition of adverse drug reactions using knowledge engineering tools such as ontology editors and reasoners.

Keywords

Pharmacovigilance; Categorial structure; Ontology.

1 Introduction

La pharmacovigilance a pour objet l'identification, l'évaluation et la prévention du risque d'événements indésirables résultant de l'utilisation des médicaments et des autres produits de santé [1]. Un effet indésirable lié à un médicament (EIM) est défini comme une réaction nocive et non voulue, se produisant aux posologies normalement utilisées chez l'homme pour la prophylaxie, le diagnostic ou le traitement d'une maladie ou la modification d'une fonction physiologique, ou résultant d'un mésusage du médicament ou produit [2]. WHO-ART (*World Health Organisation – Adverse Reaction Terminology*) [3] et MedDRA (*Medical Dictionary for Drug Regulatory Activities*) [4] sont les terminologies utilisées en pharmacovigilance pour le codage des EIM et l'analyse statistique des données. Ces terminologies sont déclarées comme obligatoires par les agences réglementaires nationales et internationales.

WHO-ART et MedDRA sont construites sur le modèle des terminologies traditionnelles comme la classification internationale des maladies. Par conséquent aucune définition formelle n'est disponible pour contraindre le sens des termes [5]. Dans un précédent

travail, nous avons mis en évidence l'intérêt du raisonnement terminologique pour implémenter certaines procédures informatiques dans le domaine de la pharmacovigilance.

- Le logiciel PharmaMiner effectue un prétraitement des données de pharmacovigilance en regroupant les termes qui désignent des conditions cliniques similaires. Plusieurs méthodes d'apprentissage sont appliquées sur les données regroupées ainsi et montrent une amélioration des performances de la fouille de données [6].
- Le logiciel PharmARTS permet de construire des requêtes complexes afin d'améliorer la recherche dans les bases de données de pharmacovigilance [7]. Par exemple, on peut demander l'ensemble des termes WHO-ART qui correspondent à des effets indésirables localisés dans la partie supérieure du tube digestif et sont associés à une hémorragie. Ce type de requête est rendu possible par la définition des termes WHO-ART de façon formelle, c'est-à-dire au moyen d'un langage informatique approprié.

Dans une première étape, nous avons étudié la construction manuelle d'une ontologie [8]. La modélisation de l'ontologie utilisée dans PharmaMiner a nécessité 300 heures de travail par un spécialiste en médecine interne. Cette ontologie est constituée de 390 concepts primitifs et de 530 concepts définis correspondant à des termes MedDRA. Nous avons ensuite essayé plusieurs approches pour l'extraction de connaissance en utilisant le traitement du langage naturel et/ou l'identification de relations dans le réseau sémantique UMLS.

En utilisant l'UMLS, nous avons considéré la mise en correspondance avec Snomed international [9] et SNOMED-CT [10]. L'alignement avec SNOMED International était basé sur les axes de la SNOMED. Afin de contraindre les définitions obtenues à partir de la SNOMED-CT, nous avons filtré les types de relations qui étaient les plus pertinentes pour le domaine. Alors que l'approche manuelle était très précise et coûteuse en temps, l'extraction de connaissances, bien que basée sur un modèle pauvre au niveau ontologique, était très pertinente pour diminuer les coûts de développement.

Avant que nous continuions le développement des définitions formelles, nous devons nous assurer d'une représentation ontologique consistante qui prend en compte non seulement les besoins minimaux pour la représentation des effets indésirables mais présente également une structure suffisamment simple pour éviter une complexité non nécessaire et une perte de temps comme dans l'approche manuelle.

Le Comité Européen de Normalisation Technical Committee 251 Working Group 2 (CEN TC 251 WG2) et plus tard l'Organisation de standardisation internationale ISO TC 215 WG3 ont élaboré et développé une approche standard pour les terminologies biomédicales appelée structure catégorielle [11].

Le Comité européen de standardisation a établi qu'il n'était pas possible de convaincre les différents pays européens de se mettre d'accord sur une terminologie clinique de référence ou de standardiser une ontologie médicale indépendante du langage. Les deux principaux arguments soutenant cette thèse étaient que les pays européens parlent des langues différentes et que les professionnels de santé qui utilisent le même langage ne donnent pas le même sens à une terminologie.

La structure catégorielle du CEN est définie comme un ensemble minimal de contraintes pour représenter une terminologie biomédicale dans un domaine précis de la santé. Ces contraintes consistent en 1) l'objectif de la structure catégorielle ; 2) une liste de catégories sémantiques ; 3) la liste des liens sémantiques entre les catégories sémantiques autorisées avec leurs catégories sémantiques associées ; 4) les contraintes minimales qui permettent la génération et la validation de phrases terminologiques bien formées.

Notre objectif est de décrire une structure catégorielle pour les EIM. Ce travail est limité à l'étude de WHO-ART. Dans la section suivante, nous exposons comment les structures catégorielles se positionnent par rapport aux recherches actuelles sur les ontologies et leur modélisation en logique de description et nous présentons une liste des structures catégorielles existantes. Dans la troisième section, nous décrivons notre ensemble de termes et la méthodologie de développement de la structure catégorielle. Dans la quatrième section, nous présentons les catégories et relations sémantiques que nous avons identifiées. Enfin, nous discutons les limites de ce travail et les perspectives.

2 État de l'art

2.1 Positionnement des structures catégorielles par rapport aux travaux sur les ontologies

Les termes et les concepts qui leur sont associés sont des éléments importants pour la représentation des connaissances en médecine et pour la documentation des problèmes médicaux des patients. Les médecins peuvent naturellement comprendre la signification des termes et échanger cette signification avec leurs collègues grâce à leurs connaissances médicales. Quand l'information est transférée du dossier papier au dossier électronique, la signification de l'information peut être perdue au cours des différents processus informatiques. Pour donner aux ordinateurs la capacité de représenter le sens de l'information, il est obligatoire de mettre en œuvre une représentation formelle de l'information médicale. La représentation formelle est basée sur un langage formel qui permet son implémentation dans un ordinateur et dont le sens peut être interprété par un médecin.

Les logiques de description sont un ensemble de formalismes logiques qui permettent de vérifier la consistance logique d'une ontologie et de faire des raisonnements. Une structure catégorielle est constituée d'une liste de catégories sémantiques et de relations entre ces catégories mais ne contient pas la liste des concepts qui se rattachent à ces catégories. La structure catégorielle est un effort de standardisation pour laquelle le mode d'implémentation est laissé au choix du développeur. On peut donc utiliser des logiques de description pour implémenter une ontologie à partir de la structure catégorielle. L'implémentation d'une structure catégorielle en logique de description peut se rapprocher de la construction d'une *top ontology*, c'est-à-dire l'architecture de haut niveau d'une ontologie contenant les principales classes et relations qui servent à partitionner le domaine.

2 .2 Liste des structures catégorielles disponibles

Il existe déjà six standards complets ISO ou CEN pour les structures catégorielles :

- CEN EN 12 264 [12] pour la définition d'une structure catégorielle ;
- CEN EN 1828 [13] pour les interventions chirurgicales ;
- ISO 1804 [14] pour les terminologies des soins infirmiers ;
- CEN EN 12 611 [15] pour les résultats de laboratoire ;
- CEN EN 13 940 [16] pour la continuité des soins ;
- CEN EN 15 521 [17] pour l'anatomie humaine.

Par ailleurs, cinq structures catégorielles sont encore en cours de validation pour les dispositifs médicaux [18-19], les médicaments à usage humain [20], les signes vitaux [21] et les signes cliniques des patients [22]. Les six standards complets seront révisés dans les cinq prochaines années par un effort commun des deux groupes de travail du CEN et de l'ISO. Cela donnera l'opportunité d'harmoniser les libellés et les définitions et de vérifier leur bon usage au niveau international. Enfin, le manuscrit EN 15 521 pour les terminologies de l'anatomie humaine a introduit des contraintes plus étendues en prescrivant des catégories sémantiques et des liens sémantiques détaillés pour le champ de connaissance de l'anatomie qui est très important pour la plupart des terminologies biomédicales.

3 Matériel et méthodes

3.1 Termes WHO-ART

WHO-ART (version 2004) est constitué de 1 857 termes préférés organisés dans une structure hiérarchique basée sur 32 classes organes. Nous avons utilisé 639 termes correspondant aux termes préférés WHO-ART choisis pour le codage des cas de pharmacovigilance dans notre extrait de la base française (saisis par le centre régional de pharmacovigilance de l'hôpital Pompidou).

3.2 Construction de la structure catégorielle

La structure catégorielle propose un cadre pour une organisation ontologique minimum afin d'assurer la standardisation de la représentation des connaissances pour la pharmacovigilance. Après avoir décrit l'objectif de la structure catégorielle dans l'introduction, nous faisons le point ici sur les catégories et relations sémantiques, et les contraintes. La structure catégorielle doit satisfaire les quatre critères décrits ci-dessus mais peut également ajouter d'autres contraintes.

Le point de départ était un modèle ontologique destiné à représenter le domaine de la pharmacovigilance. À partir de ce modèle, nous avons identifié les catégories et liens sémantiques minimaux et les contraintes qui représentent la structure catégorielle. Les catégories et liens sémantiques, et les contraintes ont été identifiés de façon itérative. Un premier schéma a été constitué à partir d'une expérience précédente dans la modélisation ontologique des effets indésirables liés aux médicaments [8]. Puis nous avons considéré chaque terme WHO-ART dans notre échantillon en décrivant la signification des termes en fonction des catégories et de liens sémantiques dans le schéma. Nous avons ajouté de nouvelles catégories sémantiques, changé les relations jusqu'à ce que le schéma permette de représenter chaque terme de façon consistante.

4 Résultats

4.1 Liste des catégories sémantiques

La figure 1 est notre proposition pour un modèle ontologique des effets indésirables. Les boîtes grisées et les relations soulignées correspondent à l'ensemble minimal de contraintes sur la structure catégorielle. Nous avons identifié 19 catégories sémantiques (tableau 1) dans notre modèle ontologique et la structure catégorielle est constituée de 8 catégories sémantiques à l'intérieur de ce modèle. « Médicament » est le nœud principal à l'intérieur

du schéma. Quand on considère les terminologies en pharmacovigilance, ce médicament n'est jamais précisé parce qu'il serait impossible de décrire toutes les combinaisons de médicament et effet indésirable. L'Administration, par exemple « surdosage » peut être lié à « Médicament » par la relation A_circonstances. Les catégories principales sont « Examen (résultat anormal de) » (ex. ALAT augmentées, hyperbilirubinémie, etc.), « Accident » (chute, etc.), et « Trouble » (ex. hépatite, anémie, éruption, etc.). Ces catégories sont reliées à « Médicament » par le lien Relation_Avec. Un « Médicament » peut être associé à un second « Médicament » (dans le cas des interactions) et un « Trouble » peut être relié à un autre « Trouble » (ex. néphropathie diabétique).

Un trouble est associé à d'autres catégories sémantiques : « Être vivant » (quand le médicament est associé à un effet immunosuppresseur), « Modalité » (par exemple au site d'injection, local, ou au site d'application), « Morphologie » (ex. ulcère, hémorragie), « Gravité » (bénin, grave, etc.), « Anatomie » (cellule, foie, etc.), « Évolution » (aigu, chronique, etc.) et « Fonction anormale » (physique et mentale). Il y a plusieurs types de troubles : « Trouble structurel », « Trouble fonctionnel », et « Trouble mixte » (structurel et fonctionnel). Un « Syndrome » est constitué de plusieurs troubles (par exemple Syndrome de Lyell).

Un « Accident » (ex. une chute) peut avoir un « Trouble » comme conséquence (ex. fracture). Les « Examens » sont reliés par le lien est_test_diagnostique_pour avec « Trouble » et un « Trouble » peut être un test diagnostique pour un autre « Trouble ». Les « Résultats d'examen » sont décrits au moyen de trois catégories sémantiques : « Milieu », « Mesure » et « Valeur ». Par exemple « EEG anormal » a_mesure « EEG » et a_valeur « anormale ».

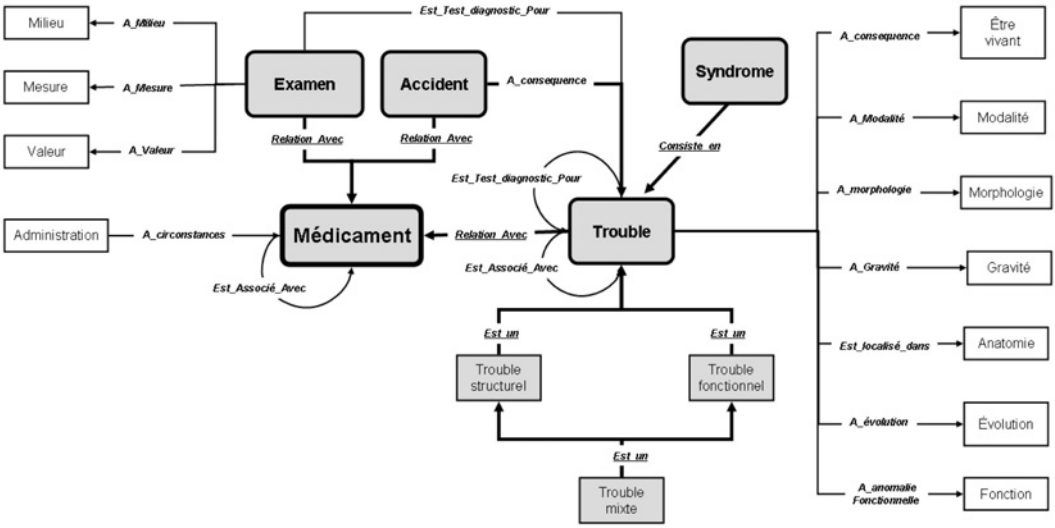


Figure 1 - Proposition pour un modèle ontologique des effets indésirables. Les catégories sémantiques sont affichées en grisé et les relations sont surlignées.

Dans la figure 1, le terme « médicament » n'est pas un exemple mais la seule valeur autorisée pour cette catégorie sémantique :

- Quand on considère les terminologies en pharmacovigilance, ce médicament n'est jamais précisé parce qu'il serait impossible de décrire toutes les combinaisons de

médicament et effet indésirable. On peut par exemple citer les termes suivants de WHO-ART : *Drug eruption*, *Drug addiction* ou *Drug fever*.

- Si on considère un cas de pharmacovigilance, le médicament est connu mais la représentation correspondante est un modèle du cas de pharmacovigilance et pas un modèle de la terminologie des EIM. La figure 2 est un exemple d'un cas de pharmacovigilance qui associe amoxicilline et acide clavulanique, et trois EIM : deux troubles (hépatite cholestatique aiguë et ictère) et un examen (hyperbilirubinémie). Cette figure illustre donc une utilisation complémentaire de la structure catégorielle ne correspondant pas à l'objectif déclaré de cette structure.

Tableau 1 - Liste des catégories sémantiques.

Catégorie sémantique	Exemple
Accident	Chute
Administration	Surdosage
Évolution	Chronique
Trouble	Hépatite
Médicament	Médicament
Fonction (anormale)	Tachycardie
Trouble fonctionnel	Hallucinations
Résultat d'examen	EEG anormal
Être vivant	*Escherichia coli*
Localisation	Foie
Mesure	EEG
Trouble mixte	Hépatite cholestatique
Modalité	Au site d'injection
Morphologie	Inflammation
Gravité	Grave
Milieu	Sang
Trouble structurel	Ulcère gastrique
Syndrome	Syndrome de Lyell
Valeur	Anormale

4.2 Liste de liens sémantiques

Nous avons identifié 16 liens sémantiques : Consiste_en, A_Circonstances, A_conséquence, A_évolution, A_anomalie_fonctionelle, Est_localisé_dans, A_Mesure, A_Modalité, A_Morphologie, A_gravité, A_milieu, A_valeur, Est_associé_avec, Est_un, Est_test_diagnostique_pour, et Relation_avec.

4.3 Contraintes minimales

L'ensemble des contraintes minimales était le suivant :

1) L'effet indésirable peut être classé comme un trouble, un accident, un examen ou un syndrome.

2) Un trouble structurel est défini par au moins une localisation et une morphologie.

3) Un trouble fonctionnel est défini par au moins une fonction anormale.

4) Il y a au moins un lien sémantique *Relation_Avec* et une catégorie sémantique « Médicament ».

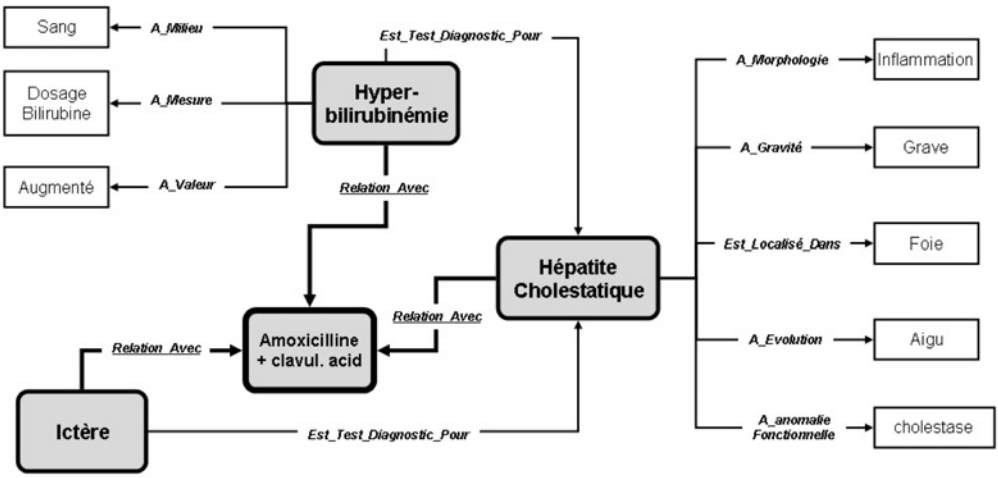


Figure 2 - Exemple d'un cas de pharmacovigilance avec hépatite cholestatique aiguë, hyperbilirubinemie, et ictère.

5 Discussion et conclusion

Ce travail est une étape préliminaire vers la proposition d'une structure catégorielle pour les EIMs. La modélisation des EIMs est différente de la modélisation d'autres terminologies dans des domaines plus précis, par exemple les interventions chirurgicales. Il existe plusieurs catégories sémantiques à représenter : Troubles, Examens, Accident, etc.

La relation causale présente une signification spéciale en pharmacovigilance. D'après Meyboom *et al.*, « l'évaluation de la causalité n'élimine ni ne quantifie le caractère incertain mais, au mieux, le catégorise de façon semi quantitative » [23]. Des méthodes ont été proposées pour l'imputation des EIMs [24]. Quand on considère un effet indésirable, les circonstances d'administration (ex. un surdosage) ou la modalité (la lésion est proche du site d'administration) sont importantes pour l'imputation de l'EIM au médicament. C'est la raison pour laquelle il était important d'introduire les catégories sémantiques « Administration » et « Modalité » dans le modèle. Cependant d'autres critères sont importants pour l'imputation tels que la relation temporelle mais n'ont pas été introduits dans le modèle parce qu'ils ne faisaient pas partie des concepts WHO-ART dans notre échantillon.

Nous avons décidé d'utiliser le lien Relation_avec entre « Médicament » et les principales catégories (« Trouble », « Examen », et « Accident ») plutôt que A_cause pour deux raisons :

1) On ne peut jamais être sûr de la relation causale entre un médicament et un effet indésirable.
2) Il existe plusieurs types de relations entre un médicament et un effet indésirable. On

établit une relation temporelle quand un patient présente un choc anaphylactique quelques minutes après l'injection d'un médicament. Une relation statistique peut être observée : par exemple le risque de chute est augmenté chez les patients traités avec certains médicaments par rapport aux autres patients [25].

5.1 Limites

Nous avons étudié un échantillon de termes WHO-ART utilisés pour le codage des cas dans un seul centre de pharmacovigilance. Certains termes spécifiques pourraient être absents et nécessiter une représentation ad hoc. Par exemple l'inefficacité thérapeutique est un type spécial d'EIM [26]. Nous avons trouvé plusieurs cas d'héparine qui étaient codés en « thrombose » plutôt qu'avec le terme « inefficacité thérapeutique ». Il n'est pas certain que la catégorie sémantique « Modalité » englobe de telles observations ou si une nouvelle catégorie sémantique est nécessaire.

Par conséquent, le modèle peut évoluer quand de nouveaux termes sont définis et une évaluation complète est nécessaire pour prendre l'ensemble de la terminologie en compte. Cette évaluation devrait bénéficier de la classification des termes au moyen d'un raisonneur afin de vérifier si le modèle est consistant. Nous prévoyons de considérer des catégories sémantiques telles que « Fonction anormale » afin de faire la différence entre la fonction physiologique (par ex. rythme cardiaque) et l'anomalie (ex. tachycardie). Les maladies et les signes ou symptômes sont en général séparés dans les autres ontologies et cela peut être nécessaire dans le modèle final afin d'assurer l'interopérabilité.

Ce travail était limité à l'étude de WHO-ART et nécessite des recherches supplémentaires afin de mettre à jour les caractéristiques supplémentaires nécessaires pour la représentation des termes MedDRA. Contrairement à WHO-ART, MedDRA inclut des classes organes spécifiques aux examens et au contexte social alors que les maladies et les examens sont regroupés à l'intérieur de la même classe organe dans WHO-ART.

Certaines parties spécifiques ne sont pas développées et pourraient bénéficier de structures catégorielles déjà publiées. Par exemple il existe une « Structure catégorielle des terminologies d'anatomie humaine » qui pourrait compléter la catégorie sémantique « Localisation » [17]. La structure catégorielle sur « Représentation des types de propriétés dédiés aux examens de laboratoire » pourrait compléter la catégorie sémantique « Résultats d'examens » [15].

5.2 Perspectives

La 11e révision de la Classification Internationale des Maladies (CIM) est un projet important pour l'Organisation mondiale de la santé dans le domaine terminologique et sera basée sur une structure catégorielle ou un modèle similaire. WHO-ART qui est développé par un centre de l'OMS nécessite ce type de modèle afin d'assurer l'interopérabilité avec la future version de la CIM-11.

Ce travail fait partie du projet VigiTermes qui a pour objectif de construire des méthodes et des outils pour une meilleure détection du signal en utilisant des méthodes d'ingénierie des connaissances. L'outil PharmARTS permet la recherche de cas de pharmacovigilance dans une base de données en fonction de critères sémantiques [6]. Il contient déjà des définitions formelles sur les effets indésirables extraits à partir de SNOMED-CT qui incluent de l'information sur la localisation anatomique, la morphologie, la fonction et l'être vivant. Nous proposons d'implémenter dans PharmARTS les catégories sémantiques décrites dans cet article pour augmenter le nombre et la performance des requêtes disponibles dans une base de pharmacovigilance au moyen de cet outil.

Remerciements

Le projet VigiTermes bénéficie d'un financement par l'Agence Nationale pour la Recherche dans le programme Technologies pour la Santé ANR-07-TECSAN-026. Nous remercions le Dr Anand Kumar pour son aide à la mise au point de cette structure catégorielle.

Références

[1] http://afssaps.sante.fr/pdf/5/glossaire_vigilances.pdf (consulté le 25/11/2008).

[2] Art. R. 5121-153 du Code de la santé publique.

[3] Uppsala Monitoring Center. *The WHO-ART Adverse Reaction Terminology*. 2003 http://www.umc-products.com/graphics/3149.pdf (last accessed Mar. 14, 2008).

[4] Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf* 1999 ; 20(2): 109-17.

[5] Bousquet C, Lagier G, Lillo-Le Louët A, Le Beller C, Venot A, Jaulent MC. Appraisal of the MedDRA conceptual structure for describing and grouping adverse drug reactions. *Drug Saf* 2005 ; 28(1): 19-34.

[6] Bousquet C, Henegar C, Lillo-Le Louët AL, Degoulet P, Jaulent MC. Implementation of automated signal generation in pharmacovigilance using a knowledge-based approach. *Int J Med Inform* 2005; 74(7-8): 563-71.

[7] Alecu I, Bousquet C, Degoulet P, Jaulent MC. PharmARTS: terminology web services for drug safety data coding and retrieval. *Medinfo* 2007 ; 12(Pt 1): 699-704.

[8] Henegar C, Bousquet C, Lillo-Le Louët A, Degoulet P, Jaulent MC. Building an ontology of adverse drug reactions for automated signal generation in pharmacovigilance. *Comput Biol Med* 2006 ; 36(7-8): 748-67.

[9] Iavindrasana J, Bousquet C, Jaulent MC. Knowledge acquisition for computation of semantic distance between WHO-ART terms. *Stud Health Technol Inform* 2006; 124: 839-44.

[10] Alecu I, Bousquet C, Jaulent MC. A case report: using SNOMED CT for grouping Adverse Drug Reactions Terms. *BMC Med Inform Decis Mak* 2008; 8 Suppl 1: S4.

[11] Rodrigues JM, Kumar A, Bousquet C, Trombert B. Standards and Biomedical Terminologies: the CEN TC 251 and ISO TC 215 Categorial Structures. A step towards increased interoperability. *Stud Health Technol Inform* 2008; 136: 857-62.

[12] CEN AFNOR EN NF 12264:2005. *Health informatics – Categorial Structure for concept systems.*

[13] CEN AFNOR EN NF 1828:2002. *Health informatics – Categorial Structure for classifications and coding systems of surgical procedures.*

[14] EN ISO 18104, 2003 – *Health informatics – Integration of a reference terminology model for nursing.*

[15] EN 1614: 2006, *Health Informatics – Representation of dedicated kinds of property in laboratory medicine.*

[16] EN 13 940: 2006, *Health Informatics – System of concepts to support continuity of care.*

[17] EN 15 521:2007. *Health informatics – Categorial Structure for terminologies of human anatomy.*

[18] ENV 12 611: 1996. *Medical Informatics – Categorial structure for system of concepts - medical devices.*

[19] ISO CD 15225:2000 – *Nomenclature – Specification for a nomenclature system for medical devices for the purpose of regulatory data exchange.*

[20] ENV 12 610:1997 *Medical Informatics – Medicinal products identification.*

[21] ENV 13734:2000 *Health Informatics – Vital signs information representation.*

[22] ISO/DTS 22789:2007, *Health informatics – Conceptual framework for patient findings and problems in terminologies.*

[23] Meyboom RH, Hekster YA, Egberts AC, Gribnau FW, Edwards IR. Causal or casual? The role of causality assessment in pharmacovigilance. *Drug Saf* 1997; 17(6): 374-89.

[24] Bégaud B, Evreux JC, Jouglard J, Lagier G. Imputation of the unexpected or toxic effects of drugs. Actualization of the method used in France. *Therapie*. 1985; 40(2): 111-8.

[25] Souchet E, Lapeyre-Mestre M, Montastruc JL. Drug related falls: a study in the French Pharmacovigilance database. *Pharmacoepidemiol Drug Saf* 2005; 14(1): 11-6.

[26] Figueras A, Pedrós C, Valsecia M, Laporte JR. Therapeutic ineffectiveness: heads or tails? *Drug Saf* 2002; 25(7): 485-7.

Adresse de correspondance

Cédric Bousquet
Service de Santé Publique et de l'Information Médicale
Bâtiment CIM 42 - Hôpital Nord
Chemin de la Marandière
42270 Saint-Priest-en-Jarez
E-mail : cedric.bousquet@chu-st-etienne.fr

Informatique et Santé, 2009 (17) : 101-112
© Springer – Verlag France

Recherche d'information multiterminologique en contexte : étude préliminaire

Ahmed-Diouf Dirieh Dibad[1, 2], Saoussen Sakji[1, 2,4], Élise Prieur[2], Suzanne Pereira[2, 3, 4], Michel Joubert[4], Stéfan J. Darmoni[1, 2]

[1]*CISMeF, Centre Hospitalier Universitaire de Rouen*
[2] *TIBS, LITIS EA 4108, Université de Rouen, France*
[3]*VIDAL, Issy-les-Moulineaux, France*
[4]*LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille, France.*

Abstract

Background: Information retrieval becomes more relevant thanks to document (semi)-automatic indexation with terminologies adapted to the domain. Being interested in the medical domain, we used the main health terminologies available in French in two different contexts: (a) during the health resource indexation process and (b) during the information retrieval. Methods: Within the CISMeF catalog (French acronym of Catalogue et Index des Sites Médicaux Francophones), we have evolved from a mono-terminological world (based on the MeSH thesaurus) to a multi-terminological universe (based on ICD10, CCAM, SNOMED terminologies and ATC classification). This new structure allows us to make the information retrieval in two different contexts: to retrieve (a) reports from the electronic health record (EHR) of patients and (b) resources from a health portal (such as CISMeF). Results: Both contexts share the same representation meta-model, however, they differ during the information retrieval, according to the size and the nature of the corpus, the role of dates and the query, which can include both structured and unstructured elements in EHR. Conclusion: In 2008, the information retrieval has already become multi-terminological in the Drug Information Portal and will become available in the catalog CISMeF at the beginning of 2009.

Keywords

International Classification Diseases; Medical Subject Headings; SNOMED; Information storage and retrieval; Terminology; Controlled vocabulary; UMLS.

1 Introduction

L'avancée des recherches sur le traitement automatique du langage naturel (TAL) et l'extraction et l'indexation automatique de termes (*semantic mining*) a fait évoluer le traitement de l'information et des connaissances. Différentes approches ont été utilisées pour obtenir une couverture du domaine médical [1] (terminologie, ontologie, base de connaissances). Pour ce faire, un nombre très important de vocabulaires (thésaurus, classification, nomenclature, etc.) ont été développés en santé : ainsi, à titre indicatif,

l'UMLS (*Unified Medical Language System*), méta-thésaurus développé à la US National Library of Medicine, en recense plus de 130, dont 10 seulement sont traduits en français. Certaines de ces terminologies standardisées et normalisées ont permis de représenter et de décrire les contenus des documents dans un but de recherche d'information (RI). La finalité de construction et les types d'utilisation de ces terminologies sont nombreux et dépendent du contexte [1].

Notre objectif est d'étudier la recherche d'information multi terminologique dans un double contexte, d'une part le contexte de connaissance dans un portail d'information de santé (en l'occurrence le catalogue CISMeF [2]) et d'autre part le contexte de soins dans un dossier électronique du patient (DEP). Deux doctorants travaillent sur ce sujet (AD & SS).

Ce papier met en relief le passage d'un monde mono terminologique à un univers multi terminologique sur la recherche d'information et explicite les avantages de la projection de ces terminologies sur ces deux contextes selon plusieurs axes (finalité, type d'utilisation). Le principal résultat attendu est la maximisation du rappel en s'appuyant sur plusieurs terminologies et des liens sémantiques entre elles. En effet, plusieurs terminologies de santé disponibles en français seront utilisées dans le cadre de nos travaux : le thésaurus MeSH, les classifications CIM10 et CCAM et la nomenclature SNOMED International et d'autres qui seront intégrés prochainement (Orphanet, LOINC, Who-Art). L'ensemble de ces terminologies seront utilisées pour indexer des ressources de santé en langue française ou les documents principaux d'un DEP (compte-rendu d'hospitalisation ou compte-rendu d'actes médicaux, comme un scanner, etc).

Nous allons décrire l'architecture de notre système de recherche d'information (SRI) dans cet univers multi terminologique et présenter la recherche d'information en contexte sous différents points de vue.

2 Matériel et méthodes

2.1 Passage d'un monde mono terminologique à un univers multi terminologique

2.1.1 Catalogue CISMeF – mono terminologique

CISMeF (acronyme de Catalogue et Index des Sites Médicaux Francophones, URL : http://www.chu-rouen.fr/cismef) [2] est un catalogue décrivant et indexant les principales sources d'information institutionnelles de santé françaises (N ≈ 64 000), ayant également un système de recherche d'information médicale (Doc'CISMeF). CISMeF est destiné aux professionnels de santé, aux étudiants mais aussi aux patients, leur famille, et d'une façon encore plus large au cyber-citoyen.

Le système Doc'CISMeF a été longtemps fondé exclusivement sur un monde mono terminologique, en se basant sur le thésaurus MeSH (*Medical Subject Headings*). Celui-ci a été créé par la US National Library of Medicine et est utilisé pour indexer les articles scientifiques de la base de données bibliographiques MEDLINE. Chaque ressource du catalogue est indexée par son contenu utilisant plusieurs ensembles de méta-données et par son contenant en utilisant les types de ressources de la terminologie CISMeF [3]. Les méta-données se réfèrent aux informations descriptives des ressources Web et ont pour finalité de faciliter et d'améliorer la recherche d'information. Dans CISMeF, elles sont essentiellement sous la forme du Dublin Core (DC) [4] ; CISMeF utilise 11 champs parmi les 15 du DC : auteur, site éditeur, description, langage, mots clés, type de ressources...

La terminologie CISMeF encapsule la version française du thésaurus MeSH [3] dans la mesure où elle représente une extension des concepts déjà existants, d'une part, et elle

emploie de nouveaux concepts, d'autre part. Dans le but de généraliser le MeSH aux ressources de santé sur Internet, des améliorations ont été réalisées depuis la création du catalogue en 1995. En plus des descripteurs (termes qui permettent l'indexation des ressources ; plus de 25 000 descripteurs dans le MeSH 2009) et des qualificatifs MeSH au nombre de 83 (qui permettent de préciser le sens d'un descripteur, et d'en souligner un aspect particulier), les notions de méta-termes et de types de ressources ont été ajoutées. La paire (descripteur/qualificatif) décrit le centre d'intérêt de la ressource. Les types de ressources sont une extension des types de publication de MEDLINE et sont utilisés afin de catégoriser la nature du contenu de la ressource (N = 293). Les méta-termes (MT) sont des super-concepts, qui ont été initialement définis pour agréger entre eux un ensemble de termes MeSH qui appartiennent à la même spécialité médicale (par exemple la cardiologie), ou des types de procédures (par exemple la neurochirurgie) ou différents aspects généraux de la santé (comme le diagnostic ou la thérapeutique). Les méta-termes ont été sélectionnés manuellement par le conservateur des bibliothèques de l'équipe CISMeF (Benoit Thirion) qui a créé manuellement des liens sémantiques entre les MT et un ou plusieurs termes MeSH, qualificatifs et types de ressources. Une étude récente de notre équipe a montré que l'utilisation des MT améliorait la recherche d'information [5] Par exemple, introduire le terme « psychiatrie » comme méta-terme constitue une stratégie plus efficace pour avoir plus de résultats : au lieu d'exploser une seule hiérarchie MeSH correspondant au terme MeSH « psychiatrie », on peut obtenir une expansion automatique de la requête en explosant les hiérarchies de tous les termes sémantiquement reliés à ce MT « psychiatrie », comme le terme MeSH « hôpital psychiatrique » ou le type de ressource « dispensaire de santé mentale ».

2.1.2 Univers multi terminologique

Dans un objectif d'une recherche d'information plus exhaustive et plus efficace, l'équipe CISMeF a pris la décision stratégique de passer d'un monde mono terminologique à un univers multi terminologique. Cet objectif doit prendre en compte la disponibilité de plusieurs terminologies, classifications, thésaurus et nomenclatures médicaux disponibles en français[1], et la mise en pratique des interactions existantes entre ces derniers (figure1).

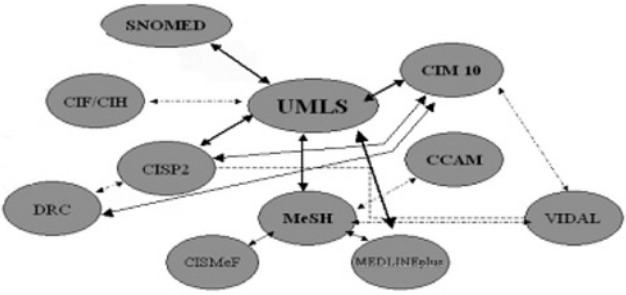


Figure 1 - Relations existantes entre les terminologies médicales.

En effet, grâce aux relations terminologiques, on peut améliorer la recherche d'information et mieux répondre à la requête de l'utilisateur et ce via l'expansion ou la reformulation de la requête. On peut chercher toutes les liaisons possibles entre les termes de la requête appartenant à une terminologie donnée et tous les termes des autres terminologies qui ont une correspondance avec les termes en question. Cette procédure permet d'élargir le

[1] Rappelons que 130 terminologies de santé existent en anglais contre une dizaine en français.

champ de la recherche de l'utilisateur selon son contexte, sans néanmoins mettre en cause la pertinence thématique de l'information ni le degré de précision du système. Par exemple, grâce à la correspondance entre le terme MeSH « appareil correction auditive » et le terme SNOMED « prothèse auditive », on peut enrichir notre résultat et retrouver toutes les ressources indexées non seulement par le terme « appareil correction auditive » mais aussi par le terme « prothèses auditives ».

2.2 Serveur multi terminologique de Santé (SMTS)

Le passage à un univers multi terminologique se traduit par la mise en place d'un serveur multi terminologique de santé (SMTS) [6], développé par CISMeF, LERTIM et la société Mondeca puis poursuivi dans le cadre du projet InterSTIS (Interopérabilité Sémantique des Terminologies dans les Systèmes d'Information de Santé Français)[2]. Ce serveur repose sur un logiciel de gestion de terminologies et d'ontologies (ITM). Grâce aux parseurs générant le format OWL développés spécifiquement pour ce projet, les terminologies de santé en français sont intégrées : MeSH, SNOMED (*Systematized NOmenclature of MEDicine*), CIM-10 (Classification Internationale des Maladies, version10), CCAM (Classification Commune des Actes Médicaux), CISP2 (Classification Internationale des Soins Primaires, deuxième édition), DRC (Dictionnaire des Résultats de Consultation), CIF (Classification Internationale du Fonctionnement, du handicap et de la santé). Ces parseurs prennent en compte ces terminologies quelles que soient leurs formes d'origine (forme SQL, fichiers texte, bases de données...) et les regroupent dans une même structure (figure 2). Les relations inter et intra terminologiques permettent d'assurer la navigation entre les terminologies.

L'outil F-MTI (*French Multi-Terminology Indexer*) [7] [8] a été développé pour indexer automatiquement et génériquement les ressources médicales, qu'elles soient des comptes rendus médicaux ou des ressources Web. Cet outil s'appuie sur les terminologies de santé parsées pour le SMTS (figure 2). F-MTI fonctionne en deux temps : une extraction des concepts des terminologies étudiées, puis une restriction vers les terminologies choisies. Cette restriction s'effectue via les relations de transcodage (relations de correspondance entre les terminologies).

2.3 Le modèle générique

Pour faciliter l'accès à toutes ces terminologies, nous allons développer un Portail de Terminologies de Santé (PTS) qui représente une porte d'entrée à ces dernières, néanmoins, sans se soucier ni de la gestion ni de la mise à jour (ce qui est le cas pour le SMTS) (figure 2).

Pour réunir dans une même structure des terminologies, des thésaurus, des nomenclatures et des classifications de natures différentes ayant chacun des spécificités particulières, il a fallu, d'abord, modéliser chacune de ces terminologies. Dans un second temps, nous avons créé un modèle générique qui tient compte de toutes ces particularités, et qui est assez générique pour avoir la possibilité d'intégrer d'autres terminologies. Pour ce faire, le modèle générique obtenu est décrit ci après (figure 3).

---

[2] Projet financé par l'ANR (appel d'offres TecSan 2007) ; URL : http://cybertim.timone.univ-mrs.fr/recherche/projets-recherche/INTERSTIS/

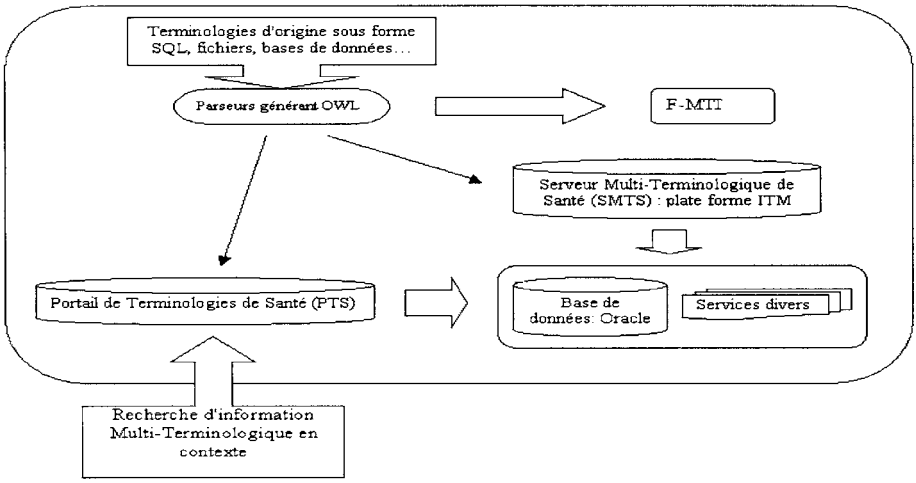


Figure 2 - Architecture de notre SRI.

Pour implémenter dans CISMeF la recherche d'information multi terminologique, il est nécessaire d'intégrer tous ces modèles dans le back-office de CISMeF. Cette implémentation pourra utiliser les outils sémantiques d'Oracle, notamment le langage SPARQL, avec gestion des triplets RDF.

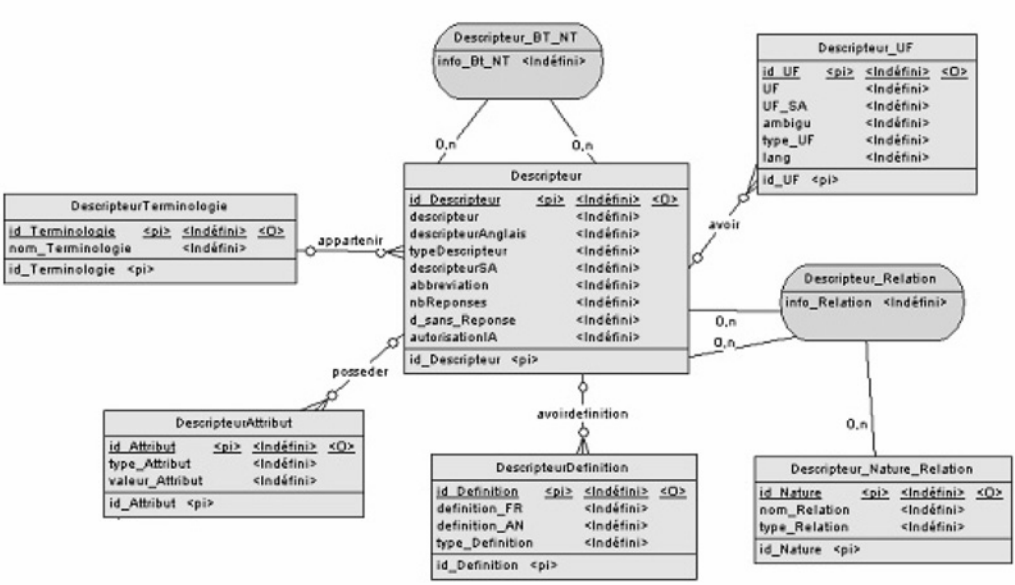


Figure 3 - Le modèle générique dans le cadre de la recherche d'information multi terminologique.

Ce modèle est centré sur l'entité **Descripteur**. Celle-ci inclut tous les termes, les mots clés, les qualificatifs, les types de ressources, les méta-termes, les éléments, les noms de

catégories, les groupes, les blocs, les chapitres qui peuvent exister au niveau des terminologies. Cette classe définit les attributs communs aux différentes terminologies. Les attributs spécifiques sont représentés par l'entité **DescripteurAttribut**, ce qui permet d'être plus générique et plus souple lors de la mise à jour du modèle. Les définitions des descripteurs sont multilingues et de types différents (DRC, MeSH, Vidal...) et sont décrites par l'entité **Descripteur_Definition**. Les synonymes sont illustrés par l'entité **Descripteur_UF**. N'étant pas considérés comme des termes d'indexation, les synonymes permettent de représenter un terme de plusieurs façons et par la suite un enrichissement de la requête de l'utilisateur lors du processus de recherche d'information.

Un descripteur peut avoir plusieurs synonymes. Chaque descripteur appartient à une des terminologies intégrées. Ces dernières sont représentées par l'entité **DescripteurTerminologie**, dont la structure est définie par l'identifiant de la terminologie (id_Terminologie) et le nom de la terminologie (nom_Terminologie). La relation **Descripteur_BT_NT** est définie entre deux descripteurs et réservée pour les relations hiérarchiques au sein d'une même terminologie, à l'inverse de l'association Descripteur_Relation qui décrit les relations non hiérarchiques intra-terminologiques (si les deux descripteurs appartiennent à la même terminologie) et les relations inter-terminologiques (si les deux descripteurs appartiennent à des terminologies distinctes). Ces derniers types de relations permettent de relier des terminologies entre elles, inspiration faite du réseau sémantique et du méta-thésaurus d'UMLS. L'entité **Descripteur_Nature_Relation** a pour but de définir les types de relations non hiérarchiques, elle contient les noms et les types des toutes les relations existantes entre deux descripteurs quelconques appartenant à la même terminologie ou non telles que : ne pas confondre, voir aussi, inclusion, exclusion... Cette classe substitue la classe *Relation* du modèle générique du fait qu'elle définit le lien entre deux autres entités.

3 Résultats

3.1 Recherche d'information multi terminologique (RIMT) en contexte

De manière générale, un processus de RI est constitué de 3 éléments : l'utilisateur, le corpus et un système de recherche d'information (SRI). L'utilisateur pourra être tout acteur intervenant dans le domaine médical. Cet utilisateur veut accéder par le contenu à une ressource de notre corpus laquelle pourrait être une ressource Web (catalogue CISMeF) ou une ressource clinique (dossier électronique du patient). Dans le cas du dossier patient, ces ressources sont indexées par une ou plusieurs des terminologies francophones disponibles qui constituent notre modèle d'indexation. Le but du SRI est de fournir à l'utilisateur un ensemble de ressources jugées pertinentes en réponse à la requête de l'utilisateur. Ce système intègre un moteur de recherche qui se fonde sur un algorithme d'appariement qui fait correspondre les concepts de la requête utilisateur avec les concepts d'une ou de plusieurs terminologies indexant les ressources du corpus.

L'algorithme d'appariement utilisé dans notre RI est l'algorithme de sac de mots. Il a été implémenté dans l'outil MAIF [9] puis amélioré et optimisé pour un environnement multi terminologique dans l'outil F-MTI [8]. Le principe de l'algorithme consiste en une simple comparaison syntaxique entre des termes normalisés par une méthode de lemmatisation ou de désuffixation. Ces termes correspondent d'une part à des termes extraits de la requête de l'utilisateur et d'autre part aux éléments désuffixés (« stems ») ou lemmatisés des concepts du modèle d'indexation.

Techniquement, les avantages sont nombreux avec une recherche d'information plus efficace en termes de rappel [8]. Le rappel mesure la proportion de documents pertinents

restitués par le système relativement à l'ensemble des documents pertinents contenus dans la collection de documents. Dans cet univers multi terminologique et avec les différentes méthodes d'expansion de requêtes [10], de nouvelles requêtes seront construites par le moteur de recherche à partir des concepts de la requête de l'utilisateur. Par exemple, actuellement dans CISMeF, la requête « Wyburn Mason » n'est pas reconnue comme un terme appartenant à la terminologie CISMeF (essentiellement le MesH). Depuis janvier 2009, cette même requête est reconnue comme un terme appartenant à deux terminologies du SMTS : Orphanet d'une part et MedDRA d'autre part. Ainsi, avec ce principe et dans le cadre de la recherche documentaire multi terminologique, on peut accéder aux ressources indexées par « Wyburn Mason », un résultat qui ne pouvait être détecté dans la base mono terminologique grâce aux descripteurs MeSH « Wyburn » et « Mason ». Cependant, une base multi terminologique et une indexation avec le terme en complet (cooccurrence des termes) offre la possibilité d'accéder, en premier lieu, à la bonne ressource, ensuite, à la définition du terme dans sa terminologie d'origine et par conséquent naviguer entre ses relations hiérarchiques et inter terminologiques si elles existent. Ainsi, on peut faire une recherche complémentaire vers d'autres sites tel que PubMed.

Un exemple en recherche documentaire déjà opérationnel de RIMT autour du projet CISMeF est la réalisation d'un Portail d'Information sur le Médicament (PIM) [11] en collaboration avec la société Vidal, dans le cadre du projet Européen PSIP [12]. Ce portail permet un accès par nom commercial, par dénomination commune internationale, par le thésaurus MeSH (y compris les Supplementary Concepts (SC)[3]), les codes ATC, les codes CAS et les différents codes sur les médicaments utilisés en France (CIS, CIP, UCD). Notre portail constitue, ainsi, un service dédié aux utilisateurs du domaine médical pour accéder aux données médicamenteuses disponibles sur Internet. Il s'agit d'un portail d'information bilingue (français/anglais) sur le médicament, dans un contexte multi terminologique et qui est relatif à des informations émanant d'institutions ou de sociétés savantes permettant d'améliorer la recherche d'information de l'utilisateur sur le médicament.

Cependant, la RIMT peut être présentée différemment dans ce double contexte (soins au sein du DEP, ressource Web au sein d'un portail comme CISMeF), et à plusieurs niveaux : (a) au niveau de la taille du corpus (plusieurs millions de comptes-rendus dans la base des DEP du CHU de Rouen à opposer aux 64 000 ressources de CISMeF) ; (b) au niveau des utilisateurs : au sein de CISMeF, il est possible de définir des sous-ensembles pour le patient, pour l'étudiant ou pour le médecin (grâce aux types de ressources respectifs : patient, matériel enseignement et recommandation). Il existe également des différences importantes en termes de recherche d'information suivant le double contexte : DEP ou portail. Au sein d'un DEP, la recherche d'information repose à la fois sur des données non structurées a priori, comme un concept médical extrait d'un compte-rendu par l'outil F-MTI (structuration a posteriori) et des données structurées a priori (comme un résultat biologique). Il faut souligner l'importance des dates dans un DEP. Enfin, il est important pour l'utilisateur de différencier si l'indexation d'un concept médical a été réalisée manuellement par un médecin ou automatiquement (par l'outil F-MTI pour ce travail). Voici un exemple de requête possible au sein du DEP : *infarctus du myocarde[CIM 10][manuel] ET bilirubine > 20 ET (même hospitalisation –même identifiant d'hospitalisation-).* Cette requête se traduit en langage naturel : rechercher les patients qui ont un infarctus du myocarde codé manuellement en CIM 10 et dont un examen biologique de bilirubine est supérieur à 20 lors de la même hospitalisation. Cet exemple illustre l'intérêt de ce travail pour améliorer l'inclusion de patients dans des essais cliniques.

---

[3] Les SC représentent les substances chimiques, ils sont non hiérarchisés, reliés à des mots clefs et interviennent dans l'indexation.

3.2 Espace de représentation (ou d'indexation)

De manière générale, l'objectif de la recherche d'information reste le même dans les deux contextes. Ceux-ci partagent le même modèle de représentation (ou d'indexation). En effet, le modèle d'indexation, pour indexer les ressources Web dans CISMeF ou pour coder les informations cliniques dans le dossier médical du patient, se fondent sur les mêmes terminologies francophones.

Dans le domaine médical, les informations sont de différentes sources et donc plusieurs terminologies sont appropriées suivant le besoin de l'utilisateur. Donc, les objectifs et les types d'utilisation diffèrent selon le contexte. En effet, le codage des données cliniques permet une évaluation médico-économique de l'activité hospitalière et de la qualité du soin du patient. Comme par exemple la classification CCAM faite pour coder les actes médicaux et utilisée dans le cadre de la tarification à l'activité (T2A) et la classification CIM10 pour les diagnostics dans un objectif d'étude épidémiologique et de santé [13]. Nous pouvons également citer la SNOMED pour un codage descriptif du DEP, le MeSH pour l'accès à la documentation scientifique, MedDRA et WHO-ART pour les effets secondaires des médicaments.

Le tableau 1 décrit dans les deux contextes (DEP ou Portail) si les principales terminologies francophones ont été utilisées manuellement (par un indexeur – documentaliste pour le MeSH, médecin pour la CCAM ou la CIM10) ou automatiquement (par l'outil F-MTI pour ce travail).

Tableau 1 : Type d'indexation des terminologies de santé francophones dans notre modèle RIMT.

Terminologies \\ Contexte	Documentaire (Portail)	Soins (DEP)
MeSH	Manuel OU EXCLUSIF Auto	Auto
CIM-10	Auto	Manuel OU Auto
CCAM	Auto	Manuel OU Auto
SNOMED IT	Auto	Auto

4 Discussion

Nous avons présenté dans ce travail les principes de la recherche d'information multi-terminologique dans différents contextes. Nous proposons notamment un modèle générique qui inclut les principales terminologies de santé en français. Ce modèle ne devrait que peu varier avec l'apport de nouvelles terminologies planifiées dans un avenir proche (LOINC, MedDRA, WHO-ART).

Pour le cas du dossier médical, le codage en texte libre est plus apprécié (car il rend possible de s'exprimer en langage naturel. Cependant, plusieurs études ont montré l'intérêt du codage structuré des données en comparaison avec le codage en mode textuel lequel est plus simple mais présente des inconvénients comme l'illisibilité, l'incomplétude et l'accessibilité [8]. Les applications spécialisées pour le codage structuré des données facilitent la RI [15-17] si bien que dans certains systèmes, les deux codages sont utilisés pour profiter des avantages des deux méthodes [18]. Toutefois, les inconvénients du codage structuré (en l'occurrence l'expressivité) sont compensés par une représentation plus efficace dans le traitement automatique des données [18]. Ce qui implique une

structuration de l'information pour mettre à l'aise l'utilisateur dans sa tâche. En effet, de la problématique entre la représentation de la connaissance et son utilisabilité résulte un conflit entre le besoin de l'utilisateur et le système [19].

4.1 Sélection et présentation de l'information

Vu le grand volume d'information à manipuler dans un dossier médical, accéder à l'information clinique sous une forme rapide et synthétisée devient impossible car le surplus d'information reste un problème majeur dans le milieu médical. Si bien que le partage d'information et la présentation des éléments du dossier médical dans un objectif d'amélioration de la qualité du soin du patient a été sujette à de nombreuses discussions. D'un point de vue fonctionnel, un médecin aura besoin pour suivre un patient d'une fiche de synthèse recueillie à partir des données se trouvant dans le dossier électronique du patient (DEP). D'où la problématique qui se pose : quelle est l'information pertinente pour un cardiologue ou un médecin généraliste dans le DEP ?
Plusieurs approches ont été implémentées pour une présentation structurée permettant de clairement identifier, capturer l'information pertinente dans l'historique du patient et ainsi aider le praticien dans son travail. En effet, la structuration de l'information facilite la RI.

Trois typologies de présentation sont utilisées dans un système électronique :

- La typologie chronologique : la plus simple et la plus utilisée où l'ensemble des documents d'un dossier médical est représenté par date. Cette vue est très intéressante lorsqu'on manipule des dossiers de faibles volumes (moins de vingt comptes-rendus) mais la RI devient difficile au-delà de cette limite (20 % des dossiers à Rouen, mais sans doute les plus consultés car ils incluent les maladies chroniques).
- L'approche Problem Oriented Medical (POM) : les dossiers médicaux orientés problèmes (chaque élément du dossier est lié à un problème médical) apparaissent en 1978. Cette approche permet d'avoir des vues conceptuelles en comparaison à l'affichage chronologique. Ces vues se fondent sur un « Gold Standard » des problèmes médicaux [20].
- La vue orientée concept dans laquelle la majorité des solutions implémentées ont utilisé des ontologies comme base de connaissances [21]- [22]. Le DEP du CHU de Rouen [23] a bénéficié du savoir-faire de l'équipe CISMeF, en utilisant les méta-termes**Erreur ! Source du renvoi introuvable.** issus de la terminologie CISMeF (MeSH) pour les adapter à la CIM10 et à la CCAM. L'objectif d'utilisation de ce super concept a été de voir comment il améliorerait la RI dans un dossier médical et que seraient les avantages d'une vue conceptuelle basée sur les spécialités médicales par rapport à la vue chronologique utilisée précédemment. Nous souhaitons dans un proche avenir étendre encore nos super-concepts à la terminologie SNOMED pour catégoriser autant que faire se peut les concepts extraits des différents compte-rendus.

Dans la même optique, d'autres solutions ont vu le jour et se résument à des techniques de synthèse automatique de documents ou à l'utilisation des annotations pour représenter les relations entre les données dans un dossier médical [25-26]. Cela dit, toutes ces solutions, pour présenter le DEP dans un objectif d'aide au codage, se fondent sur une liste de problèmes médicaux pour filtrer et/ou classer les informations cliniques et ainsi obtenir une vue conceptuelle ou une synthèse structurée. Une liste exhaustive (proposé par Campbell 98), existe et a été implémentée avec succès dans les DEP [27].

4.2 Intégration des données textuelles (non structurées) et des données structurées

Plusieurs travaux [28-29] ont été réalisés sur l'intégration des données structurées et des données non structurées. La première approche était d'implémenter un SRI utilisant un

système relationnel [30]. Par ailleurs, les fonctionnalités avancées de la RI (recherche booléenne, recherche par proximité, recherche par classement, etc.) ont toujours utilisé le langage SQL standard [29]. L'efficience des systèmes de base de données relationnels pour les données structurées est déjà connue et ils sont moins coûteux que des systèmes spécialisés RI. Donc plusieurs implémentations de cette approche ont été réalisées sur certains modèles de représentation RI (Modèle Booléen, Modèle probabiliste avec l'algèbre relationnel) et étendues vers des systèmes relationnels supportant les inférences et les agrégations [31] qui sont des méthodes importantes dans la RI. Mais l'inconvénient est que cette approche pose un problème de stockage connu dans les systèmes relationnels [28].

5 Conclusion

En collaboration avec la société Oracle dans le projet PSIP, nous étudions l'intérêt d'intégrer le langage d'interrogation SPARQL dans notre algorithme de recherche d'information. Ce travail nécessitera l'export des terminologies en OWL pour créer un ensemble de triplets RDF.

Au premier trimestre 2009, la recherche d'information deviendra multi terminologique au sein du catalogue CISMeF, en incluant toutes les terminologies du SMTS, également incluses dans l'outil F-MTI (figure 2).

Remerciements

Ce travail a été partiellement financé par le projet Européen 7e PCRD, PSIP Grant agreement n° 216130.

Références

[1] Zweigenbaum P. Encoder l'information médicale : des terminologies aux systèmes de représentation des connaissances. *Innovation Stratégique en Information de Santé* 1999 ; (2-3): 27-47 .1999.

[2] Darmoni SJ, Leroy JP, Thirion B, *et al.* CISMeF a Structured Health Resource Guide. *Methods Inf Med* 2000 ; 39(1): 30-35.

[3] Douyère M, Soualmia LF, Névéol A, Rogozan A, Dahamna B, Leroy JP, Thirion B, Darmoni SJ. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. *Health Info Libr J* 2004; 21(4): 253-61.

[4] Dekkers M., Weibel S. State of the Dublin Core Metadata Initiative, *D-Lib Mag. V9 N° 40.* 2003.

[5] Gehanno JF, Thirion B, Darmoni SJ. Evaluation of meta-concepts for information retrieval in a quality-controlled Health Gateway. *AMIA Annu Symp Proc* 2007 ; 11: 269-73.

[6] Joubert M, Dahamna B, Delahousse J, Fieschi M, Darmoni SJ. SMTS® : Un Serveur Multi-Terminologies de Santé. *JFIM* 2009.

[7] Bento Pereira S. *Indexation multi-terminologique de concepts en santé.* Thèse d'informatique. Université de Rouen; 2008. 230p

[8] Pereira S, Neveol A, Kerdelhué G, Serrot E, Joubert M, Darmoni SJ. Using multi-terminology indexing for the assignment of MeSH descriptors to health resources in a French online catalogue. *AMIA Annu Symp Proc* 2008 ; 6:586-90.

[9] Névéol A, Rogozan A, Darmoni SJ. Automatic indexing of online health resources for a French quality controlled gateway. *Information Processing & Management* 2006; 42(3): 695-709.

[10] Qiu Y, Frei HP. Concept Based Query Expansion. *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval.* 1993; pp. 160-169.

[11] Letord C, Sakji S, Pereira S, Dahamna B, Kergoulay I, Darmoni SJ. Recherche d'information multi-terminologique : application à un portail d'information sur le medicament en Europe. *Ingénierie et Recherche Biomédicale/ Biomedical Engineering and reserch, Volume In Press*, 2008.

[12] PSIP (Patient Safety through Intelligent Procedures in medication). URL: www.psip-project.eu (Dernier accès le 18 novembre 2008).

[13] Pereira S, Névéol A, Massari P, Joubert M, Darmoni SJ. Construction of a semi-automated ICD10 coding help system to optimize medical and economic coding. *Proc. MIE.* 2006 ; 845-50.

[14] Meystre S, Haug PJ. Natural language processing to extract medical problems from electronic clinical documents: Performance evaluation. *Journal of Biomedical Informatics* 2006 ; 39 , 6 : 589-599; 1532-0464.

[15] Bleeker SE, Derksen-Lubsen G, van Ginneken AM, van der Lei J, Moll HA. Structured data entry for narrative data in a broad specialty: patient history and physical examination in pediatrics. *BMC Med Inform Decis Mak* 2006; 629.

[16] Doupi P, van Ginneken AM. Physical Examination Data: A Modelling Challenge Structured physical examination data: a modeling challenge. *Stud Health Technol Inform* 2001; 84(Pt 1): 614-8.

[17] Poon AD, Fagan LM, Shortliffe EH. The PEN-Ivory project: exploring user-interface design for the selection of items from large controlled vocabularies of medicine. *Journal of the American Medical Informatics Association* 1996 ; Vol 3, 168-183.

[18] Johnson S.B., Bakken S., Dine D., Hyun S., Mendonça E., Morrison F., Bright T., Vleck T.V, Wrenn J, and Stetson P. An Electronic Health Record Based on Structured Narrative. *J Am Med Inform Assoc* 2008; 15: 54-64.

[19] Trent Rosenbloom S., Miller R.A., Johnson K.B, Elkin P.L., Brown S.H. Interface Terminologies: Facilitating Direct Entry of Clinical Data into Electronic Health Record Systems. *J Am Med Inform Assoc* 2006; 13: 277-288.

[20] Rector AL. Clinical terminology: why is it so hard? *Methods Inf* Med 1999; 38: 239-52. 1999

[21] Zeng Q, Cimino JJ, Zou KH. Providing concept-oriented views for clinical data using a knowledge-based system: an evaluation. *J Am Med Inform Assoc* 2002; 9(3):294-305.

[22] Bayegan E, Nytro O, Grimsmo A. Ontologies for knowledge representation in a computer-based patient record. Ontologies for Knowledge Representation in a Computer-Based Patient Record. *14th IEEE Int. Conf. on Tools with Artificial ICTAI.* 2002.

[23] Massari P, Smuraga I, Froment L, Boudehent S, Czernichow P, Streiff J, Baldenweck M, Hecketsweiler P. Application de gestion, les dossiers informatisés du CHU de Rouen. *Cinquièmes Journées Francophone d'Informatique Médicale.* Genève 9-10 juin 1994.

[24] Massari P, Pereira S, Thirion B, Derville A, Darmoni SJ. Use of Super-Concepts to customize electronic medical records data display. *Stud Health Technol Inform.* 2008; 136: 845-50.

[25] Afantenos SD, Karkaletsisa V, Stamatopoulosb P. Summarization from medical documents: a survey. *Artificial Intelligence in Medicine* 2005 ; 33(2): 157-177.

[26] Bringay S, Barry C, Charlet J, Krim G. Une fonctionnalité d'annotation pour le dossier patient informatisé. *Journées Francophones d'Informatique Médicale,* Lille 12-13 mai. 2005.

[27] Bainbridge M, Salmon P, Rappaport A, Hayes G, Williams J, Teasdale S. The Problem Oriented Medical Record - just a little more structure to help the world go round?, *Clinical Computing Special Interest Group (CLICSIG) of the PHCSG. Proceedings of the 1996 Annual Conference of The Primary Health Care Specialist Group of the British Computer Society.* 1996.

[28] Frieder O, Chowdhury A, Grossman D, McCabe MC. On the Integration of Structured Data and Text: A Review of the SIRE Architecture. *In proceedings of the DELOS Workshop on Information Seeking, Searching, and Querying in Digital Libraries.* 2000.

[29] Grossman DA, Frieder O, Holmes DO, Roberts DC. Integrating Structured Data and Text: A Relational Approach. *Journal of the American Society of Information Science,* 1997 ; 48(2).

[30] Campbell JR. Strategies for problem list implementation in a complex clinical enterprise. *Proceedings / AMIA Annual Fall Symposium* 2008; 285-9.

[31] Farah M, Vanderpooten D. L'Agrégation en Recherche d'Information: In *Proceedings 4e conférence francophone en Recherche d'Information et Applications (CORIA'07),* pages 125-136. 2007.

Adresse de correspondance

Professeur Stephan Darmoni

Équipe CISMeF

CHU de Rouen

Laboratoire LITIS EA 4108

1 rue de Germont

76031 Rouen, France

E-mail : stefan.darmoni@chu-rouen.fr

Informatique et Santé, 2009 (17) : 113-124
© Springer – Verlag France

Enrichissement des bases de connaissances en biologie par extraction de marqueurs de confiance dans la littérature scientifique

Ines Jilani[1] et Marie-Christine Jaulent[1]

[1]INSERM, UMR S 872, Éq. 20, Les Cordeliers, Paris, F-75006 France ; Université Pierre et Marie Curie-Paris6, UMR S 872, Paris, F-75006 France ; Université Paris Descartes, UMR S 872, Paris, F-75006 France

Abstract

The characterization of biomedical knowledge, taking into account the degree of confidence expressed in texts by authors themselves or given by other external hints carried out by the impact factor of the journal, or the study type for example, is an important issue in the biomedical domain. The authors of scientific texts use grammatical and lexical devices to qualify their assertions, voluntarily or not. We named these markers of qualification "confidence markers". We present here the results of our efforts to collect confidence markers (often associated to epistemic modality) from full texts and abstracts, to classify them on the basis of semantics, and their use within a knowledge extraction system. We propose in this study, an implementation of these confidence markers for functional annotation of the human gene Apolipoprotein (APOE) thought to be involved in Alzheimer's disease. As a result, we obtain, through the extraction system, triplets: (G, F, PMID), in which G is the gene APOE, F is its function found in texts and the PMID of the article from which this knowledge was extracted. Moreover, a multidimensional projection space is proposed for representing the extracted knowledge depending on confidence criteria associated with it.

Keywords

Linguistics; Artificial Intelligence; Confidence Intervals; Semantics: Computational Biology; Medical Informatics Computing

1 Introduction

La maîtrise des techniques de séquençage du génome et les manipulations génétiques ont ouvert de nouvelles perspectives en biologie et conduisent actuellement à une explosion de données et de connaissances. Plusieurs initiatives visent à formaliser et à structurer ces données et connaissances en se focalisant sur différents aspects de la biologie, par exemple : SwissProt[1] (description de protéines humaines), GOA[2] (description de fonctions et localisations de gènes et de protéines dans des organismes modèles), OMIM[3] (association

[1] http://www.expasy.org/sprot/ (accédé le 22 novembre 2008).

[2] http://www.ebi.ac.uk/GOA/ (accédé le 22 novembre 2008).

[3] http://www.ncbi.nlm.nih.gov/omim/ (accédé le 22 novembre 2008).

entre les gènes et leurs produits et les maladies héréditaires humaines), HGMD[4] (mutations des gènes humains), Ingenuity[5] (interactions entre gènes/protéines).

Parmi les sources de données et connaissances qui nourrissent les bases biologiques citées plus haut, la littérature scientifique est la plus importante. Le travail assuré par les annotateurs (professionnels, postdocs, etc), qui extraient et interprètent ces connaissances, est essentiellement manuel et se trouve limité par le volume d'informations qu'il faut analyser. En effet, chaque année, environ 1 million de nouvelles publications sont enregistrées dans Pubmed[6]. Cette base de données bibliographiques propose, en juin 2008, l'accès à plus de 16,5 millions de publications[7].

Malgré ces initiatives, il a été récemment estimé que la majeure partie (80 % en 2005) des données et des connaissances biologiques restent aujourd'hui encore disponibles uniquement dans un format textuel [1]. D'une manière générale, aussi bien pour les biologistes que pour les annotateurs (également biologistes mais dédient leur activité à la tâche d'annotation), la recherche d'information dans la littérature est une procédure longue et fastidieuse. Dans ce contexte, plusieurs méthodes se développent pour permettre l'extraction automatique de connaissances biologiques à partir de la littérature scientifique. Ces méthodes, basées sur des approches d'analyse syntaxique et sémantique de corpus textuels et des techniques de traitement automatique des langues, sont nombreuses [2-4] et évaluent comparativement leurs résultats au cours de compétitions telles que BioCreAtIvE (*Critical Assessment of Information Extraction Systems in Biology*) [5] et TREC [6]. Dans notre laboratoire, nous avons proposé une approche par patrons lexico-syntaxiques pour retrouver de l'information sur les fonctions des gènes dans les textes [7].

Que ce soit pour les annotateurs ou pour les systèmes informatiques, l'accès aux publications et leur analyse sont rendus difficiles par le manque d'homogénéité dans l'écriture et la désignation des entités manipulées. Cela se reflète d'une part dans l'accord moyen entre les annotateurs, qui reste faible, environ 30 % pour l'annotation en protéomique fonctionnelle [8] et d'autre part les mesures de sensibilité, spécificité ou précision encore peu élevées qui sont obtenues par les approches informatiques [5].

La perte du contexte est une limite de l'extraction d'informations sur les gènes, comme leurs fonctions, les maladies qui leur sont associées, leurs interactions. Par exemple, si la connaissance provient d'une procédure expérimentale ou d'une référence à un autre article, alors le contexte permettant à la personne qui reçoit la connaissance extraite de l'associer à une certaine fiabilité et confiance est perdu. Aucune des approches citées précédemment ne prend en compte le contexte ou ne délivre d'informations sur la fiabilité de la connaissance extraite.

Afin de pallier ces limitations, nous nous sommes intéressées à l'identification et à la formalisation de cette information additionnelle avec pour objectif de pondérer la connaissance extraite et d'associer une quantification de la fiabilité aux connaissances explicitées dans les bases existantes. Notre objectif à long terme est de détecter automatiquement les marqueurs de confiance dans les textes qui constituent une trace du contexte et de les intégrer aux patrons lexico-syntaxiques que nous avons déjà développés. L'indice de fiabilité calculé en fonction du contexte permettra d'apporter une homogénéisation (inter-annotateurs et inter-bases) des annotations des gènes et des

[4] http://www.hgmd.cf.ac.uk/ac/index.php (accédé le 22 novembre 2008).

[5] http://www.ingenuity.com/ (accédé le 22 novembre 2008).

[6] http://www.ncbi.nlm.nih.gov/pubmed/

[7] NLM Technical Bulletin - June 26, 2008 MLA 2008: NLM Online Users' Meeting Remarks http://www.nlm.nih.gov/pubs/techbull/mj08/mj08_mla_dg.html (last visited on Nov, 19th 2008).

protéines, de même que des informations pour juger de la fiabilité des informations proposées. La décision finale sur le statut et la validité de ces informations reviendra aux utilisateurs biologistes.

Dans cet article, nous proposons une étude sur l'utilisation de marqueurs de confiance dans la littérature biologique. Après un état de l'art sur les travaux autour des notions de « nuanceur » et « modalité » dans les textes biologiques, nous présentons une étude qui, à partir d'un corpus textuel, nous a permis d'élaborer une classification des marqueurs de confiance dans le domaine. Nous avons testé ce modèle au contexte de la maladie d'Alzheimer et, plus précisément, à la recherche de connaissances concernant le gène *apolipoprotein* E (APOE) qui est soupçonné par les biologistes d'être lié à la maladie d'Alzheimer chez l'Homme. Nous présentons et discutons les résultats de cette étude.

2 État de l'art

Les connaissances concernant les gènes/protéines dans les textes de biologie ne sont pas toujours exprimées de manière catégorique et irréfutable. En effet, les auteurs de ces textes ont assez souvent recours à des formulations linguistiques exprimant le doute ou tout au moins exprimant la distance qu'ils instaurent entre leur énoncé et leur opinion propre.

1. *Our results prove that GRK1 is involved in phototransduction*
2. *Our results suggest that GRK1 might be involved in phototransduction*

Dans les exemples ci-dessus, pour le cas 1, l'extraction *involve(GRK1 → phototransduction)* est justifiée car il n'y a pas de doute sur l'information présentée, elle est même renforcée par l'utilisation d'un marqueur ayant un fort degré de certitude « *prove* ». Pour le cas 2, la même extraction *involve(GRK1 → phototransduction)* ne saurait être justifiée, puisque l'utilisation de marqueurs de faibles degrés de certitude est faite « *suggest* » et « *might* ».

La perte du contexte lors de l'extraction de la connaissance est donc problématique dans ces cas.

Ces dispositifs linguistiques représentent ce que plusieurs auteurs ont déjà évoqué dans la littérature linguistique et nomment « *hedges* », ou « *qualifiers* », ou « *modifiers* », ou « *downtoners* », ou « *boosters* ». Ils peuvent prendre diverses formes grammaticales (verbes lexicaux, adverbes, noms, adjectifs...) et sont souvent assimilés à la modalité linguistique, mais pas toujours.

2.1 La notion de modalité

En linguistique, on s'accorde généralement à définir la modalité comme un domaine sémantique qui englobe un ensemble de valeurs (éventuelle, hypothétique, probable, possible, nécessaire...) exprimant l'opinion, l'attitude ou le jugement de l'énonciateur par rapport au contenu propositionnel de l'énoncé. Le domaine modal est généralement divisé en deux sous-domaines : le déontique, qui englobe l'obligation et la permission et l'épistémique qui s'applique à la possibilité. Coates [9] indique que la modalité épistémique est en rapport avec les suppositions de l'auteur ou avec son estimation des possibilités et, dans la plupart des cas, elle indique la confiance de l'auteur (ou le manque de confiance) dans la véracité de la proposition exprimée. L'opposition entre déontique et épistémique se révèle cependant insuffisante compte tenu de la complexité des phénomènes modaux. Certains auteurs suggèrent des modalités aléthiques, appréciatives, intersubjectives (ordre, conseil, reproche...), bouliques (volonté), implicatives (condition,

conséquence...), temporelles ; ou encore incluent la négation dans cette catégorie. Les différents types de modalités prises en compte le plus souvent, dans les travaux de linguistiques, sont définis précisément dans la référence [10]. La modalité retrouvée le plus fréquemment dans les textes de biologie est de nature épistémique et c'est celle-ci qui va nous intéresser.

Les marqueurs de la modalité épistémique présents dans les textes de biologie sont divers. Ils peuvent être de nature lexicale (verbes modaux, adverbes modaux, proposition en incise) ou morphologique (modes verbaux).

2.2 La notion de « nuanceur »

Le premier à avoir évoqué la notion de nuance dans les textes est Lakoff [11], il nomme ces marqueurs « *hedges* » (« esquives » en français) et les décrit comme étant « des mots dont le rôle est de rendre les choses plus ou moins floues ». Cette définition a été reprise par la suite pour faire référence à tous les dispositifs linguistiques utilisés par les auteurs dans les textes pour nuancer leurs expressions.

Hyland [12] a recherché ces marqueurs dans 26 articles scientifiques de biologie cellulaire et moléculaire (75 000 mots). Hyland indique que ces marqueurs signalent une anticipation de la part de l'auteur d'une possible opposition à sa déclaration et que la fonction de ces marqueurs est de trois types : 1) « *X may cause Y* » : spécifie l'état actuel des connaissances, 2) « *These data indicate that....* », « *the model implies...* », « *it was assumed that...* » spécifient le désir de l'auteur d'anticiper des conséquences négatives possibles que serait l'invalidation de ses résultats, 3) « *We propose that...* », « *I believe that...* » contribuent au développement de la relation auteur-lecteur. Dans cette étude, un marqueur (« *hedge* ») est relevé tous les cinquante mots.

Presque une décennie plus tard, Light [13] mène une étude sur l'utilisation du langage spéculatif dans les abstracts de Medline. Il focalise sa recherche sur l'expression des niveaux de croyance et prend en compte l'importance du contexte de l'abstract. Pour cette étude, les auteurs considèrent les annotations au niveau de la phrase; ils définissent une phrase « spéculative » comme étant une phrase qui contient au moins un fragment « spéculatif ». Selon eux, il y aurait 11 % des phrases des abstracts de Medline qui contiennent des fragments « spéculatifs ». Enfin, cette étude ne présente qu'une catégorisation (faite manuellement par quatre personnes) des abstracts, selon que la phrase est : « très spéculative » (« *high speculative* »), « peu spéculative » (« *low speculative* ») ou « formelle » (« *definite* »).

Mercer [14] présente une méthodologie pour le développement d'un outil d'indexation des citations dans les articles scientifiques prenant en compte le contexte de la citation. L'étude menée a montré que les « nuanceurs » (appelés « *hedging cues* » par les auteurs) sont plus utilisés dans des contextes de citation qu'ailleurs dans les textes. Plus tard, autour de la même idée, le travail de Mizuta [15], basé sur l'article de Teufeul et Moens [16], présente un système automatique pour résumer des articles du domaine informatique. L'étude se focalise sur des textes du domaine de la biologie. Vingt articles sélectionnés au hasard sont annotés, à l'aide d'un arbre de décision. Les annotations sont, par exemple, *Outline* (exposition des grandes lignes de l'article), *Method* (méthode), *Insight* (aperçu, idée), *IMP* (implication), etc. Pour détecter les différents segments dans les textes, des indices linguistiques sont utilisés. Par exemple, si le fragment « *Our results support the hypothesis that...* » est détecté dans la phrase, la zone sera identifiée comme *Insight*. Si un contexte modal est détecté (introduit par « *might* », « *may*», « *could* », etc. ou des adverbes « *probably*», « *perhaps* », etc.), alors la zone sera reconnue comme *IMP*.

Rizomilioti [17] adopte une vue étendue de la modalité, en s'intéressant à ce qu'il appelle les « *downtoners* » (dispositif lexical qui signale le manque de confiance du locuteur comme « *might* », « *maybe* ») et à leurs opposés, les « *boosters* » (tels que « *obviously* », « *clear* »). Il s'intéresse également à ce qu'il nomme « les indicateurs de certitude » tels que « *show* » et « *demonstrate* » (verbes lexicaux), « *finding* » et « *conclusion* » (noms).

En résumé, la notion de nuanceur étend celle de la modalité en incluant les « approximateurs » tels que « *somewhat* », « *a number of* », les clauses en « *if* », l'utilisation de la voix passive, la position personnel du locuteur « *according to* », « *in my opinion* », « *to our mind* », et des conditions expérimentales limitées proposées par Hyland [12], telles que « *we do not know whether...* », « *viewed in this way the concept...becomes obsolete* ».

Comme le suggère cet état de l'art, les notions d'« hedge» et de « modalité épistémique » ont été beaucoup étudiées dans la communauté linguistique. Toutes les études menées et citées plus haut sont de nature qualitative. Les auteurs n'ont pas modélisé les marqueurs, ni intégré leur utilisation pour la pondération d'informations dans des systèmes d'extraction de connaissances.

2.3 Synthèse

La définition de la modalité épistémique et par élargissement des nuanceurs, que nous prenons en considération dans notre travail, est celle de Coates [9]. C'est-à-dire que sont pris en compte les cas où l'assertion est peu certaine mais aussi ceux où elle est très sûre. Nous y ajoutons cependant les cas où l'assertion est affirmative, sans marqueurs particuliers, et représente une proposition véhiculant un haut degré de confiance.

À côté des nuanceurs et de la modalité, nous distinguons également d'autres types d'éléments pouvant donner des indices sur la confiance que l'on peut accorder à une connaissance extraite d'un texte. Ces éléments concernent un niveau global de confiance accordé à la connaissance extraite et qui s'appuient sur une méta connaissance extérieure au texte (comme par exemple, la notoriété des auteurs).

C'est l'ensemble de ces marques que nous regroupons sous le nom de « marqueurs de confiance ». Ils représentent tout type d'élément intrinsèque ou extrinsèque à l'article, pouvant véhiculer implicitement ou explicitement une information servant à documenter la fiabilité de la connaissance extraite. On distingue donc :

- les nuanceurs (« *hedges* ») ;
- la modalité linguistique et grammaticale ;
- le facteur d'impact (« *impact factor* ») du journal dont provient l'article ;
- le type d'étude menée dans l'article (étude expérimentale, théorique, essai clinique, rapport...) ;
- l'affiliation des auteurs de l'article ;
- la date de l'article ;
- ...

Dans l'étude menée, nous ne considérons que les trois premiers types de marqueurs.

3 Matériel

Nous utilisons trois corpus constitués à partir de Pubmed[8] que nous avions préalablement constitués lors de notre travail sur les patrons lexico-syntaxiques [18].

[8] http://www.ncbi.nlm.nih.gov/sites/entrez

- CORP1 est obtenu avec une liste de 160 gènes humains candidats, suspectés d'être liés à la maladie d'Alzheimer [19]. Ce corpus est une collection de 355 résumés, contenant 817 phrases, 213 618 mots et a une taille de 1 Mo.
- CORP2 est composé de 68 textes entiers en relation avec les résumés de CORP1, disponibles dans Pubmed Central[9]. Il contient 27 912 phrases, 1 123 873 mots et a une taille de 2,4 Mo.
- CORP3 est une collection de 348 résumés collectés de Pubmed grâce à une requête contenant une liste de 160 gènes du nématode identifiés par les biologistes comme étant liés à la maladie d'Alzheimer. CORP3 contient 825 phrases, 201 753 mots et a une taille de 1 Mo.

WordNet® [20] est une importante base de données lexicales de l'anglais : des noms, des verbes, des adjectifs et des adverbes sont regroupés par synonymie, chacun exprimant un concept distinct. Cette base de données a été utilisée dans notre étude pour enrichir les marqueurs de confiance extraits manuellement, en permettant l'identification de leurs synonymes.

4 Méthode

Notre méthode consiste à collecter manuellement le plus grand nombre possible de marqueurs de confiance (à partir de CORP1, CORP2 et CORP3), en les enrichissant de leurs synonymes grâce à WordNet®. Nous proposons une classification de ces marqueurs et nous définissons d'autres critères de confiance susceptibles de modifier le degré de confiance des connaissances extraites (par exemple, l'impact facteur de la revue). Ainsi, nous avons pu établir un modèle pour les marqueurs de confiance. Nous avons appliqué ce modèle à l'extraction de connaissances à un gène particulier susceptible d'être lié à la maladie d'Alzheimer chez l'homme.

4.1 Collecte des marqueurs de confiance

En nous appuyant sur l'état de l'art, nous avons collecté les marqueurs de confiance manuellement, retenant chaque dispositif linguistique suggérant que l'auteur était en train de nuancer son assertion, la présence de l'auteur dans le texte, les indices de la présence d'un contexte particulier de l'information présentée dans le texte, etc.

4.2 Enrichissement de la liste des marqueurs de confiance

Les marqueurs collectés sont les traces des dispositifs linguistiques utilisés par les auteurs dans les documents analysés. Ils ne fournissent qu'une vue partielle de la diversité d'expression des marqueurs. Il est donc nécessaire d'étendre ces dispositifs à leurs variantes morphosyntaxiques et à leurs synonymes. Nous avons utilisé les ressources de WordNet® pour une extension manuelle des marqueurs à leurs synonymes.

4.3 Classification des marqueurs de confiance et définition d'autres marqueurs de confiance

Les marqueurs doivent être classés en classes homogènes pour faciliter leur utilisation au sein d'un outil d'extraction automatique de connaissances et dans le but de pondérer ces

[9] http://www.pubmedcentral.nih.gov/

mêmes connaissances. Ces classes sont ensuite graduées en fonction de la confiance ou de la défiance que le marqueur ajoute à la connaissance à laquelle il se rapporte. Nous proposons deux types de classification qui sont présentés dans la partie résultat.

Dans le cadre de cette étude, et à titre expérimental, nous avons fait le choix de prendre en compte un critère additionnel pour pondérer l'information extraite : l'impact facteur (IF) du journal. L'IF d'un journal est calculé sur une période de trois ans, et peut être considéré comme étant le nombre moyen de fois que les papiers publiés dans ce journal sont cités, deux ans après leur publication. Le choix de ce facteur est discutable comme représentation de la confiance associée à une connaissance mais c'est une information facile à obtenir, qui est utilisée ici à titre expérimental. Il est de plus très apprécié par les biologistes, car il donne une « valeur » supplémentaire aux connaissances présentes dans les articles publiés dans des journaux à haut facteur d'impact.

4.4 Modèle pour les marqueurs de confiance

La dernière étape de la méthode implique d'établir un modèle pour l'ensemble final des marqueurs. Ce modèle de marqueurs de confiance doit être utilisé au sein d'un outil d'extraction automatique de connaissances fonctionnelles sur les gènes, à partir de textes scientifiques, ainsi que les informations concernant la confiance qui les accompagnent. Nous avons décrit dans un article précédent la méthode des patrons lexico-syntaxiques, basée sur les expressions régulières et les graphes, qui permet d'extraire des connaissances à partir des textes [7]. Nous avons utilisé la même approche pour définir le modèle des marqueurs de confiance. Il a été construit manuellement.

Enfin, nous avons appliqué ce modèle au contexte de la maladie d'Alzheimer et, plus précisément, à la recherche de connaissances concernant le gène apolipoprotein E (*APOE*) qui est soupçonné par les biologistes d'être lié à la maladie d'Alzheimer chez l'Homme. Nous avons construit un corpus spécifique (CORP_APOE) pour l'application de ce modèle. Ce corpus a été obtenu en questionnant Pubmed avec le gène *APOE* et ses synonymes. Nous considérons toutes les fonctions biologiques d'*APOE* que nous extrayons avec notre système d'extraction de connaissances pour obtenir des triplets (*G, F, PMID*), où *G* est le gène *APOE*, *F* est sa fonction et *PMID* est l'identifiant Pubmed de l'article à partir duquel la connaissance est obtenue. Ensuite, nous détectons tous les marqueurs de confiance qui pourraient nuancer la connaissance extraite, nous pondérons cette connaissance en fonction des classes auxquelles les marqueurs appartiennent. Finalement, nous représentons la confiance associée à chaque connaissance extraite dans un espace multidimensionnel selon les critères de confiance exposés précédemment.

5 Résultats

Dans cette section, nous détaillons les résultats obtenus pour chaque étape de la méthode proposée.

5.1 Collecte des marqueurs de confiance

Une liste de 250 marqueurs de confiance collectés manuellement a été établie. Ces marqueurs peuvent être des auxiliaires modaux (*must*, *will not*, etc.), des adverbes modaux (*possibly*), de manière (*clearly*), de temps (*often*), de degré (*largely*), de lieu (*here*), des informations de contexte (*in our case*), des informations sur la source (*on the basis of*), certains verbes (*assume*, *confirm*, *appear*, etc.), etc.

5.2 Enrichissement de la liste des marqueurs de confiance

Une fois collectés, nous avons étendu manuellement les marqueurs à leurs variantes morphologiques. Cette recherche s'effectue à partir des lemmes des marqueurs de confiance. Par exemple, nous avons étendu le marqueur de confiance : "*previous study*" à "*previous studies*" (ajoutant ainsi les formes plurielles ou singulières), et "*remain unknown*" à "*remains unknown*" (forme flexionnelle du verbe). Cet enrichissement nous a permis de construire pour les 3 corpus étudiés, une liste de 478 marqueurs.

Nous avons également mené un deuxième type d'extension pour lequel les synonymes des marqueurs de confiance extraits ont été pris en compte. Par exemple, nous avons étendu le marqueur de confiance "*we anticipate*" à "*we expect*" (*expect* étant le synonyme de *anticipate* dans WordNet®), "*previous study*" à "*previous work*" et "*previous report*" (*work* et *report* sont des synonymes de *study*). Il y a un débat concernant l'utilisation, l'exhaustivité, la précision et la pertinence linguistique de WordNet® mais cette ressource est considérée comme appropriée pour le contexte dans lequel nous travaillons. En effet, le processus d'extension basé sur les synonymes utilisés pour ce travail ne modifie en rien les résultats finaux, il augmente juste le nombre de marqueurs de confiance à 700. Notons qu'ignorer cette étape d'acquisition des variantes et des synonymes des marqueurs de confiance nous aurait fait perdre des informations pour la pondération des résultats, puisque la liste des marqueurs dont nous disposions n'aurait pas été assez exhaustive.

5.3 Classification des marqueurs de confiance et définition des autres critères de confiance

La nécessité de classer les marqueurs de confiance en fonction de leur sémantique a émergé afin de donner différents poids à la façon dont ils caractérisent la connaissance. Nous avons retenu quatre catégories différentes, décrites ci-dessous dans un ordre de confiance ascendante :

1. **Interrogation ou tâtonnements de l'auteur.** Cela correspond à une connaissance qui est encore non prouvée et qui requiert une démonstration. Cette connaissance peut aussi correspondre à une interprétation de l'auteur en l'absence de preuves concluantes dans l'article.
   ex. : *"remain to be confirmed", "has yet to be identified", "?"*
2. **Distance suggérée par l'auteur par rapport à son assertion ou à une connaissance présentée dans le texte.** Cette subtile distinction a été suggérée par Hyland [12]. Cela peut aussi correspondre à une restriction de la connaissance à un certain contexte (ex.: le contexte de l'article ou d'une expérience en particulier).
   ex. : *"our findings suggest that", "in this case we conclude that", "it is possible that"*
3. **Études d'autres chercheurs, références à d'autres travaux, articles ou méthodes.** Ces connaissances sont associées à un haut niveau de confiance, car nous partons du principe que si un article est cité, les informations qui y sont présentes sont considérées comme vraies par l'auteur, ou au pire supposées vraies. Ces connaissances n'ont donc pas besoin de plus amples démonstrations dans l'article dans lequel elles sont rapportées.
   ex.: *"previous observation", "it is now believed that", "it has been proposed that"*
4. **Démonstration ou preuve faite par l'auteur.** Cela correspond aux travaux de l'auteur présentés dans son article. Cela implique l'utilisation de marqueurs associés à des connaissances auxquelles on accorde un très haut niveau de confiance car elles

sont basées sur des démonstrations présentées dans l'article considéré.
ex. : *"we reveal that", "we show here that", "our results indicate that"...*

Ces quatre catégories regroupent les marqueurs de confiance de *Type 1.* Sur les 478 marqueurs détectés initialement et étendus à leurs variantes morphologiques, 56 se trouvent dans la classe 4, 65 dans la classe 3, 214 dans la classe 2 et 134 dans la classe 1. Les 9 marqueurs restants concernent la négation (*no, not, unable,* etc.) et ne sont pas pris en compte en tant que classe de *Type 1.*

Nous avons également distingué des marqueurs qui modifient le niveau de confiance à l'intérieur des quatre classes présentées précédemment. Ces marqueurs caractérisent la subjectivité dans les textes, décrivant des choses ou des évènements qui sont possibles mais pas observés ou pas certains. Ces marqueurs sont de *Type 2.* Ils sont représentés dans la figure 1, allant de la négation à l'affirmation, c'est-à-dire du moins probable au plus probable.

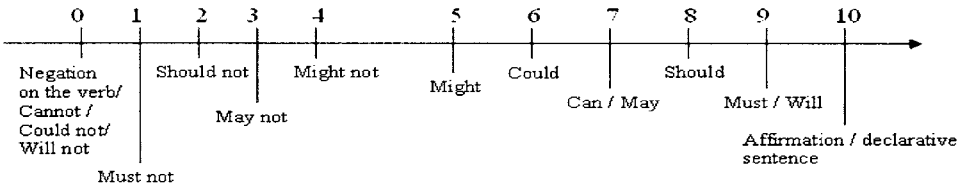


Figure 1 - Graduation des marqueurs de confiance de Type 2.

La négation (poids 0) indique que l'auteur présente une assertion négative, i.e. l'auteur considère que la connaissance présentée est fausse. Par exemple, si l'auteur présente l'information suivante : « le gène A n'est pas lié à la fonction B » alors nous aurons la connaissance sous forme de couple (gène A, fonction B) avec le poids 0, car la connaissance est sous la forme négative. Ce genre de connaissance a donc le poids de confiance le plus faible (*"unable", "not", "neither...nor"...). Must not, should not, may not* et *might not* traduisent plus de confiance qu'une négation claire. *Might, could, can/may* et *should* expriment des probabilités croissantes. Enfin, *must/will* et les phrases affirmatives expriment le plus haut degré de confiance.

À travers l'utilisation du critère de l'impact facteur (IF), nous faisons l'hypothèse que l'IF du journal est en relation avec la fiabilité des informations biologiques et médicales qu'il publie. Cela n'implique pas que les journaux à faible IF publient des expériences biologiques erronées, mais simplement que les biologistes ont l'habitude de plus se fier aux informations provenant de journaux à fort IF. Le journal de publication d'un article peut être retrouvé assez aisément grâce à son numéro ISSN (*international standard serial number*), disponible au format XML dans le résultat de recherche Pubmed, dans la balise *<ISSN IssnType="Print">**0003-2697**</ISSN>* (le numéro en gras correspond à l'ISSN du journal de parution). La liste des IF des journaux peut être utilisée ensuite pour pondérer les connaissances issues de ces journaux.

5.4 Modèle de critère de confiance pour l'extraction automatique de connaissances

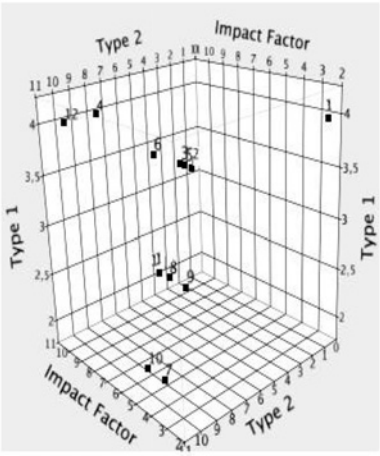
Dans un travail précédent, nous avons utilisé l'approche des patrons lexico-syntaxiques (expressions régulières) pour l'extraction d'informations à partir de textes en biologie. Nous réutilisons ce modèle pour les marqueurs de confiance. Tout d'abord, les lemmes de

chaque marqueur de confiance sont définis, afin de détecter toutes les formes possibles du mot. Ensuite, les synonymes des marqueurs ont été collectés à l'étape d'enrichissement de la liste des marqueurs de confiance. Par exemple, des marqueurs de confiance comme "*we anticipate*" et "*we expect*" forment l'expression *we<have>*(<anticipate>+<expect>)*, où les mots entre < > correspondent à la forme canonique du mot, + correspond au *OU* logique et * correspond à *0* ou *n* occurrences. Quand certains marqueurs sont identifiés comme étant synonymes dans WordNet®, nous les incluons dans la même expression. C'était le cas pour "*we hypothesise*" et "*we suspect*", et, avec l'extension avec leurs synonymes, cela a créé l'expression régulière suivante :

we<have>(<hypothesise>+ <speculate>+ <expect>+<predict>+<suspect>).*

Des noms, adjectifs, adverbes, ainsi que des verbes, peuvent être étendus avec la ressource WordNet®, comme dans l'expression suivante :

<have><be>(previously+now)*(largely+widely+extensively+generally)*<confirm>* qui récupère les phrases *"have been previously confirmed"* ainsi que *"is now largely confirmed"* ou *"is widely confirmed"*. À travers ces exemples, nous pouvons observer la capacité des expressions régulières à identifier des évènements linguistiques ciblés.



1 (ApoE epsilon4 allele, vascular depression, 17337010)
2 (APOE, arteriosclerosis, 17132969)
3 (Apolipoprotein E, aggressive behavior, 17077632)
4 (ApoE, differentially regulated, 17077142)
5 (APOE4 allele, late onset AD, 17069783)
6 (ApoE, cognitive impairment, 16764677)
7 (ApoE, pupil response, 16764677)
8 (apolipoprotein E (APOE) epsilon4 allele, sporadic AD, 16741934)
9 (Apolipoprotein E epsilon 4, increased vulnerability to cell death, 16736246)
10 (ApoE4, HSV-1 latency, 16699018)
11 (apolipoprotein E epsilon4 allele, late onset AD, 16681804)
12 (ApoE, repairing injured neurons, 16585500)

Figure 2 - Visualisation graphique des triplets extraits selon leurs degrés de confiance.

5.5 Application au contexte du gène apolipoprotein E

Nous avons utilisé le système développé dans le cadre d'un travail antérieur [7] pour extraire 650 éléments de connaissances à partir du corpus CORP_APOE. Ces éléments sont extraits sous la forme de triplets *(G, F, PMID)*, où *G* est le gène, *F* sa fonction et *PMID* l'identifiant Pubmed de l'article d'où l'annotation fonctionnelle provient. Nous appliquons les expressions régulières construites pour les marqueurs aux phrases qui sont à l'origine des triplets ce qui nous a conduit à l'extraction de 12 éléments de connaissances associés à des marqueurs de confiance. Pour chacun des 12 triplets, le type 1, le type 2 et l'IF sont explicités. La connaissance biologique extraite est placée dans un espace 3D selon son niveau de confiance de *Type 1*, son niveau de confiance de *Type 2* et de l'IF du journal source (figure 2).

6 Conclusion

Le travail décrit dans cet article consiste à intégrer l'identification de marqueurs de confiance dans un système automatique d'extraction de connaissances à partir de la littérature en biologie [7]. Les marqueurs ont été détectés manuellement à partir de corpus d'abstracts et de textes entiers tirés de Pubmed. Ces marqueurs ont été étendus grâce à WordNet® et classés en quatre classes de type 1, et 10 classes de type 2 selon leur sémantique et la force avec laquelle ils modifient la confiance que l'on porte à l'information.

L'extraction et la caractérisation de connaissances concernant la fonction biologique de gènes et leurs produits, basés sur un système de ce type, améliorent notre compréhension de la fiabilité de ces informations, particulièrement si la connaissance extraite est projetée dans un espace multidimensionnel. Nous travaillons actuellement avec des experts biologistes pour évaluer l'efficacité d'une telle présentation.

Nous planifions de travailler sur la détection automatique de nouveaux marqueurs de confiance. En effet, la détection manuelle des marqueurs de confiance, dans le cadre de cette étude, a été réalisée sur 3 corpus de tailles relativement modestes, une détection automatique sur un corpus plus large permettrait de compléter de façon plus exhaustive notre liste actuelle de marqueurs.

Notons que dans l'étude qui est présentée, nous ne tenons pas compte du style de l'auteur ni de son milieu culturel ou institutionnel. Nous avons également remarqué que ces marqueurs pouvaient être parfois regroupés au sein d'une même phrase avec une sémantique particulière. De plus, une autre dimension possible serait la prise en compte du type d'article étudié (article de revue, essai clinique, etc.). À long terme, l'objectif de recherche de notre projet consiste à proposer des outils automatiques qui faciliteront l'accès aux informations disponibles dans la littérature scientifique. Ces informations seront pondérées par un indice de leur fiabilité, calculé en fonction des contextes et des données de référence.

Les marqueurs de confiance présentés dans ce travail peuvent également être retrouvés dans des articles du domaine médical. La biologie ne consiste finalement qu'à un champ applicatif des travaux réalisés et évoqués dans cet article sur l'extraction de connaissances à partir de textes et leurs valeurs de confiance qui y sont associées.

Références

[1] Koehler J. Editorial. *Briefings in Bioinformatics* 2005; 6(3): 220-221.

[2] Rice SB, Nenadic G, Stapley BJ. Mining protein function from text using term-based support vector machines. *BMC Bioinformatics* 2005 ; 6 (Suppl 1) : S22.

[3] Krallinger M, Pradon M, Valencia A. A sentence sliding window approach to extract protein annotations from biomedical articles. *BMC Bioinformatics* 2005; 6 (Suppl 1): S19.

[4] Ono T, Hishigaki H, Tanigami A, Takagi T. Automated extraction of information on protein-protein interactions from the biological literature. *Bioinformatics* 2001; 17(2): 155-161.

[5] Blaschke C, Leon EA, Krallinger M, Valencia A. Evaluation of BioCreAtIvE assessment of task 2. *BMC Bioinformatics* 2005; 6(Suppl 1): S16.

[6] Hersh W. Report on TREC 2003 genomics track first-year results and future plans. *ACM SIGIR Forum* 2004; 38(1): 69-72

[7] Jilani I, Grabar N, Jaulent MC. Fitting the finite-state automata platform for mining gene functions from biological scientific literature. In *SMBM* (*Symposium on Semantic Mining in Biomedicine*), Germany, 2006.

[8] Camon EB, Barrell DG, Dimmer EC, Lee V, Magrane M, Maslen J, Binns D, Apweiler R. An evaluation of GO annotation retrieval for BioCreAtIvE and GOA. *BMC Bioinformatics* 2005; 6(Suppl 1): S17.

[9] Coates J. *The Semantics of the Modal Auxiliaries*. Routledge 1983, 259 p.

[10] Le Querler N. *Typologie des Modalités*. Presses Universitaires de Caen 1996.

[11] Lakoff G. Hedges: A study in meaning criteria and the logic of fuzzy concepts, *Journal of Philosophical Logic* 1972; 2, 458-508.

[12] Hyland K. Talking to the Academy: Forms of Hedging in Science Research Articles, *Written Communication* 1996; 13, 251-281.

[13] Light M, *et al.* The Language of Bioscience: Facts, Speculations, and Statements In Between. *HLT-NAACL 2004 Workshop: BioLINK 2004, Linking Biological Literature, Ontologies and Databases*. Association for Computational Linguistics, Boston, Massachusetts, USA, 2004; 17-24.

[14] Mercer RE, Di Marco C. A Design Methodology for a Biomedical Literature Indexing Tool Using the Rhetoric of Science. *HLT-NAACL 2004 Workshop: BioLINK, Linking Biological Literature, Ontologies and Databases*. Association for Computational Linguistics, Boston, Massachusetts, USA, 2004; 77-84.

[15] Mizuta Y, *et al.* Zone analysis in biology articles as a basis for information extraction, *Int J Med Inform* 2006; 75, 468-487.

[16] Teufel S, and Moens M. Summarizing scientific articles: experiments with relevance and rhetorical status, *Computational Linguistics* 2002; 28, 409-445.

[17] Rizomilioti V. Exploring Epistemic Modality in Academic Discourse Using Corpora. In US, S. (ed), *Information Technology in Languages for Specific Purposes*. 2006; 53-71.

[18] Grabar N, Jaulent MC, Chambaz A, Lefebvre C, Neri C. Sifting abstracts from Medline and evaluating their relevance to molecular biology. *Stud Health Technol Inform* 2006; 124: 111-6.

[19] Lefebvre C, Aude JC, Glemet E, Neri C. Balancing protein similarity and gene co-expression reveals new links between genetic conservation and developmental diversity in invertebrates. *Bioinformatics* 2005; 21(8): 1550-8

[20] WordNet, An Electronic Lexical Database, C. Fellbaum ed., The MIT Press, Cambridge, Mass, 1998.

Adresse de correspondance
Ines JILANI

SPIM, UMR_S 872, eq.20 - Centre des Cordeliers
15 rue de l'École de Médecine

75006 Paris
E-mail : ines.jilani@spim.jussieu.fr

Informatique et Santé, 2009 (17) : 125-136
© Springer – Verlag France

Interopérabilité sémantique : comparaison de la représentation du score d'APGAR en HL7V3 et OpenEHR

Marc Cuggia[1], Sahar Bayat[1], Patrice Poulain[2], Patrick Pladys[3], Hélène Robert[4], Régis Duvauferrier[1]

[1]*Faculté de Médecine de Rennes, INSERM U986, IFR 140, Université Rennes 1 France*
[2]*Département de Gynécologie-Obstétrique et Médecine de la Reproduction, CHU de Rennes*
[3]*Réseau Périnatal Bien Naître en Ille-et-Vilaine 2, Rue de l'Hôtel-Dieu, 35000 Rennes, France*
[4]*Pôle médico-chirurgical de Pédiatrie et de génétique clinique, Université de Rennes 1*

Abstract
Semantic interoperability, a prerequisite to eHealth projects, relies on sharing both information and knowledge models between information systems. Two of the standards of information models are HL7 v3 and the European norm, EN13606/OpenEHR. The paper compares both standards on a fragment of the prenatal medical record, the APGAR score. The HL7v3 perinatology DMIM specification and the OpenEHR APGAR archetype were used. HL7v3 appears to be more formal than OpenEHR and able to represent in an easier way the clinical context. For both standards, the binding to reference terminologies such as LOINC is poor.

Keywords

Interoperability, Information model, Standards, Terminology

1 Introduction

Depuis moins d'une décennie, grâce au développement des technologies de l'information et de la communication, un certain nombre d'activités médicales basculent dans le nouveau paradigme de l'e-santé. Eysenbach définit l'e-santé comme l'intersection de la médecine, de l'informatique médicale, et de la santé publique. L'e-santé correspondant à « un ensemble de services de santé, basé sur les technologies de l'information et de la communication, capable de fournir et valoriser l'information médicale ».

La connexion de l'ensemble de ces services aussi divers soient-ils (dossiers médicaux partagés, messageries sécurisées, prescriptions connectées, etc.) supposent que les systèmes d'information qui les supportent soient interopérables.

Une voie de résolution de l'interopérabilité consiste en ce que les systèmes partagent des modèles d'information capables de capturer les données, et des modèles de connaissance (ou terminologiques comme SNOMED, LOINC) permettant de coder de façon non ambiguë les données [1].

Il existe actuellement plusieurs normes (*de jure* et *de facto*) concurrentes de modèle d'information susceptibles d'être utilisées par les industriels. En particulier pour le dossier médical électronique, on trouve d'une part le Clinical Document Architecture (CDA) défini dans la version 3 d'HL7 [2], et la norme européenne EN 13606 [3]. Cette dernière est fortement inspirée du standard OpenEHR [4]. La question du choix du modèle se pose.

Quels en sont les apports et les inconvénients respectifs ?

Nous proposons ici de comparer la représentation d'un fragment du dossier médical périnatal, en l'occurrence le score d'APGAR dans les formalismes HL7V3 et EN13606/OpenEHR et en particulier la représentation de l'information et le lien établi avec les terminologies de références. Cette étude s'inscrit dans le cadre d'un projet de faisabilité d'un dossier médical périnatal partagé breton.

2 État de l'art

2.1 Le score d'APGAR

Le score inventé en 1952 par Virginia Apgar [5] est utilisé pour évaluer la santé d'un nouveau-né à la naissance. Il est compris entre 0 (arrêt cardiaque) et 10 (normal). Le calcul du score est basé sur l'observation de paramètres cliniques pour chacun desquels on donne une note de 0 à 2. Le score global est la somme de ces notes. Le tableau 1 présente les différents critères et leur cotation.

Tableau 1 - Détails du score d'Apgar.

Paramètres	0	1	2
Battements cardiaques	Absents	< 100 / minute	> 100 / minute
Mouvements respiratoires	Absents	Lents, irréguliers	Réguliers, vigoureux, avec cri
Tonus musculaire	Nul, flasque	Faible : légère flexion des extrémités	Fort : quadri-flexion, mouvements actifs
Réactivité à la stimulation	Nulle	Faible : léger mouvement, grimace	Vive : cri, toux
Coloration	Globalement bleue ou pâle	Corps rose, extrémités bleues	Totalement rose

Le score d'Apgar est évalué classiquement à différents moments (juste après l'extraction, à 1 minute, 5 minutes, voire 10 minutes après la naissance). Ce score va conditionner la prise en charge immédiate du nouveau-né, en particulier par des manœuvres de réanimation. La modélisation du score d'APGAR est disponible dans les normes HL7 V3 et OpenEHR/EN13606.

2.2 Les standards de modèle d'information

2.2.1 HL7 version 3

À partir de 1996, HL7 a développé une nouvelle génération de standards, regroupés sous la dénomination « version 3 » dont la caractéristique est qu'ils reposent et dérivent totalement d'un modèle de classe formel appelé RIM (*Reference Model Information*). Le RIM est une représentation graphique des données cliniques. Il s'agit d'un modèle à partir duquel est déterminé de manière explicite les messages échangés au sein ou entre les systèmes d'information. La méthodologie de dérivation à partir du RIM s'appuie sur un formalisme proche d'UML [6]. La spécification des classes à partir du RIM identifie d'abord les éléments génériques du domaine à modéliser puis, par itérations, va aux plus spécifiques.

Les liens entre les classes du RIM et la sélection et/ou la complétion des classes correspondantes dépendent des conditions techniques, légales et organisationnelles du domaine dans lesquelles on se trouve. Pour ces raisons, la définition des messages s'appuie sur un clonage des classes du RIM dans un sous-modèle (appelé *Domain Message Information Model* ou (DMIM)) spécifique au domaine.

À partir du DMIM, le processus de spécification des messages se poursuit par l'extraction dans des *Refined Message Information Models* (RMIM) des sous-ensembles des classes nécessaires à la spécification des messages. Enfin, la définition opérationnelle des messages, c'est-à-dire utilisables et implémentables par les industriels, s'effectue au sein du *Hiearchical Message Description* (HMD).

Il convient de préciser que l'appellation D-MIM et R-MIM viennent récemment de changer au profit respectivement de DIM (*Domain Information Model*) et CIM (*Constrained Information Model*).

Comme on le voit, il s'agit d'un processus complexe. Dans le cas où certaines classes sont communes à l'ensemble des domaines, HL7/V3 prévoit de formaliser des messages trans-domaines appelés *Common Message Element Type* (CMET). Cette approche permet à la fois de respecter les spécificités d'un domaine, tout en réutilisant ce qui peut l'être.

2.2.2 OpenEHR

L'originalité de GEHR/OpenEHR est d'avoir introduit la notion d'archétype [4, 7]. Cette approche dissocie le modèle d'information en deux niveaux : le modèle de référence et les archétypes proprement dits.

Le modèle (de classes) de référence

C'est l'équivalent du RIM d'HL7. Mais bien qu'il soit spécifique au domaine médical, il reste relativement général. Ce modèle contient un nombre restreint de classes (par exemple : `COMPOSITION`, `SECTION`, `DATA_STRUCTURE`, `ENTRY`, etc. Les 2 premières classes citées permettent de représenter la structure ou le découpage d'un dossier médical. Par exemple, un compte-rendu opératoire sera typiquement dérivé de la classe `COMPOSITION`. La partie décrivant l'histoire de la maladie d'une observation sera de type `SECTION`. La classe `DATA_STRUCTURE` décrit le type de structuration syntaxique des données (par exemple un élément de données isolé, une liste ou un tableau).

Le découpage de l'observation médicale (antécédent, histoire de la maladie, examen clinique, etc.) reposera sur l'utilisation de la classe `SECTION`. La classe `ENTRY` est particulière puisqu'elle permet de représenter les informations médicales proprement dites. `ENTRY` possède 4 sous-classes (`OBSERVATION`, `EVALUATION`, `INSTRUCTION` et `ACTION`).

La classe `OBSERVATION` est utilisée pour représenter une observation, l'état clinique ou paraclinique du patient. `OBSERVATION` sera utilisée pour représenter par exemple l'histoire de la maladie, la tension artérielle du patient ou sa numération formule sanguine.

La classe `EVALUATION` est à prendre au sens d'une hypothèse formulée par le clinicien. Elle permet de représenter les hypothèses diagnostiques ou les risques potentiels encourus par le patient devant des facteurs de risques. `INSTRUCTION` et `ACTION` permettent respectivement de représenter les prescriptions au sens large et les actions menées en réponse aux prescriptions. Ces classes permettent de représenter le détail des étapes d'un workflow.

Les archétypes

Dans le deuxième niveau, les concepts médicaux et de soins tels que la pression artérielle, les résultats de laboratoire, un score d'APGAR, sont modélisés en tant qu'archétypes. Un archétype est donc une modélisation de concepts du domaine, exprimés par l'utilisation de contraintes sur les structures sous-jacentes d'un modèle de référence. L'exemple d'une contrainte est de restreindre la classe générique du modèle de référence « Observation » à celle par exemple de l'archétype « Pression Artérielle ». L'archétype possède 3 parties distinctes :

- des données descriptives ;
- les règles de contraintes ;
- et une section dite « définitions ontologiques ».

Les données descriptives contiennent l'identifiant unique de l'archétype, un code machine décrivant le concept (généralement clinique) et enfin diverses métadonnées décrivant l'auteur, la version, le sujet ou l'objectif de l'archétype. Enfin est spécifié si un archétype est une spécialisation d'un autre archétype. Les règles de contraintes sont au cœur de l'archétype et définissent les restrictions sur les classes du modèle de référence. Ces restrictions seront par exemple la cardinalité, le type syntaxique, des valeurs min-max, etc.

Dans les archétypes, l'expression des règles de contraintes s'effectue grâce à un langage dédié de type objet : l'Archétype Définition Langage (ADL) [8], dérivé du langage Object EIFFEL [9].

Enfin, la troisième partie, appelée selon nous abusivement « Ontology », contient les définitions textuelles des éléments de données utilisés dans l'archétype, ainsi que le vocabulaire utilisé dans les éléments de données traduits dans différentes langues. Ce vocabulaire d'interface est aligné (on parle de *biding*) avec un ou plusieurs vocabulaires de référence (tel que Snomed ou LOINC).

3 Résultats

3.1 Analyse de la représentation du score Apgar en HL7/V3

Nous avons réutilisé la modélisation publiée par Goosen [10]. Il s'agit d'un D-MIM développé pour ce domaine dans le cadre du projet national du dossier médical informatisé néerlandais [10]. Notons que ce modèle n'a pas à ce jour été implémenté. Au sein de ce D-MIM, on trouve la modélisation du score d'APGAR (figure 1).

On remarque que le score d'Apgar dérive de la classe ACT [x_Apgar_score] ❶. L'attribut class_cd<==OBS ajoute une contrainte à la classe ACT et permet de spécifier qu'il s'agit d'une observation. Le mood_cd (ou mood code) exprime le contexte de la classe. Il s'agit ici d'une mesure ponctuelle (attribut EVN pour évènement). Les pointillés modélisent la notion de choix exclusifs (« *check box* » en anglais), c'est-à-dire que le score d'Apgar est un acte qui peut être effectué au choix à 1 minute, 5 minutes et 10 minutes.

Le score d'APGAR [x_Apgar_Score] est relié au dossier de naissance [A_birth_Data] via une relation récursive du RIM (Act_relationship) ❷, en l'occurrence [AR_Has_apgar_score] qui ici est de type composition (<== COMP). Dans cette modélisation, on dissocie la classe [X_Apgar_score], relative à une mesure à un moment donné, de la classe [A_Apgar_score] ❸, qui représente le score proprement dit. Cette dernière classe est liée à la première par une relation de dérivation (DRV).

On retrouve les contraintes au niveau des attributs permettant d'exprimer le score : par exemple la valeur bornée de 0 à 10 et l'expression mathématique de son calcul. La classe A_Apgar_score dérive elle-même de classes spécifiques, modélisant les paramètres cliniques mesurés, la fréquence cardiaque par exemple ❹.

La cardinalité figure pour chaque relation entre classes. Le score d'APGAR est dérivé d'une et une seule classe observation [A_Heart_Frequency_score].

En ce qui concerne l'alignement, seules deux classes sont reliées à un référentiel sémantique, en l'occurrence LOINC ❺. Il existe une alternative à la modélisation de la classe [x_Apgar_Score] et de la classe [A_Apgar_score]. Une seule classe aurait pu être

produite, dont l'attribut CD aurait été valorisé par un code LOINC correspondant au temps de l'observation (1, 5, 10 minutes).

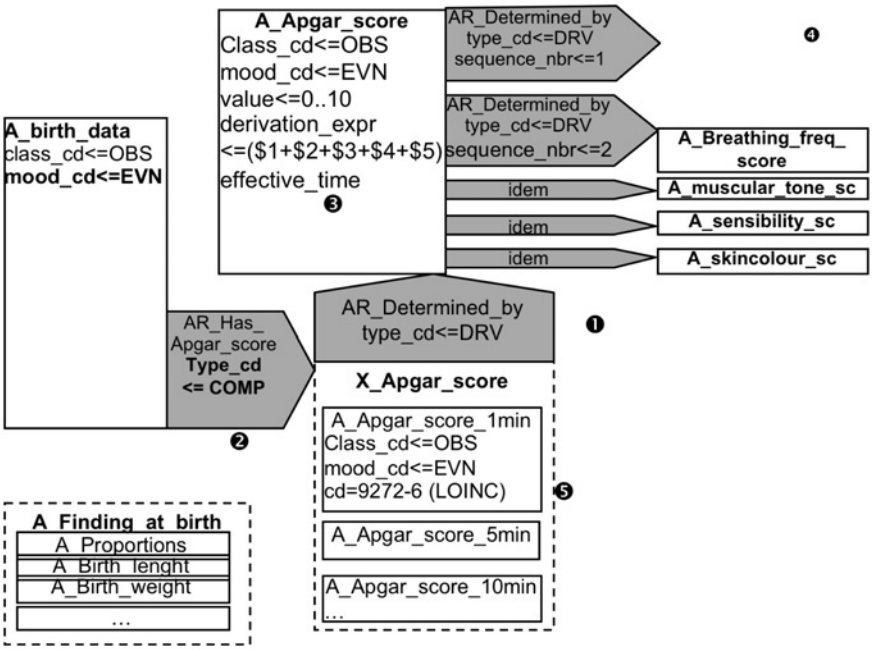


Figure 1 - Détail de la modélisation du score d'APGAR dans le DMIM de périnatologie modélisé par Goossen et al.

On aurait également pu ne pas détailler le score, en ne gardant que le total, mais on a choisi de représenter les différentes composantes du score. Cette information est utile au clinicien, lorsqu'il veut connaître l'origine de la perturbation du score.

3.2 Analyse de la modélisation du score APGAR en OpenEHR/EN13606

Comme la plupart des archétypes « cliniques », l'archétype décrivant le score d'Apgar [openEHR-EHR-OBSERVATION.apgar.v1] dérive de la classe OBSERVATION qui, elle-même, dérive de la classe ENTRY.

On retrouve les quatre parties de modélisation d'un archétype (Description, Data, Protocol, Event Series). La partie « Description » contient un texte de définition de l'Apgar ; la partie « Protocole » contient sous forme textuelle la procédure de mesure du score ; la partie « Data » regroupe les items cliniques à examiner (*heart rate, breathing* etc.) ; la partie « Event Series » modélise la temporalité du recueil de l'Apgar (1, 2, 5,10 minutes).

Pour comprendre la modélisation de l'archétype, il faut retourner à son codage en ADL. Apgar Score dérive bien de la classe OBSERVATION①. Apgar est une mesure historisée (History)②, de type mesure ponctuelle (POINT_EVENT)③. Par défaut, la mesure se fait à une minute (argument |PTM1| pour Point Time 1 minute)④.

Puis sont définis les items à recueillir. Les éléments de données sont organisés dans une structure de type ITEM_ LIST⑤, chaque élément est associé à une cardinalité (par exemple, pour le rythme cardiaque, elle est de 0 ou 1), à un identifiant permanent au sein de l'archétype qui servira éventuellement à l'alignement (par ex: [at0005]) ⑥, enfin à un ensemble de valeurs (*value*) associées également à un identifiant.

```
definition
  ① OBSERVATION[at0000] ∈ {    -- Apgar score
      data ∈ {
        ② HISTORY[at0002] ∈ {    -- history
            events cardinality ∈ {1..*; unordered} ∈ {
                POINT_EVENT[at0003] occurrences ∈ {0..1} ∈ {    -- 1 minute    ③
                    offset ∈ {
                        DV_DURATION ∈ {
                            value ∈ {|PT1M|}    ④
                        }
                    }
                    data ∈ {
                      ⑤ ITEM_LIST[at0001] ∈ {    -- structure
                          items cardinality ∈ {1..6; ordered} ∈ {
                            ⑥ ELEMENT[at0005] occurrences ∈ {0..1} ∈ {    -- Heart rate
                                value ∈ {
                                    0||local::at0006],     -- No heart beat
                                  ⑦ 1||local::at0007],     -- Less than 100 beats per minute
                                    2||local::at0008]      -- Greater than or equal to 100 beats per minute
                                }
```

Par exemple, la valeur 0 est attribuée à l'identifiant [AT006]⑦, qui lui-même sera défini plus loin, d'abord textuellement comme « pas de battement cardiaque » et éventuellement relié à un concept d'une terminologie de référence, ce qui n'est pas le cas ici.

Le score total est également défini mais, contrairement à la modélisation HL7, il est mis au même niveau que les autres éléments de données. De plus, ne figure pas d'expression mathématique permettant de le calculer (simplement les bornes de 0 à 10).

```
ELEMENT[at0025] ∈ {    -- Total
    value ∈ {
        DV_COUNT ∈ {
            magnitude ∈ {|0..10|}
```

Pour exprimer les différents temps de mesure du score (1, 2, 5, 10 minutes), les éléments de données ne sont pas redéfinis pour chaque temps. Un adressage (use_node) interne permet de renvoyer à la définition de la mesure.

```
POINT_EVENT[at0026] occurrences ∈ {0..1} ∈ {    -- 2 minute
    offset ∈ {
        DV_DURATION ∈ {
            value ∈ {|PT2M|}
        }
    }
    data ∈ {
        use_node ITEM_LIST /data[at0002]/events[at0003]/data[at0001]
    }
```

Dans la partie dite « ontologique » de l'archétype, on retrouve l'identifiant [atxxx] de chaque élément de donnée, associé à un libellé et à une définition textuelle.

```
["at0006"] = <
    description = <"No heart beat is present - palpation of the base of umbilical cord">
    text = <"No heart beat">
```

Enfin, concernant le lien avec les terminologies externes, comme nous l'avons vu, seuls deux éléments de données sont reliés à LOINC.

```
term_binding = <
    ["LNC205"] = <
        items = <
            ["/data[at0002]/events[at0003]/data[at0001]/items[at0025]/"] = <[LNC205::9272-6]>
            ["/data[at0002]/events[at0028]/data/items[at0025]/"] = <[LNC205::9271-8]>
```

À partir de l'éditeur d'archétype distribué par la société Ocean Informatics[1], nous avons généré automatiquement un formulaire qui permet de mieux se représenter comment les différents éléments de données sont organisés dans l'archétype Apgar.

[1] https://wiki.oceaninformatics.com/confluence/display/TTL/Archetype+Editor+Releases

4 Discussion

4.1 Le formalisme des modèles :

Sur le plan du formalisme, HL7V3 présente une cohérence bien supérieure à OpenEHR. En effet, le modèle HL7 du score APGAR s'intègre dans la modélisation globale d'un domaine (ici le DMIM sur la périnatalogie). Nous avons vu plus haut que le DMIM lui-même dérive du RIM. Cette structure, en un seul tenant, donne une très grande cohérence à l'ensemble. Encore faut-il que soient respectées les règles de bonne pratique de dérivation du RIM en DMIM, c'est-à-dire se basant sur la notion de domaine fonctionnel et non sur d'autres critères.

Le formalisme OpenEHR est plus « lâche ». Si le lien avec le modèle de référence est formel, ce lien est sémantiquement pauvre : d'une part, parce que le modèle de référence d'OpenEHR est beaucoup moins fourni que le RIM et d'autre part, parce qu'il n'y a pas de modélisation globale d'un domaine.

Ce manque de formalisme se retrouve au sein même des archétypes : les éléments de données sont regroupés « en vrac ». Il n'y pas, par exemple, de hiérarchisation entre le score total d'Apgar et les différents paramètres permettant de le mesurer. HL7 introduit non seulement la relation de dérivation entre score global et paramètres du score, mais en plus spécifie l'expression mathématique permettant son calcul.

4.2 La sémantique véhiculée

L'archétype Apgar contient de nombreuses informations descriptives (objectifs, protocoles, partie « ontologie » définissant les éléments de données, etc.). Cette documentation est supérieure au modèle HL7, mais malheureusement elle est textuelle et peut difficilement être exploitable par les machines. Les deux modèles encapsulent des éléments de données, les regroupent et les structurent, générant de fait une sémantique contextuelle. Par exemple, dans l'archétype Apgar, nous avons un élément de donnée [at0014] correspondant à l'atonie (Limp=mou) (Figure 2).

```
ELEMENT[at0013] occurrences ∈ {0..1} ∈ {     -- Muscle tone
    value ∈ {
        0|[local::at0014],      -- Limp
        1|[local::at0015],      -- Some flexing of arms and legs
        2|[local::at0016]       -- Active motion
```

Figure 2 - Représentation du tonus musculaire dans l'archétype APGAR.

Cet élément de données ne peut être interprété correctement qu'en contexte avec l'ensemble des autres éléments de données. Il s'agit de l'atonie, comme état du tonus musculaire, mesuré pour le calcul du score d'Apgar. Le DMIM nous indique de plus que cette mesure est un composant du dossier de périnatalité. Ainsi, les archétypes ou les DMIM modélisent des contextes.

4.3 Un faible taux d'alignement

Les deux exemples montrent un faible lien avec des terminologies de référence. Pour chacun des modèles, seuls 2 éléments de données ont été codés en LOINC. Le choix de cette terminologie n'a jamais été justifié. L'Apgar étant un score basé sur des observations cliniques, la Snomed CT aurait été sans doute la bonne terminologie de référence.

4.4 Le télescopage des modèles d'information et de connaissance.

Nous illustrons notre propos par un exemple : dans l'APGAR, nous retrouvons l'élément de données « tonus musculaire ». Que ce soit pour HL7 ou OpenEHR, cet élément hérite de la classe `OBSERVATION`.

Or, comme le montre la Figure 3, Snomed CT présente deux concepts *tonus musculaire* `[6918002 : muscle tone]` et `[6918002 : muscle tone finding]` héritant de classes strictement différentes `[363787002 : Observable entity]` et `[404684003 : clinical finding]`.

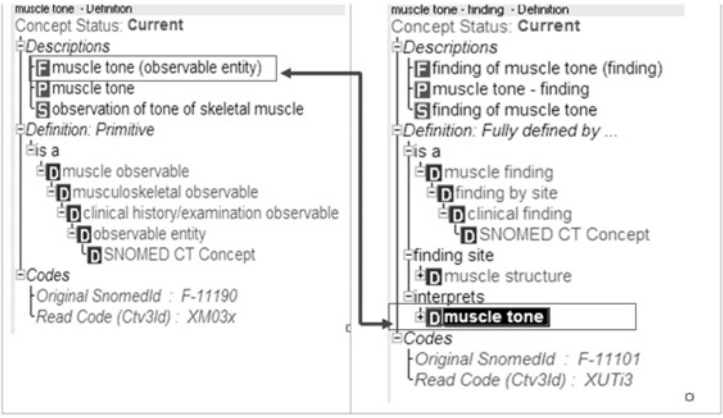


Figure 3 - Double représentations du concept tonus musculaire dans Snomed CT.

Il s'agit bien à chaque fois du tonus musculaire, mais dans le premier cas, il est considéré comme un phénomène potentiellement observable (on peut imaginer une question « tonus musculaire ? »), dans l'autre cas il s'agit du phénomène réellement présent, c'est-à-dire du signe (*finding*) que l'on a recueilli (le tonus musculaire du patient). Les deux concepts sont liés par une relation formelle « interprets ».

On voit donc que modèles d'information et de connaissance représentent des concepts identiques de deux façons différentes. Dès lors, comment réaliser sans ambiguïté l'alignement ? Pour l'archétype Apgar, on peut considérer que l'élément de données [at0013] constitue une question « Quel est la tonicité du bébé ? ». Il pourrait de fait être lié au concept observable `[6918002 : muscle tone]`.

Les réponses possibles pourraient être codées avec des fils du concept sémiologique `[6918002 : muscle tone finding]` et, s'ils n'existent pas, avec de nouveaux concepts post-coordonnés.

Le télescopage entre modèles d'information et de connaissance a été identifié par Qamar et Rector [11, 12]. Avec l'alignement, nous nous trouvons précisément à mettre en liaison deux modèles formels, celui du modèle d'information et celui de connaissance. Nous considérons en effet Snomed CT comme structure et ressource de représentation formelle des connaissances du domaine médical.

Ce problème, bien connu des ontologistes, pourrait être résolu par deux approches complémentaires et qui ont pour objectif de réunifier les deux modèles.

La première consiste à réaligner les 2 modèles avec un troisième dit de haut niveau, on parle ici *d'upper ontology*. Par exemple, il s'agirait ici de réaligner le RIM d'HL7 et les 19 classes primitives de la Snomed CT, en utilisant une des ontologies de haut niveau actuellement disponible, comme l'ontologie Dolce par exemple [13].

L'autre méthode dite de mapping ou ontologique vise à trouver des équivalences de classes directement entre les deux modèles. Un travail exploratoire récent [14] a entamé le mapping entre HL7 et Snomed CT.

4.5 Le problème de la modélisation du contexte

Comme nous l'avons vu, le modèle d'information génère une sémantique contextuelle qui, de fait, rend les éléments de données décrits à la fois dépendants et indissociables. Cette sémantique contextuelle apporte une nouvelle connaissance opérationnelle qui n'est pas forcément présente dans les modèles de connaissance ontologiques.

Par exemple, dans l'archétype APGAR, on trouve la procédure pour la mesure du score, ainsi que les différents éléments de recueil. Lorsqu'on examine la modélisation du score APGAR en Snomed CT (figure 4), on constate qu'elle consiste en une taxonomie des différentes valeurs du score aux différents temps. Cette structuration permet de classer les scores, mais ne contient aucun concept relatif à la façon de le mesurer.

Figure 4 - Représentation conceptuelle du score d'APGAR en Snomed CT.

Les modèles d'information apportent donc de nouvelles connaissances contextuelles opérationnelles du domaine. L'interopérabilité sémantique nécessite la prise en compte de ces connaissances.

4.6 Le problème du vocabulaire propre au modèle d'information

Les modèles d'information ne sont donc pas uniquement des modèles de structuration syntaxique de l'information, mais modélisent de fait une sémantique. On a vu que cette sémantique est théoriquement apportée par le modèle de connaissance, via l'alignement.

En réalité, ce n'est pas systématiquement le cas. Que ce soit pour HL7 ou OpenEHR/EN13606, les deux normes ont développé leur propre vocabulaire du domaine.

Par exemple, les classes filles de ENTRY du modèle de référence d'Open EHR sont appelées Observation, Évaluation, Instruction, Action avec un vocabulaire proposé par la norme. Dans les archétypes, un vocabulaire plus spécialisé est défini de façon non formelle.

C'est également le cas pour HL7. Les classes situées très haut dans la hiérarchie du RIM possèdent des attributs déjà sémantiquement très spécifiques et qualifiés, avec un vocabulaire propre à la norme : par exemple, la classe « *entity* » du RIM possède l'attribut « *numéro de téléphone* ».

4.7 La gouvernance de la production des modèles

Cette question est aujourd'hui essentielle et conditionne directement la capacité à communiquer des systèmes d'information en santé de la prochaine décennie. Comment coordonner la production des modèles d'information et des référentiels sémantiques à grande échelle ?
Il s'agit de couvrir l'ensemble du domaine médical par la construction de milliers d'archétypes et/ou de centaines de DMIM, de Templates, de CMET, etc. Les référentiels terminologiques devront également évoluer vers plus de formalisme et intégrer de nouveaux domaines, comme la génomique ou la médecine translationnelle.
Cette tâche colossale sera entravée par plusieurs facteurs que nous énumérons ci-dessous.

4.8 La recherche d'un consensus

La connaissance médicale présente ceci de particulier qu'elle est complexe à modéliser. Les pratiques et les points de vue varient en fonction des pays, des spécialités médicales, voire des individus. Par conséquent, la recherche d'un consensus, à un niveau national ou international, pour structurer l'information dans un des deux modèles, semble être une gageure.

4.9 L'évolution des connaissances

La connaissance biomédicale évolue de façon importante. De nouveaux domaines, tels que la génomique, apparaissent et introduisent de nombreux nouveaux concepts qui modifient les représentations antérieures que nous avions du vivant. Comment alors faire évoluer les systèmes pour qu'ils puissent véhiculer ces nouveaux concepts ?

5 Conclusion

Les systèmes d'information connaissent un essor de plus en plus important dans le monde. Les projets e-santé développés dans la plupart des pays industrialisés aboutiront à plus ou moins long terme à des dossiers électroniques partagés, alimentés par les différents systèmes d'information. Pour atteindre ces objectifs, il faudra lever les verrous posés par l'interopérabilité. L'adoption d'un ensemble de normes et de standards par les industriels est la condition sine qua non pour aboutir à une intégration à grande échelle.

HL7V3 et OpenEHR/EN13606 sont les deux normes qui se dégagent actuellement d'un paysage normatif encore trop complexe. L'apport des archétypes d'openEHR/EN13606 est indéniable, mais il persiste des problèmes liés à leur formalisme et à leur production. Il est donc raisonnable de penser qu'HL7 s'imposera à terme, d'une part en raison des moyens considérables qui lui sont consacrés et d'autre part, du fait que l'on observe une

convergence généralisée des autres normes et standards vers la V3.

Les nouveaux cadres normatifs (modèles d'information ou terminologiques) se caractérisent par leurs structures formelles. Nous avons pu illustrer au travers de ce travail que de nombreuses interrogations persistent quant à leur mise en œuvre. Nous pensons que l'adoption de ces standards va impacter en profondeur les systèmes d'information et demandera aux éditeurs de logiciels un effort très conséquent de réingénierie posant alors la question du rapport bénéfice/coût.

Références

[1] Rector AL. Clinical terminology: why is it so hard. *Methods of Information in Medicine* 1999; 38(4-5): 239-252.

[2] Dolin RH, *et al.* HL7 Clinical Document Architecture, Release 2. *J Am Med Inform Assoc* 2006; 13(1): 30-9.

[3] CEN. Comité de Normalisation Européen. 2008 [cited 2008; Available from: http://www.cen.eu/cenorm/homepage.htm.

[4] Beale T. Archetypes and the EHR. *Stud Health Technol Inform* 2003; 96: 238-44.

[5] Apgar V. The newborn (Apgar) scoring system. *Pediat Clin N Am* 1966; 13: 645-650.

[6] Blobel BG, Engel K, Pharow P. Semantic interoperability--HL7 Version 3 compared to advanced architecture standards. *Methods Inf Med* 2006; 45(4): 343-53.

[7] Beale T. Archetypes: Constraint-based Domain Models for Future-proof Information Systems. OOPSLA 2002 workshop on behavioural semantics, 2002.

[8] Beale, T. and S. Heard, Archetype Definition Language (ADL). OpenEHR specification, the openEHR foundation, 2005.

[9] Meyer B. Eiffel: a language and environment for software engineering. *Journal of Systems and Software* 1988; 8(3): 199-246.

[10] Goossen WT, *et al.* Electronic patient records: domain message information model perinatology. *Int J Med Inform* 2003; 70(2-3): 265-76.

[11] Qamar R, Kola J, Rector AL. Unambiguous data modeling to ensure higher accuracy term binding to clinical terminologies. *AMIA Annu Symp Proc* 2007; 608-13.

[12] Qamar R, RectorAL. Semantic issues in integrating data from different models to achieve data interoperability. *Stud Health Technol Inform* 2007; 129(Pt 1): 674-8.

[13] Gangemi A, *et al.* Sweetening Ontologies with DOLCE. *LECTURE NOTES IN COMPUTER SCIENCE* 2002; 166-181.

[14] Ryan A. Towards semantic interoperability in healthcare: ontology mapping from SNOMED-CT to HL7 version 3. 2006: Australian Computer Society, Inc. Darlinghurst, Australia.

Adresse de correspondance

Marc Bayat

Département d'Information Médicale

CHU de Rennes, Pontchaillou

2 rue Henri Le Guilloux

35033 Rennes cedex, France

E-mail : marc.cuggia@chu-rennes.fr

Chapitre 4

Modèles, méthodes et outils pour l'amélioration des pratiques

Informatique et Santé, 2009 (17) : 139-150
© Springer – Verlag France

Étude contrastive des documents vulgarisés et scientifiques de santé : sur la piste de la morphologie

Jolanta Chmielik[1,2], Natalia Grabar[1,3]

[1]INSERM, UMR_S 872, eq.20 Paris, France ; Université Paris Descartes, Paris, France
[2]Commission européenne, Délégation à la recherche, Bruxelles, Belgique
[3]DIH, HEGP, AP-HP, Paris, France

Abstract

The amount of the health information on Internet is constantly growing although little is done to guide users towards documents which are the most appropriate to their expertise level. One of the distinctions is concerned with the specialization and technical level of documents: indeed documents differently specialized are co-existing while this difference is not explicitly indicated to users. The objective of this work is to propose clues for the automatic distinction between scientific and popularized medical content. Particularly, we propose to study the morphological level: morphological units, like bases and affixes, considered through their lexemes and morphological families. We apply the NLP tools, which provide access to the morphological content of documents. Our experiments and results highlight that this level may indeed provide clues for the distinction of the specialization of documents, and that they appear to be significant and stable across medical fields studied. Our work is done with French documents within three medical fields (cardiology, pneumology and diabetes).

Keywords

Medical Informatics; Access to Information; Language; Natural Language Processing; Documentation.

1 Introduction

Une analyse récente de l'utilisation de l'Internet [1] montre qu'environ 80 % des requêtes effectuées sont liées à la santé. Ce chiffre souligne la préoccupation que manifestent les citoyens vis-à-vis de leur santé. Il montre également que l'Internet propose des volumes d'information de santé extrêmement importants. La qualité de ces documents n'est pourtant pas égale. Les initiatives comme CISMeF [2] ou HON [3] se concentrent sur la qualité et la transparence éthique des pages de santé : les documentalistes et experts du domaine médical effectuent une analyse des pages à indexer ou à certifier. Mais par ailleurs, les documents de santé montrent aussi une différence quant à leur spécialisation et technicité. Ainsi, les documents scientifiques spécialisés (créés par des professionnels de santé à destination des professionnels de santé) et les documents vulgarisés (créés à destination des usagers non experts en médecine) co-existent. À ceci s'ajoutent aussi les documents médicaux de type didactique qui s'adressent aux étudiants en médecine. En fonction de la technicité des documents, qui se traduit souvent par les termes et les mots employés et les

contextes où ils apparaissent, les documents peuvent être plus ou moins difficiles à comprendre par des utilisateurs non experts en médecine. Cette hétérogénéité technique n'est pas transparente aux usagers, alors qu'il pourrait être utile de l'indiquer de manière explicite.

Il a été ainsi constaté que le niveau de technicité peut avoir un impact sur la compréhension des informations de santé par les non-spécialistes et, par la suite, sur la communication des patients avec les professionnels de santé et la réussite des soins médicaux qui leur sont administrés [4-5]. Afin de guider les utilisateurs non experts vers des sources d'information qui leur sont plus appropriées, les moteurs de recherche pourraient faire la distinction entre les documents spécialisés et vulgarisés. Notons que les portails médicaux comme HON [3], CISMeF [2] ou le moteur de recherche général GoogleCoopsanté [6], proposent cette distinction mais elle est essentiellement basée sur une catégorisation manuelle de pages et de sites Internet. Dans la situation actuelle, où le volume d'information augmente sans cesse sur l'Internet, il est souhaitable de chercher à faire la distinction entre les documents scientifiques et vulgarisés automatiquement. Pour cela, il faut tout d'abord déterminer les critères sur lesquels cette distinction peut être basée.

Nous faisons l'hypothèse qu'une description et une analyse détaillées de différents niveaux linguistiques des documents scientifiques et vulgarisés peuvent fournir des critères complémentaires pour la détection des niveaux de spécialisation des documents. Dans ce travail, notre intérêt porte plus spécifiquement sur le niveau morphologique, qui de plus ne semble pas avoir été étudié dans la littérature. L'objectif que nous poursuivons concerne donc une étude des caractéristiques morphologiques des documents médicaux, vulgarisés et experts, et leur utilisation pour l'émergence et la sélection de critères morphologiques qui pourraient servir pour la distinction automatique de discours médicaux. Une telle étude se présente en effet comme une étape préalable à une classification automatique de discours. Nous utilisons les outils du traitement automatique de langues.

Dans la suite de ce papier, nous présentons d'abord l'état de l'art, ensuite le matériel et les méthodes appliquées. Nous présentons et discutons les résultats et terminons avec des perspectives sur ce travail.

2 État de l'art

Il existe plusieurs travaux en distinction automatique du niveau de spécialisation dans le domaine médical. Les formules linguistiques de lisibilité [7-9] prennent en compte les critères comme la longueur moyenne de mots ou de phrases. Ainsi, plus les mots sont longs, plus ils sont considérés comme savants. Ces formules peuvent être combinées avec le vocabulaire spécialisé (*ie,* MeSH [10]), ce qui permet de prendre en compte la dimension médicale des documents du domaine, où les mots courts peuvent aussi être difficiles à comprendre [11]. Les approches qui appliquent les algorithmes d'apprentissage sont également très répandues et se basent sur différents types de descripteurs (n-grams de caractères [12], pondération manuelle [13] ou automatique [14] des termes MeSH, critères stylistiques [15] ou discursifs [16] des documents, leur niveau lexical [17]). Ce domaine d'activité est extrêmement actif actuellement et, de plus en plus, l'accent est mis sur la combinaison de différents descripteurs [18-20]. En revanche, des études approfondies d'un type de descripteurs restent rares. A cet effet, citons par exemple l'alignement de vocabulaires scientifiques et vulgarisés [21, 22] ou l'étude assez détaillée du niveau syntaxique [23].

3 Matériel et méthodes

Dans les sections 3.1 à 3.3, nous présentons la manière dont nous avons constitué et sélectionné le matériel morphologique qui nous servira de base à ce travail. Nous décrivons ensuite (sec. 3.4) les méthodes proposées pour l'étude contrastive des discours médicaux spécialisés et vulgarisés. Nous utilisons les outils du traitement automatique de langues et des outils et programmes Perl.

3.1 Corpus de documents scientifiques et vulgarisés

Notre travail est effectué sur les documents en français. Ces documents sont collectés au travers du portail CISMeF, qui propose des ressources satisfaisant les critères de qualité de l'information de santé. Nous exploitons deux types d'informations sur chaque ressource recensée par CISMeF : (1) leur caractérisation selon les domaines médicaux effectuée grâce à l'indexation avec les mots-clés MeSH. Nous nous intéressons à trois spécialités médicales : cardiologie, pneumologie et diabète ; (2) leur caractérisation selon les types de discours et les types de pratiques sociales distinctes : grand public, professionnels de santé, et étudiants en médecine. Les ressources indexées sont proposées par CISMeF sous forme d'adresses URL téléchargeables. Les documents ont été téléchargés automatiquement à partir des listes d'URL, à l'aide d'un « aspirateur de Web » wget, paramétré pour télécharger les URL et les URL filles (par exemple, pour accéder aux documents fractionnés en plusieurs pages). Les documents ainsi téléchargés ont été ensuite filtrés pour sélectionner les documents au format HTML ou XML, que l'on peut convertir en texte plus facilement. Grâce à l'annotation CISMeF, ces documents ont été regroupés en des ensembles différenciés en fonction de leur spécialisation médicale et de leur étiquetage comme expert, étudiant et vulgarisé. Au total, nous disposons des trois corpus : cardiologie, pneumologie et diabète, dont chacun se compose de trois parties : expert, étudiant et vulgarisé.

Tableau 1 - Taille des trois corpus étudiés.

Domaine	Discours	Nombre de documents	Nombre d'occurrences
Cardiologie	expert	1 583	942 409
	étudiant	205	449 765
	vulgarisé	143	157 382
	total	1 931	1 549 556
Pneumologie	expert	724	600 524
	étudiant	127	213 379
	vulgarisé	134	96 559
	total	985	910 462
Diabète	expert	213	181 039
	étudiant	23	44 847
	vulgarisé	52	29 817
	total	236	255 703

Le tableau 1 indique la taille de ces corpus. La colonne *Nombre de documents* indique le nombre de documents réunis dans chaque corpus (*Total* indique la somme des trois discours réunis). La colonne *Nombre d'occurrences* indique le nombre d'occurrences dans

chaque ensemble. Nous pouvons voir que la taille des corpus varie d'une spécialité médicale à l'autre et d'un type de discours à l'autre. Le corpus cardiologie est le plus volumineux de tous, et le corpus diabète est le plus petit. Par ailleurs, en général, les textes « expert » sont les plus nombreux dans chaque spécialité. Les documents « étudiant » viennent en deuxième position et, derniers en termes de volume, arrivent les textes vulgarisés.

Ce matériel, constitué au travers le portail CISMeF, nous sert de gold standard dans notre étude. Les différences dans la taille des corpus correspondent sans doute à la réalité Internet, mais peuvent éventuellement introduire des biais dans les résultats que nous pourrons obtenir. Nous supposons par contre que la caractérisation des documents selon les types de ressources (*article de périodique, rapport technique, recommandation, brochure pédagogique pour les patients, information patient et grand public, cours, cas clinique, etc.*) ou l'indexation avec les mots-clés MeSH, étant effectuée par des professionnels du domaine, est fiable.

3.2 Accès au niveau morphologique des documents textuels

Afin d'accéder au niveau morphologique de textes de nos corpus, nous appliquons des outils du traitement automatique de langues : (1) l'étiqueteur TreeTagger [24] assigne à chaque mot des documents une étiquette morpho-syntaxique et effectue une lemmatisation. Par exemple, *angioblastiques* est étiqueté comme adjectif (*ADJ*) et lemmatisé en singulier *angioblastique/ADJ*, *antiinflammatoire* est étiqueté comme pronom (*PRO*), etc. ; (2) le lemmatiseur Flemm [25] reprend la lemmatisation proposée et la corrige ou bien ajoute des traits morphologiques supplémentaires. Ainsi, grâce à Flemm, l'étiquetage de *antiinflammatoire/PRO* est corrigé en adjectif ; (3) enfin, l'analyseur morphosémantique DériF [26] effectue l'analyse des termes en fonction de leur structure morphologique et sémantique. Il fonctionne en quatre étapes, que nous illustrons sur l'exemple de *angioblastique/ADJ* :

1. Calcul de l'arbre d'analyse d'un lemme étiqueté :

   [[angi N] [blast N*] ique ADJ]*
2. Reprise de l'arbre sous forme de famille ordonnant l'ensemble des bases successives reconnues par l'analyseur :

   (*angioblastique/ADJ*, [angi,N*]:blast/N*)
3. Représentation en langage naturel de la relation sémantique entre le lemme et ses bases (glose sémantique) :

   " Qui est en relation avec cellule embryonnaire et vaisseau "
4. Description d'autres traits sémantiques acquis automatiquement :

   Constituants = /angi/blast/ique

   Type = anatomie

   Relations possibles = (eql:ang/blast, eql:angé/blast, eql:angéio/blast, eql:vas/blast, eql:vascul/blast, isa:ang/cyt, isa:angi/cyt, isa:angé/cyt, isa:angéio/cyt, isa:vas/cyt, isa:vascul/cyt, see:ang/bréph, see:angi/bréph, see:angé/bréph, see:angéio/bréph, see:vas/bréph, see:vascul/bréph)

La partie de la méthode qui suit exploite cette analyse morphosémantique pour le choix des bases qui seront analysées et pour la constitution de familles morphologiques.

3.3 Préparation de matériel morphologique

Nous avons ainsi considéré qu'un des critères de différentiation au niveau morphologique entre le discours des spécialistes et celui du grand public concerne l'usage qui est fait des bases : d'une part les lexèmes qu'elles permettent de former, d'autre part la systématicité et la fréquence avec laquelle ces unités sont utilisées dans les documents analysés.

3.3.1 Sélection de bases

Nous avons sélectionné, pour chaque type de discours et pour chaque spécialité médicale, des bases spécifiques à ces corpus et représentant les plus grandes familles morphologiques. Ainsi par exemple la base *cardio*, appartenant au sous-domaine de la cardiologie, forme un nombre de lexèmes parmi les plus élevés dans le corpus cardio-étudiant : 57 lexèmes. Sur ce critère nous avons retenu *cardio* et comparé sa productivité dans les corpus cardio-expert et cardio-vulgarisé (respectivement, 26 et 20), de la même manière qu'au travers des corpus de pneumologie et diabète, qui nous servent de données témoins pour cette base.

3.3.2 Constitution de familles morphologiques

Une famille morphologique regroupe l'ensemble des lexèmes reliés par des relations formelles et sémantiques au travers des règles de formation de lexèmes. Dans notre cas, chaque famille morphologique est formée autour d'une base. Nous nous basons pour ceci sur l'analyse morphosémantique du lexique par Derif. L'analyse morphosémantique de chaque lexème est pris en compte, comme la notation entre les crochets dans ces exemples :

angioblastique/ADJ: [[angi N] [blast N*] ique ADJ]*

cardiogénique/NOM: [[[cardio N] [gène V*] ique ADJ] NOM]*

électrocardiographie/NOM: [[électr N] [[cardio N*] [graphie N*] NOM] NOM]*

echocardiographie/NOM: [[echo N] [[cardio N*] [graphie N*] NOM] NOM]*

cardiovasculaire/ADJ: [[cardio N] [vascul N*] aire ADJ]*

angiocardiographie/NOM: [[angi N] [[cardio N*] [graphie N*] NOM] NOM]*

Les lexèmes composés, comme *angioblastique* dans les exemples présentés, appartiennent à autant de familles que le nombre de bases qu'ils comportent. Ainsi, *angioblastique* enrichit deux familles (*angi, blast*), et *cardiogénique* enrichit aussi deux familles (*cardio, gène*), etc.

3.4 Méthodes pour l'étude contrastive des discours

Afin de faire émerger des caractéristiques morphologiques saillantes, nous abordons l'analyse contrastive des discours médicaux de deux points de vue : au travers de la taille de familles morphologiques et au travers des fréquences que les bases sélectionnées montrent en corpus. Nous pensons en effet que les critères, issus de ces deux types d'observations, peuvent être détectés automatiquement en corpus et montrer une spécificité suffisamment importante pour la distinction automatique de discours.

La taille de familles morphologiques est étudiée au travers du nombre de lexèmes que chaque base étudiée forme. Ainsi, plus une base est productive, plus sa famille morphologique sera grande. Nous prenons en compte deux valeurs pour mesurer la taille de familles : brute et normalisée. La taille brute correspond exactement au nombre de lexèmes formés. La taille normalisée représente une valeur normalisée par la taille des corpus.

Les fréquences de lexèmes en corpus sont étudiées sur les données étiquetées et lemmatisées par Flemm (pour éviter le biais de la variation des lexèmes). De la même manière, les fréquences des lexèmes sont observées au travers des valeurs brutes et normalisées.

4 Résultats et discussion

4.1 Matériel morphologique sélectionné

Les bases sélectionnées à partir des données de Derif appartiennent à deux catégories morphologiques : les bases dites supplétives ou savantes, qui n'apparaissent pas de manière isolée dans la langue mais toujours couplées avec d'autres éléments morphologiques (p. ex : *gastr(o)* qui se réalise par exemple au travers *gastrique*) et les bases autonomes (par ex. : *bronches, bactérie*). Nous avons ainsi 37 bases supplétives (*pulmon, vascul(o), vas(o), muscul, séro, artério, angi(o), athér(o), endo, uro, myo, cardio, ite, patho, pharyngo, gastr(o), glyc(o), graphie, hépat(o), néphr(o), phléb(o), psych(o), lip(o), neur(o), thromb(o), osté(o), pneum(o), cyt(o), gène, rhin(o), hém(o), ose, ectomie, tomie, scopie, ome, oïde*) et 7 bases autonomes (*lipide, diabète, insuline, veine, bronche, bactérie, thérapie*). Ces bases sont différemment représentées dans les corpus (spécialités, discours) étudiés : certaines sont répandues dans tous les corpus, d'autres restent spécifiques (en termes d'occurrence) à une spécialité ou un discours.

Lors de la sélection des bases, nous nous sommes rendus compte de quelques limites dans l'analyse morphosémantique de DériF, mise à part l'incomplétude de son lexique de référence (par rapport auquel l'existence des lexèmes et des bases est identifiée) et la couverture partielle des bases et affixes (comme *-eux*) du français analysés par Derif. Ainsi, Derif propose parfois une analyse morphosémantique erronée, comme dans l'exemple du lexème *gymnase/NOM,* analysé comme l'*« enzyme du nu »* :

gymnase/NOM: [[gymno A] [ase N*] NOM]*

" (Partie de -- Type particulier de) enzyme caractérisé par la propriété : nu "

Constituants = /gymno/ase/

Type = produit chimique

Une autre limite, qui avait plus de répercussion sur notre travail, concerne la notation des bases reconnues. Dérif identifie et regroupe les bases selon leur forme (chaîne de caractères). Par exemple, les quatre bases suivantes (*hém(o)*, *héma*, *hémat(o)* et *èm*) proviennent du même mot grec *haima*, signifiant « *relatif au sang* ». Dérif les segmente correctement mais continue de les différencier formellement. Ainsi, trois bases sont distinguées dans l'analyse morphologique des lexèmes : [hém] (*hém(o), héma*), [hémato] et [èm]. Un regroupement de familles morphologiques effectué sur le critère formel va inévitablement former trois familles morphologiques distinctes, alors que le sens véhiculé par ces bases est le même. D'ailleurs, la glose sémantique associée reste la même (*relatif au sang, liée au sang*). Cet état de choses suggère que le regroupement de familles morphologiques devrait être basé autant sur la segmentation de lexèmes effectuée par Derif que sur la glose sémantique qu'il en propose.

4.2 Familles morphologiques

Nous avons constitué un ensemble de 46 familles morphologiques. La plupart de ces familles apparaissent dans tous les corpus. De manière générale, ces familles sont les mieux représentées dans les corpus cardiologie et pneumologie, ce qui n'est pas surprenant

car ces corpus sont beaucoup plus grands que le corpus diabète. Par ailleurs, certaines familles morphologiques occurrent dans les corpus d'une spécialité médicale parmi les trois qui sont étudiées et, par conséquent, peuvent être absentes d'autres corpus. Par exemple, la famille de la base *athér*, que l'on peut associer au domaine de cardiologie, est absente des corpus vulgarisés de diabète et de pneumologie, mais apparaît dans les discours expert et étudiant de ces spécialités. Notons que trois familles (*rhin*, *tomie* et *phléb*) sont totalement absentes dans le corpus diabète (ce qui pourrait être expliqué par la petite taille de ce corpus).

Le nombre total de lexèmes que regroupent les 46 familles est de 2 295. Mais tous ces lexèmes n'apparaissent pas dans tous les corpus. Par exemple, dans le corpus cardio-étu, la famille *angio* comporte 19 lexèmes (*macroangiopathie, angiocardiographie, angioneurotique, microangiopathie, angiocoronarographique, angiocardiographique, cinéangiographie, angiopneumographie, angiomatose, angiographique, angiopathie, angiome, angioscopie, télangiectasie, angioscanner, angiographie, angioplastie, angiographique, angioplasticien*), alors qu'elle en comporte 15 dans le corpus cardio-expert. Notons que *angiomateux* n'est pas inclus dans cette famille : le lexème n'est pas segmenté en bases et affixes ni analysé car l'affixe *-eux* n'est pas traité dans la version actuelle de l'outil. Ceci correspond à une limite de notre matériel, mais comme la situation est la même pour toutes les bases, nous supposons que cela relativise cette limite.

4.3. Étude comparative de la taille des familles morphologiques

La taille de chaque famille morphologique correspond à la productivité de la base qui permet de la fonder. Plus une famille est grande, plus la base est utilisée dans le corpus (par les locuteurs qui ont produit les documents), plus elle permet de former des lexèmes et plus son lexique est diversifié.

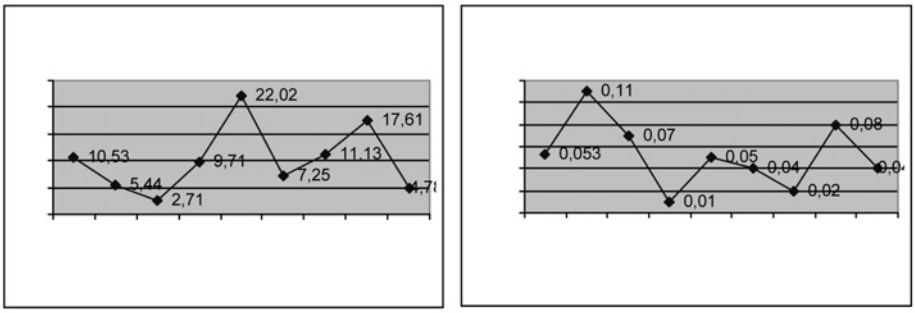


Figure 1- Taille moyenne des familles morphologiques : brute et normalisée.

La figure 1 indique la taille de familles morphologiques : brute (fig.1 gauche) et normalisée (fig.1 droite). Nous pouvons en effet observer une variabilité selon les spécialités et les discours. Nous n'allons pas nous arrêter sur les résultats observés sur le corpus diabète, car ce corpus est relativement petit, mais surtout sur les deux autres corpus. Nous pouvons ainsi remarquer que les documents étudiants montrent toujours des familles morphologiques plus grandes que les deux autres discours, que ce soit avec les résultats bruts ou normalisés. En effet, les corpus étudiants réunissent des documents didactiques pour l'enseignement de la médecine. Compte tenu de leur destination, ces textes sont généralement plus précis et traitent des sujets avec plus de détails, ce qui explique la

diversité du vocabulaire médical, et donc la taille de familles morphologiques plus grande que dans les autres discours.

Nous avons ainsi constaté que, dans toutes les trois spécialités, la taille normalisée de familles en corpus expert est inférieure à celle de textes vulgarisés (fig. 1 droite), alors que les valeurs brutes indiquent le contraire (fig. 1 gauche). L'observation d'une diversité plus grande du vocabulaire médical dans les corpus expert a été faite par un autre travail [11], où les auteurs étudient le langage médical suédois dans le domaine de maladies cardiovasculaires. Comme dans notre travail, une comparaison est effectuée entre les documents provenant des portails professionnels et les sites web grand public. Si nos deux travaux arrivent à des conclusions similaires, la différence quant au matériel utilisé persiste : nous positionnons notre travail au niveau de la morphologie (et unités lexicales correspondantes), alors que Kokkinakis *et al.* [11] étudient l'usage (et reconnaissance) des termes MeSH. Par ailleurs, lorsque nous analysons les valeurs normalisées, nous constatons que la taille de familles morphologiques normalisée est plus grande dans les textes vulgarisés que dans les textes expert. Ce cas de figure n'a pas été envisagé dans le travail cité. Justement, ce que les valeurs normalisées montrent c'est que, parmi toutes les occurrences des corpus, la proportion des bases étudiées, et du lexique correspondant, est plus grande dans les corpus vulgarisés. Une des raisons de cette situation est que la taille de corpus expert est la plus importante (tableau 1) : la normalisation de la taille des familles par la taille des corpus va inévitablement faire baisser cette valeur plus qu'elle ne le fait pour les corpus vulgarisés. En effet, les corpus vulgarisés sont en moyenne 6 fois moins volumineux que les corpus expert. Il est possible qu'une analyse de la taille de familles morphologiques dans des corpus de tailles comparables aurait aboutit à des résultats différents.

4.4 Étude comparative de l'utilisation de familles morphologiques en corpus

La figure 2 présente les fréquences normalisées d'apparition des familles morphologiques (au travers de leurs lexèmes) dans les corpus.

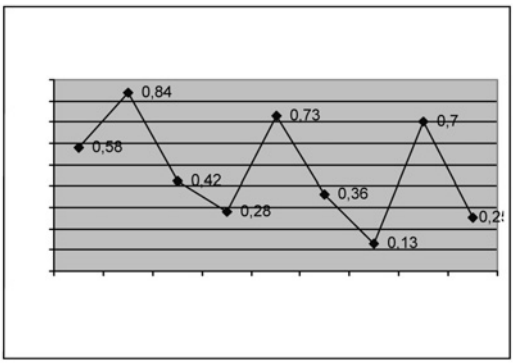


Figure 2 - Fréquence normalisée des familles morphologiques en corpus.

Dans les trois spécialités médicales, cette utilisation est la plus élevée dans les corpus étudiant (0,84 en diabète, 0,73 en cardio, 0,70 en pneumo). Dans la section précédente, nous avons vu que dans le discours étudiant les familles étaient les plus fournies. Maintenant, nous pouvons voir de plus que dans ce discours la répétition est également importante. En deuxième position, viennent les corpus vulgarisés (0,36 en cardio, 0,25 en

pneumo). Enfin, l'utilisation et répétition, une fois normalisées, sont les plus basses dans les corpus expert (0,28 en cardio, 0,13 pneumo). Les documents expert sont certainement consacrés à des questions ciblées et précises et, par ailleurs, la grande taille des corpus expert semble pénaliser les valeurs normalisées.

La plus grande différence, quant aux fréquences et répétitions, se creuse entre les textes étudiant et expert, ensuite entre les textes étudiant et vulgarisés. Voilà quelques exemples de familles morphologiques qui montrent des fréquences élevées. La famille *diabète* est très fréquente en corpus diabète (expert et vulgarisé) ; la famille *hém* en corpus diabète étudiant ; la famille *pulmon* en corpus pneumologie. Le corpus cardiologie se distingue par trois bases *vascul* dans les documents expert, *cardio* dans les documents étudiant, et *graphie* dans les documents vulgarisés. En effet, ces familles morphologiques sont étroitement liées aux trois spécialités médicales : diabétologie (*diabète, hém*), cardiologie (*vascul, cardio, graphie*) et pneumologie (*pulmon*).

5 Conclusion et perspectives

Dans le travail présenté, l'objectif poursuivi concerne la distinction automatique des discours vulgarisé, expert et étudiant dans le domaine médical. Plus particulièrement, nous effectuons une étude détaillée des caractéristiques du niveau morphologique de ces documents. L'étude est faite de manière contrastive : nous analysons les corpus expert, étudiant et vulgarisés en parallèle et ensuite nous effectuons une comparaison de leurs spécificités. Les documents du corpus sont collectés au travers du portail CISMeF, leur annotation par les documentalistes de ce portail sert donc de « gold standard » pour cette étude. Ce travail mise sur l'utilisation des outils du traitement automatique de langues qui permettent d'accéder au niveau morphologique de la langue.

Nous avons abordé les objectifs de deux points de vue. Tout d'abord, nous avons étudié la productivité de 46 bases morphologiques. La productivité des bases correspond au nombre de lexèmes qu'elles forment. L'étude a été faite séparément dans chaque discours. Ensuite, nous avons observé les fréquences d'utilisation de ces bases en corpus, également séparément dans chaque discours.

Ces deux analyses montrent qu'il existe en effet des caractéristiques morphologiques spécifiques à chacun des trois discours : scientifique (expert), didactique (étudiant) et vulgarisé. Les résultats obtenus corroborent dans une certaine mesure. Ainsi, nous avons constaté que les textes didactiques destinés à l'enseignement de la médecine se caractérisent par la plus grande redondance de termes médicaux, ceci dans les trois spécialités étudiées (cardiologie, pneumologie, diabète). En effet, la répétition fait partie intégrante du processus de l'apprentissage. Si l'on ajoute à cela la taille de familles morphologiques plus élevée que dans d'autres discours, nous voyons que les textes étudiant se distinguent à la fois par la diversité (taille de familles morphologiques) et la densité (fréquences) du vocabulaire médical. De part de ces deux caractéristiques, les textes didactiques se distinguent largement des textes destinés au grand public ou aux professionnels. Concernant les discours vulgarisé et expert, notre étude montre que la redondance du vocabulaire est plus élevée en textes vulgarisés qu'en textes expert, tandis que sa diversité est plus grande dans les textes expert lorsque l'étude est faite avec des données brutes. Une fois normalisées, les données montrent que les textes experts sont proportionnellement moins redondants que les textes vulgarisés.

La perspective principale de notre travail concerne l'utilisation des critères morphologiques pour la distinction automatique de discours. En effet, les résultats obtenus

dans le présent travail suggèrent que la distinction automatique pourra fournir des résultats assez fiables. Dans cette étude, nous nous sommes concentrés sur 46 familles morphologiques, alors que l'ensemble de familles morphologiques pourrait être pris en compte. Toutefois, le succès de cette piste dépend de l'efficacité des outils du traitement automatique de langues et nous avons vu que l'analyse morphosémantique, bien qu'efficace pour de très nombreux lexèmes, montre encore quelques limites (*i.e.*, analyse erronée des lexèmes, manque de normalisation des bases, analyse incomplète). Au travers de la taille des familles morphologiques, nous avons abordé la notion de productivité des bases. Il serait intéressant d'explorer plus en détail la productivité morphologique de nos corpus en appliquant des approches proposées à cet effet [27]. Notons que la productivité des bases et affixes semble être en effet spécifique des discours et des genres [28]. Cette étude pourrait être étendue à des documents provenant d'autres sources (*i.e.*, Santé Canada) et à d'autres spécialités médicales.

Références

[1] Fox S. On line Health Search 2006. *Most Internet users start at a search engine when looking for health information online. Very few check the source and date of the information they find.* Technical report, Pew Internet & American Life Project, 2006; Washington DC

[2] www.chu-rouen.fr/cismef

[3] www.hon.ch

[4] AMA. Health literacy: report of the council on scientific affairs. Ad hoc committee on health literacy for the council on scientific affairs, American Medical Association. *JAMA*, 1999; 281(6) : 552–7.

[5] McCray A. Promoting health literacy. *Journal of American Medical Informatics Association*, 2005; 12: 152–163.

[6] www.google.com/coop

[7] Flesch R. A new readability yardstick. *Journal of Applied Psychology*, 1948; 23: 221–233.

[8] Gunning R. *The art of clear writing*. New York, NY : McGraw Hill, 1973.

[9] Björnsson & Härd af Segerstad, Segezstad B. Lixpä franska och tio andra språk. Stockholm:Pedagogiskt centrum, Stockholms skolförvaltning. 1979.

[10] *Medical Subject Headings*. National Library of Medicine, Bethesda, Maryland. http://www.nlm.nih.gov/mesh/meshhome.html.

[11] Kokkinakis D, Gronostaj MT. Comparing lay and professional language in cardiovascular disorders corpora. In A. Pham T., James Cook University, Ed., WSEAS *Tran sactions on BIOLOGY and BIOMEDICINE*, 2006; 429-437.

[12] Poprat M, Marko K, Hahn U. A language classifier that automatically divides medical documents for experts and health care consumers. In *MIE* 2006; 503–508.

[13] Zheng W, Millos E, Watters C. Filtering for medical news items using a machine learning approach. In *AMIA* 2002; 949–53

[14] Borst A, Gaudinat A, Boyer C, Grabar N. Lexically based distinction of readability levels of health documents. *MIE* 2008.

[15] Grabar N, Krivine S, Jaulent MC. Classification of health web pages as expert and non expert with a reduced set of cross-language features. In *AMIA* 2007; 284–8.

[16] Goeuriot L, Grabar N, Daille B. Caractérisation des discours scientifique et vulgarisé en français, japonais et russe. *TALN* 2007

[17] Miller T, Leroy G, Chatterjee S, Fan J, Thoms B. A classifier to evaluate language specificity of medical documents. In *HICSS*, 2007; 134-140.

[18] Wang Y. Automatic recognition of text difficulty from consumers health information. In IEEE, Ed., *Computer-Based Medical Systems* 2006; 131-136.

[19] Zeng-Treiler Q, Kim H, Goryachev S, Keselman A, Slaugther L, Smith CA. Text characteristics of clinical reports and their implications for the readability of personal health records. In *MEDINFO* 2007; pp. 1117–1121.

[20] Leroy G, Helmreich S, Cowie J, Miller T, Zheng W. Evaluating Online Health Information: Beyond Readability Formulas. In *AMIA* 2008; pp. 394-8.

[21] www.consumerhealthvocab.org

[22] Deléger L, Zweigenbaum P. Paraphrase acquisition from comparable medical corpora of specialized and lay texts. In *AMIA* 2008; 146-50.

[23] Zeng Q, Tse T. Exploring and developing consumer health vocabularies. *JAMIA* 2006; 13 : 24-29.

[24] Schmid H. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of International Conference on New Methods in Language Processing* 1994; 44-49.

[25] Namer F. Flemm: Un analyseur Flexionnel de Français à base de règles. *Traitement automatique des Langues* 2000; 41(2) : 523-47.

[26] Namer F. Automatiser l'analyse morphosémantique non affixale: le système DériF. Cahiers de Grammaire 2003; 28, 31–48.

[27] Bayen H. Quantitative aspects of morphological productivity. *Yearbook of Morphology* 1991; 109-149.

[28] Baayen H. Derivational productivity and text typology. *Journal of quantitative linguistics* 1994; 1(1) : 16-34.

Adresse de correspondance

Natalia Grabar

SPIM, UMRS Inserm 872 éq. 20

Centre de Recherche des Cordeliers

15, rue de l'École de Médecine

75006 Paris, France

E-Mail : natalia.grabar@spim.jussieu.fr

Informatique et Santé, 2009 (17) : 151-162
© Springer – Verlag France

Un service Web pour l'annotation sémantique de données biomédicales avec des ontologies

Clément Jonquet, Nigam Shah, Mark A. Musen

Center for Biomedical Informatics Research, Stanford University School of Medicine, USA

Abstract

The range of publicly available biomedical data is enormous and is expanding fast. This expansion means that researchers now face a hurdle to extracting the data they need from the large numbers of data that are available. Biomedical researchers have turned to ontologies and terminologies to structure and annotate their data with ontology concepts for better search and retrieval. However, this annotation process cannot be easily automated and often requires expert curators. Plus, there is a lack of easy-to-use systems that facilitate the use of ontologies for annotation. This paper presents the Open Biomedical Annotator (OBA), an ontology-based Web service that annotates public datasets with biomedical ontology concepts based on their textual metadata. The biomedical community can use the annotator service to tag datasets automatically with ontology and terminology terms (from UMLS and NCBO). We have used the annotator service internally to index several online datasets (e.g., ArrayExpress, PubMed, ClinicalTrial.gov). The index is directly queriable in the NCBO BioPortal ontology repository (www.bioontology.org). Such semantic annotations facilitate translational discoveries by integrating annotated data.

Keywords

Medical Informatics Applications; Information Storage and Retrieval; Controlled Vocabularies; Unified Medical Language System; Terminology; Semantics; Knowledge (*from MeSH*) + Ontology; Annotation; Semantic Annotation (*not in MeSH*).

1 Introduction

La richesse et la diversité des données biomédicales aujourd'hui accessibles au public permettent l'intégration translationnelle de multiples études bioinformatiques (*translational bioinformatics research*) [1-[2]. Cependant, les découvertes qui pourraient être réalisées par la fouille des données biomédicales sont limitées car la plupart des ressources publiques ne sont généralement pas décrites à l'aide de terminologies et d'ontologies. Un chercheur qui étudie les variations alléliques d'un gène voudrait connaître toutes les voies métaboliques qui sont touchées par ce gène, les médicaments dont les effets pourraient être modulés par les variations alléliques de ce gène, les maladies qui pourraient être causées par ce gène. Il peut également être intéressé par les essais cliniques qui ont étudié ces médicaments ou ces maladies. La connaissance nécessaire pour répondre à ces questions est aujourd'hui publique et disponible dans les ressources de données biomédicales en ligne, le problème est désormais de trouver cette information. La communauté biomédicale reconnaît d'ores et déjà l'importance des terminologies et des ontologies pour faciliter l'intégration de données et pour permettre de nouvelles découvertes [3]. Cependant, la

variété des données est très importante et celles-ci sont rarement annotées à l'aide de concepts décrits dans des ontologies biomédicales.[1] Le plus souvent, les éléments d'une ressource (e.g., recueil de données expérimentales, diagnostics, maladies, échantillons, descriptions d'essais cliniques, publications, images) sont annotés avec des métadonnées textuelles qui décrivent cet élément. Le problème est que ces descriptions textuelles sont rarement structurées et que le plus souvent elles n'utilisent pas des termes définis dans des ontologies biomédicales. Il existe donc un challenge qui consiste à produire pour ces descriptions textuelles des annotations (ou labels, tags) qui utilisent des termes d'ontologies et faciliteront la recherche et l'indexation de ces données ainsi que leur intégration [4, [5]. Par exemple, une recherche des essais cliniques pour le concept `carcinoma` doit considérer également les éléments indexés avec les termes `Malignant epithelial tumor` et `Epithelial Neoplasm` car l'ontologie SNOMED-CT nous précise que le premier terme est un synonyme du concept `carcinoma` et l'ontologie NCI Thesaurus nous précise que ce concept est un sous-type (`is_a`) de `Epithelial Neoplasm`. Il existe quelques exemples pour lesquels *l'annotation sémantique* (*i.e.*, l'annotation à l'aide de concepts définis dans des ontologies) des données s'est avérée très utile. Par exemple, Gene Ontology est largement utilisée pour décrire les fonctions moléculaires, les composants cellulaires et les processus biologiques des produits de gènes. Gene Ontology permet, entre autres, d'intégrer ces descriptions dans plusieurs bases de données, en gardant un formalisme commun. Autre exemple, quand une nouvelle citation PubMed est créée, le titre et le résumé de l'article correspondant sont indexés (grâce à des annotations manuelles) avec des termes de MeSH améliorant significativement la performance des recherches d'articles. Toutefois, en dehors de certains bons exemples, l'annotation sémantique de données biomédicales reste encore marginale. C'est pourquoi les succès mentionnés servent de motivation à notre travail.

L'annotation sémantique de données biomédicales n'est pas une pratique courante pour plusieurs raisons [6] :

- Les annotations ont le plus souvent besoin d'être créées manuellement par des experts ou directement par les auteurs des données.
- Le nombre d'ontologies biomédicales disponibles est important. En outre, ces ontologies changent régulièrement et se chevauchent les unes les autres. Elles sont dans des formats différents et ne sont pas toujours accessibles via des interfaces de programmation (API) qui permettent aux utilisateurs de les utiliser « programmatiquement ».
- Les utilisateurs ne connaissent pas toujours la structure des ontologies pour faire les annotations eux-mêmes. Parfois, ils ne connaissent pas l'existence des ontologies qu'ils pourraient utiliser.
- L'annotation est souvent une tâche supplémentaire ennuyeuse et sans retour immédiat pour l'utilisateur.

Dans le cadre du projet National Center for Biomedical Computing (NCBO), nous travaillons sur un système d'indexation de ressources biomédicales appelé *Open Biomedical Resources* (OBR) [1][2]. Dans cet article, nous présentons brièvement les résultats de ce travail. Nous présentons principalement un service Web d'annotation, appelé *Open Biomedical Annotator (OBA)*, qui permet aux chercheurs d'utiliser les ontologies biomédicales pour annoter leurs données automatiquement. L'annotateur traite

1 Dans le reste de l'article, nous utilisons seulement l'expression "ontology", pour terminologie et ontologie. Pour le travail présenté ici, les deux modèles peuvent être considérés équivalents puisque le niveau sémantique nécessaire (*i.e.*, nom, synonyme, alignement, relations is_a) existe dans les deux modèles.

les métadonnées textuelles brutes pour les taguer avec des concepts définis dans des ontologies biomédicales et utilise la connaissance représentée dans les ontologies pour étendre ces annotations. Les annotations sont ensuite renvoyées aux utilisateurs.

2 Matériel et méthodes

Le processus d'annotation du service OBA est composé de deux étapes principales (figure 1). Tout d'abord, le texte brut de l'utilisateur est traité par un outil de reconnaissance de concept qui détecte la présence d'un concept par reconnaissance syntaxique. Cet outil utilise un dictionnaire (ou lexique). Ce dictionnaire est une liste de termes qui identifient des concepts définis dans des ontologies.[2] Il est construit en récupérant à partir d'un ensemble d'ontologies biomédicales tous les noms ou synonymes qui identifient syntaxiquement les concepts. Le choix de l'ensemble d'ontologies à utiliser pour créer le dictionnaire dépend du type de données biomédicales que l'annotateur doit traiter. Par exemple, si un utilisateur veut annoter des expressions génétiques avec des noms de maladie, les ontologies SNOMED-CT et NCI Thesaurus peuvent être utilisées. Le résultat de cette première étape du processus est un ensemble d'*annotations directes*.

L'ensemble d'annotations directes est ensuite traité par un ou plusieurs composants d'expansion sémantique qui créent de nouvelles annotations à partir des annotations directes et en utilisant la connaissance représentée dans les ontologies. Par exemple :

- Le composant *fermeture transitive is_a* traverse la hiérarchie parent-enfant (ou type sous-type) des ontologies afin de créer de nouvelles annotations avec les concepts parents d'un concept constituant une annotation directe. Par exemple, si un texte est annoté directement avec un concept de NCI Thesaurus, `melanoma`, ce composant d'expansion sémantique peut générer une nouvelle annotation avec les concepts `skin tumor` et `neoplasms` car NCI Thesaurus spécifie les relations : `melanoma is_a skin tumor` et `skin tumor is_a neoplasms`. Le niveau maximum dans la hiérarchie à considérer pour l'expansion est paramétrable.[3]
- Le composant *distance sémantique* utilise la notion de similarité sémantique[4] [7-[9] afin de créer de nouvelles annotations avec les concepts similaires d'un concept constituant une annotation directe. Par exemple, si un texte est annoté directement avec un concept de MeSH, `melanoma`, ce composant d'expansion sémantique peut générer une nouvelle annotation avec les concepts `Apudoma` et `Neurilemmoma` car MesH spécifie ces trois concepts comme frères. La distance maximum à considérer pour l'expansion est paramétrable.
- Le composant *mapping* utilise les alignements (ou mappings) [10] entre ontologies afin de créer de nouvelles annotations avec les mappings d'un concept constituant une annotation directe. Par exemple, si un texte est annoté directement avec le concept `NCI/C0025202` (`melanoma` dans NCI Thesaurus), ce composant d'expansion sémantique peut générer une nouvelle annotation avec les concepts `SNOMEDCT/C0025202` (`melanoma` dans SNOMEDCT) et `38865/DOID:1909`

[2] Un concept est unique dans une ontologie (*i.e.*, classe). Un terme est une chaîne de caractères possible qui identifie un concept. Habituellement, un concept a plusieurs termes (*e.g.*, nom, synonymes, label).

[3] Ce composant d'expansion sémantique assume la transitivité de la relation is_a pour toutes les ontologies et terminologies utilisées. Cependant, si un utilisateur considère que cette transitivité n'est pas valide, le niveau maximal à considérer dans la hiérarchie peut alors être paramétrer à 1.

[4] Les distances sémantiques sont généralement basées soit sur la structure de l'ontologie, soit sur un corpus de données, ou les deux.

(`melanoma` dans Human Disease) parce que l'UMLS Metathesaurus et NCBO BioPortal fournissent ces alignements. Le type d'alignements à considérer pour l'expansion est paramétrable.

L'annotateur est conçu de façon à pouvoir sélectionner et paramétrer les composants d'expansion sémantique à la demande. Les performances du service sont liées aux paramètres choisis par l'utilisateur. Par exemple, la fermeture transitive is_a prend beaucoup de temps à traiter, même avec une hiérarchie pré-calculée. À la suite de la deuxième étape, les annotations directes ainsi que les *annotations expansées* sont classées par ordre de pertinence et renvoyées à l'utilisateur. Les annotations créées par l'annotateur ont pour sémantique : *ces données concernent (ou traitent de) ces concepts*. Cette sémantique peut être exploitée pour diverses applications comme la recherche ou l'intégration de donnée.

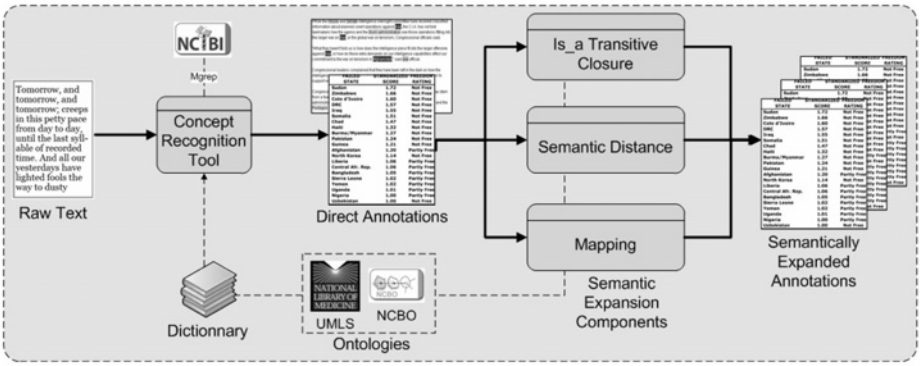


Figure 1 - Processus d'annotation du service OBA.

Les identifiants de concepts utilisés par l'annotateur proviennent soit de l'UMLS (*Unified Medical Language System*) Metathesaurus et sont formés par la concaténation des *ABbreviated Source name* (SAB) et des *Concept Unique Identifier* (CUI),[5] soit ils proviennent des *Uniform Resource Identifier* (URI) définis par le *National Center for Biomedical Ontology* (NCBO).

Chaque annotation possède un *contexte* qui précise si l'annotation est directe ou expansée et détaille l'origine de l'annotation. Par exemple, pour une annotation directe, le terme qui a été reconnu est précisé ainsi que sa position dans le texte. Dans le cas d'une annotation expansée, le concept à partir duquel l'annotation a été créée est précisé. Les annotations peuvent être renvoyées à l'utilisateur dans différents formats (texte, tab-delimited, XML ou OWL). La description des résultats renvoyés par l'annotateur est disponible [11].

3 Résultats et évaluation

3.1 Service Web d'annotation (OBA)

Notre implémentation de l'annotateur utilise toutes les terminologies dans UMLS (anglaises seulement, soit 94) ainsi que toutes les ontologies (complètement utilisables et fonctionnelles) de BioPortal (132 ontologies au moment de l'écriture). Ces ontologies nous donnent un dictionnaire contenant 2 608 228 concepts et 5 186 484 termes, ce qui en fait

[5] Voir la documentation de UMLS : http://www.nlm.nih.gov/research/umls/meta2.html

un des plus gros dictionnaires disponibles aujourd'hui.

Dans la première étape du processus d'annotation, le service utilise Mgrep [12], un outil de reconnaissance de concept développé par le *National Center for Integrative Biomedical Informatics* (NCIBI) à l'Université du Michigan qui rapporte un degré de précision élevé (> 95 %) pour la reconnaissance de noms de maladie [13]. Mgrep utilise une structure de radix-tree qui permet une reconnaissance rapide et efficace de texte à partir d'un dictionnaire. Nous avons réalisé une évaluation comparative de Mgrep avec MetaMap [14], l'outil de reconnaissance de concept qui sert de référence dans le domaine biomédical. Dans cette évaluation, nous avons utilisé quatre ressources de données et quatre dictionnaires différents pour évaluer les deux outils en termes de précision, rapidité d'exécution, et de passage à l'échelle. Le tableau 1 reporte certains de ces résultats pour deux des dictionaires. Mgrep s'est montré très rapide avec une précision plus élevée pour la plupart des ressources de données traitées et pour les quatre dictionnaires. Par exemple, Mgrep obtient 88 % de précision par rapport à 75 % pour MetaMap dans le traitement d'expressions génétiques de GEO. En outre, Mgrep est ouvert à tout type de dictionnaire et n'est pas limité, comme MetaMap, à UMLS. Nous avons donc pu utiliser Mgrep pour traiter les ontologies NCBO qui sont généralement disponibles dans les formats OBO (*Open Biomedical Ontology*) et OWL (*Web Ontology Language*). Pour plus de détail sur cette évaluation comparative, voir la référence [15].

Tableau 1 - Précision de Mgrep et MetaMap avec deux dictionnaires.

Ressource de données biomédicales	Disease dictionary (Concepts d'UMLS avec comme type sémantique « disease or syndrome »		Biological processes (Gene Ontology) dictionary	
	Mgrep	MetaMap	Mgrep	MetaMap
PubMed (subset) www.ncbi.nlm.nih.gov/PubMed	0,23	0,091	0,77	0,76
ClinicalTrials.gov http://clinicaltrials.gov	0,87	0,71	0,6	0,63
Gene Expression Omnibus www.ncbi.nlm.nih.gov/geo	0,88	0,755	0,93	0,73
ARRS GoldMiner (subset) http://goldminer.arrs.org	0,73	0,548	0,58	0,33

Dans la deuxième étape du processus d'annotation, le service fait actuellement appel à un composant fermeture transitive is_a et un composant mapping. Le service Web est déployé avec une interface SOAP (*Simple Object Access Protocol*) et une interface REST (*REpresentational State Transfer*) - http://obs.bioontology.org.

3.2 Index de ressources biomédicales (OBR)

Nous avons utilisé l'annotateur pour créer un système d'indexation de ressources biomédicales appelé *Open Biomedical Resources* (OBR) [1][2]. Nous avons traité : (1) les données d'expression génétiques de ArrayExpress et Gene Expression Omnibus (GEO) ; (2) les descriptions d'essais cliniques de Clinicaltrials.gov ; (3) des légendes et descriptions d'images de ARRS Goldminer ; et (4) des titres et résumés de citation PubMed. Le tableau 2 présente à la fois le nombre actuel d'éléments annotés et le nombre d'annotations créé pour chaque ressource (ces chiffres ont été produits avec une version plus ancienne du dictionnaire). L'index fournit des annotations (au moins 1) pour 99 % des éléments de PubMed et 100 % des éléments des autres ressources traitées. Le nombre moyen d'annotations est situé entre 359 et 824 par élément, avec une moyenne de 27%

d'annotations directes. L'index est interfacé par une API services Web (REST).

Tableau 2 - Ressources annotées dans l'index OBR.

Ressource de données biomédicales	Nombre d'éléments	Taille de la ressource (Mb)	Nombre d'annotations directes (Mgrep)	Nombre total d'annotations	Nombre moyen d'annotations/ élément
PubMed (subset) www.ncbi.nlm.nih.gov/PubMed	1 050 000	146,1	30 822 190	174840 027	160
ArrayExpress www.ebi.ac.uk/arrayexpress	3 371	3,6	502 122	1 849 224	525
ClinicalTrials.gov http://clinicaltrials.gov	50 303	99	16 108 580	48 796 501	824
Gene Expression Omnibus www.ncbi.nlm.nih.gov/geo	2 085	0,7	165 539	772 608	359
ARRS GoldMiner (subset) http://goldminer.arrs.org	1 155	0,5	134 229	662 687	564
TOTAL	1 106 914	249,9	47 732 660	226 921 047	(moy) 486,4

Cet index de ressources permet à un utilisateur de rechercher des données biomédicales annotées pour un concept donné. Il permet également, pour un identifiant d'un élément d'une ressource (e.g., PubMedID, GEO datasetID), d'obtenir toutes les annotations qui ont été créées pour cet élément. La création d'un tel index favorise l'intégration translationnelle de différentes ressources de données biomédicales.

3.3 Intégration dans BioPortal

NCBO [15] développe et maintient une application Web appelée BioPortal qui permet d'accéder et d'utiliser des ontologies biomédicales (http://bioportal.bioontology.org) [16-17]. Le portail contient une grande collection d'ontologies, telles que Gene Ontology, National Cancer Institute Thesaurus, International Classification of Diseases, Foundational Model of Anatomy. Les utilisateurs peuvent consulter, rechercher et commenter la collection d'ontologies soit en ligne, soit via une API de services Web (REST).

L'index de ressources OBR est directement accessible dans BioPortal. Lorsqu'un utilisateur visite un concept donné, il a accès (lien Web) à l'ensemble des éléments annotés avec ce concept comme l'illustre la figure 2.

4 Discussion

4.1 Scénarios supportés par l'annotateur

Il existe de nombreux cas d'utilisation pour le service d'annotation autre que la création de l'index de ressources biomédicales présenté précédemment. Le service est actuellement en cours d'évaluation dans plusieurs projets extérieurs qui illustrent certains de ces cas d'utilisation :

(1) Les chercheurs travaillant sur Trialbank (www.trialbank.org) à l'Université de Californie, San Francisco, créent des annotations d'essais cliniques sur le VIH/sida afin de développer une application Web pour visualiser et comparer les essais. Ils évaluent notre service d'annotation pour traiter les champs « health condition », 'intervention' et 'outcomes' d'essais cliniques de ClinicalTrials.gov.

(2) Les chercheurs de l'Université de l'Indiana évaluent l'intégration de l'annotateur dans leur système de gestion de recherche scientifique appelé Laboratree (http://laboratree.org), de sorte que toute annotation textuelle créée dans Laboratree

puisse être complétée d'annotations sémantiques.

(3) Les développeurs de Collabrx (http://collabrx.com) utilisent l'annotateur dans leur plate-forme Rex pour traiter le contenu généré par les utilisateurs, de façon à déterminer quels dictionnaires (*i.e.*, quelles ontologies) sont les plus appropriés pour leur contenu.

(4) Les chercheurs du Jackson Lab (www.jax.org) évaluent l'utilité de l'annotateur pour trier des articles à l'aide de concepts reconnus dans les titres et les résumés de ces articles.

Chacun de ces groupes bénéficie d'une meilleure interopérabilité de leurs données à l'aide des annotations créées avec l'annotateur.

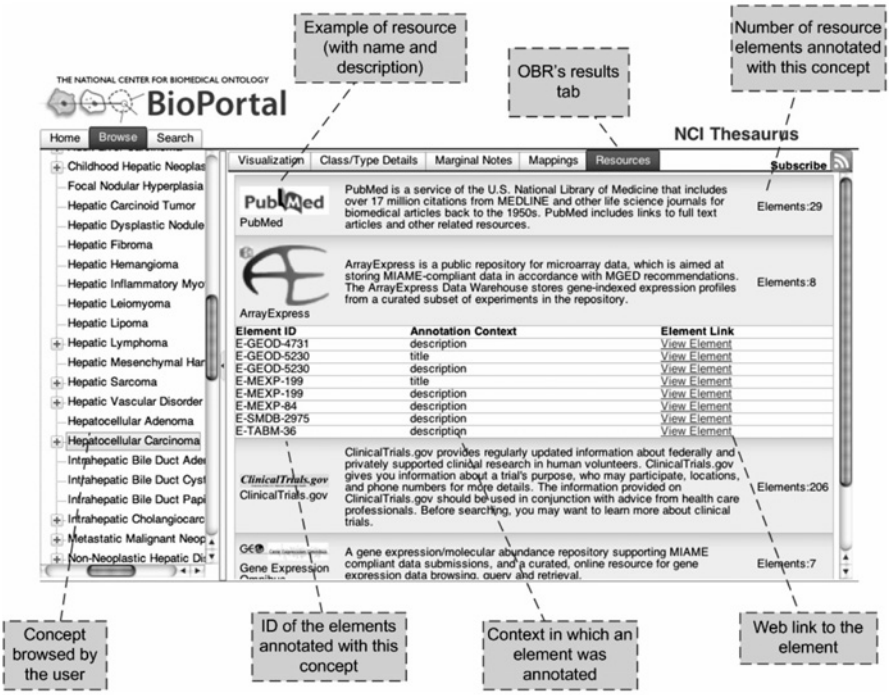


Figure 2 - Interface utilisateur dans BioPortal.

4 .2 État de l'art

Dans le domaine biomédical, l'annotation automatique et l'indexation de ressources est un sujet important. Il existe plusieurs outils de reconnaissance de concept qui permettent d'identifier des entités d'ontologies (concept ou relation) à partir de texte. Par exemple, IndexFinder [18], MetaMap [14], CONANN [19], SAPHIRE [20], Baud *et al.* [21] et Mgrep [12-13]. MetaMap, qui identifie des concepts de UMLS Metathesaurus, est généralement utilisé comme référence en termes d'évaluation. Notre choix pour Mgrep a été fait sur la base de critères tels que la flexibilité, la rapidité, l'ouverture et le passage à l'échelle comme décrit précédemment. CONANN est très similaire à l'annotateur que nous proposons et est également disponible en ligne en tant que service. CONNAN identifie les meilleures correspondances possibles entre un texte et un ensemble de concepts, tandis que

Mgrep identifie le plus grand nombre de concepts (de façon à générer le plus grand nombre d'annotations possible). Toutefois, CONNAN est limité à UMLS et n'effectue pas l'étape cruciale d'expansion sémantique. De manière assez générale, la connaissance représentée dans les ontologies est rarement utilisée pour l'expansion d'annotations, ce qui donne à notre service un avantage important.

Nous pouvons également noter que l'usage de telles pratiques d'expansion est plus fréquent dans le domaine de la recherche d'information. De nombreux projets utilisent les annotations sémantiques pour améliorer les performances de moteur de recherche. Par exemple, MedicoPort [22] utilise la sémantique de UMLS pour faire de l'expansion de requête. Moskovitch *et al.* [5] utilisent des annotations sémantiques (*concept-based search*) et démontrent l'importance du contexte (*context-sensitive search*) lors de l'annotation de documents structurés. Ils utilisent les relations is_a lors de leur étape d'expansion sémantique. De façon similaire, Essie [23] montre qu'une combinaison judicieuse de la structure des documents et de l'expansion sémantique de concept est une approche utile pour la recherche d'information. HealthCyberMap [24] utilise des ontologies et des distances sémantiques pour visualiser le contenu de ressources biomédicales. La plupart de ces outils sont limités à UMLS ou à un petit nombre d'autres ontologies. Cette limitation, donne à notre service un avantage important car il utilise, en plus de UMLS, toutes les ontologies (utilisables) de NCBO BioPortal.

Khelif *et al.* [25] présentent un travail similaire à l'index de ressources OBR. Les auteurs ont annoté la GeneRIF en utilisant la plate-forme GATE [26] qui leur permet d'extraire non seulement des concepts, mais des relations. Ils utilisent une application à base de graphes conceptuels appelée Corese [27] pour faire l'expansion sémantique. Névéol *et al.* [28] présentent également un travail sur l'indexation de ressources biomédicales avec des termes de MeSH.

4.3 Perspectives

Les travaux futurs se concentreront sur trois grands axes : (1) l'amélioration de l'étape de reconnaissance de concepts grâce à des techniques de traitement automatique des langues. En outre, il serait intéressant de reconnaître également des relations ; (2) l'amélioration de la personnalisation du service (paramètres utilisés pour les composants d'expansion sémantiques et sélection des ontologies à utiliser) ; (3) l'amélioration de l'étape d'expansion sémantique par la composition de composants existants et par le développement de nouveaux composants.[6] Par exemple, une distance sémantique peut être extraite de l'index OBR et utilisée par l'annotateur.

La plupart des ontologies disponibles étant en anglais, nous ne travaillons pour le moment qu'avec des ressources en anglais. Cependant, la méthodologie décrite précédemment est parfaitement valide pour d'autres langues.

5 Conclusion

L'annotation sémantique de données biomédicales avec des ontologies joue un rôle crucial pour l'interopérabilité et l'intégration des données ainsi que pour favoriser les découvertes translationnelles. Cette situation est vraie de manière générale dans le domaine des e-sciences. La nécessité de passer du Web actuel à un Web sémantique pourvu d'un contenu riche annoté à l'aide d'ontologies a clairement était identifié [29]. Répondre à ce

[6] Les composants utilisés actuellement (relations is_a et alignements) sont les plus faciles et évidents à utiliser.

besoin exige la mise à disposition de services (utilisables aussi bien par des humains et que par des agents logiciels) qui peuvent être facilement intégrés dans les processus actuels de curation et d'annotation de données.

Nous avons présenté un service pour l'annotation sémantique de données biomédicales. Notre annotateur se distingue des précédents efforts pour plusieurs raisons :

- Il s'agit d'un service Web qui peut facilement être intégré dans des processus et des programmes (vision architecture orientée service).
- Il utilise des ontologies à la fois pour créer des annotations directes et pour l'étape d'expansion sémantique.
- Il utilise une des plus grandes collections d'ontologies biomédicales publiques disponibles (UMLS Metathesaurus et NCBO BioPortal).[7]

L'annotateur est actuellement utilisé au sein du projet NCBO afin d'annoter un grand nombre de ressources biomédicales. Le service d'annotation et l'index créé à partir de celui-ci sont tous les deux accessibles via une API de services Web (http://obs.bioontology.org). L'annotateur est disponible et est, d'ores et déjà, utilisé par la communauté. Des retours d'expérience intéressants sont à prévoir sous peu et motiveront les évolutions et améliorations futures.

Remerciements

Ce travail est supporté par le programme : National Center for Biomedical Computing (NCBC) National Institute of Health roadmap initiative; NIH grant U54 HG004028. Nous remercions également Manhong Dai and Fan Meng de l'Université du Michigan (NCIBI) ainsi que Chuck Kahn pour l'accès à la ressource Goldminer.

Références

[1] Jonquet C, Musen MA, Shah. NH. A System for Ontology-Based Annotation of Biomedical Data. In: Bairoch A, Cohen-Boulakia S, and FroidevauxC. Editors, *International Workshop on Data Integration in the Life Sciences, DILS'08*, volume 5109 of *Lecture Notes in BioInformatics*, pages 144–152, Evry, France, June 2008. Springer-Verlag.

[2] Shah NH, Jonquet C, Chiang AP, Butte AJ, Chen R, Musen MA. Ontology-driven Indexing of Public Datasets for Translational Bioinformatics. *BMC Bioinformatics*, 2008. Expected end of 2008.

[3] Bodenreider O, Stevens R. Bio-ontologies: Current Trends and Future Directions. *Briefings in Bioinformatics* 2006; 7(3): 256–274.

[4] Nigam H. Shah, Daniel L. Rubin, Kaustubh S. Supekar, and Mark A. Musen. Ontology-based Annotation and Query of Tissue Microarray Data. In *American Medical Informatics Association Annual Symposium, AMIA'06*, pages 709–713, Washington DC., USA, November 2006.

[7] L'avantage d'un tel dictionnaire n'est bien sûr pas que quantitatif. Utiliser toutes ces ontologies dans la même application permet pour la première fois une large intégration dirigée par les données des ontologies de UMLS et BioPortal. Par exemple, nous travaillons sur l'extraction d'alignement entre concept de différentes ontologies dirigé par les annotations de l'index OBR. Si des milliers de données sont systèmatiquement annotés avec les mêmes concepts, il semble possible d'en extraire une relation entre ces concepts.

[5] Robert Moskovitch, Susana B. Martins, Eytan Behiri, Aviram Weiss, and Yuval Shahar. A Comparative Evaluation of Full-text, Concept-based, and Context-sensitive Search. *American Medical Informatics Association* 2007 ; 14(2): 164–174.

[6] Shah NH. *Encyclopedia of Database Systems*, chapter Biomedical Data/Content Acquisition, Curation, page In press. Springer-Verlag, New York, NY, USA, 2009.

[7] Caviedesa JE, Cimino JJ. Towards the development of a conceptual distance metric for the UMLS. *Biomedical Informatics* 2004 ; 37(2): 77–85.

[8] Al-Mubaid H, Nguyen HA. A Cluster-Based Approach for Semantic Similarity in the Biomedical Domain. In *28th IEEE EMBS Annual International Conference*, 2713–2717, New York, NY, USA, September 2006.

[9] Pedersen T, Pakhomov SVS, Patwardhan S, Chute CG. Measures of semantic similarity and relatedness in the biomedical domain. *Biomedical Informatics* 2007 ; 40(3): 288–299.

[10] Euzenat J, Shvaiko P. *Ontology matching*. Springer-Verlag, Berlin Heidelberg, DE, 2007.

[11] Jonquet C, Musen MA, Shah NH. Help will be provided for this task: Ontology-Based Annotator Web Service. Research report BMIR-2008-1317, Stanford University, CA, USA, May 2008.

[12] Dai M, Shah NH, Xuan W, Musen MA, Watson SJ, Athey BD, Meng F. An Efficient Solution for Mapping Free Text to Ontology Terms. In: *American Medical Informatics Association Symposium on Translational Bioinformatics, AMIASTB'08*, San Francisco, CA, USA, March 2008.

[13] Xuan W, Dai M, Mirel B, Athey B, Watson SJ, Meng F. Interactive Medline Search Engine Utilizing Biomedical Concepts and Data Integration. In: *BioLINK SIG: Linking Literature, Information and Knowledge for Biology*, 55–58, Vienna, Austria, July 2007.

[14] Aronson AR. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In *American Medical Informatics Association Annual Symposium, AMIA'01*, 17–21, Washington, DC, USA, November 2001.

[15] Bhatia N, Shah NH, Rubin DL, Chiang AP, Musen MA. Comparing Concept Recognizers for Ontology-Based Indexing: MGREP vs. MetaMap. Research report BMIR-2008-1332, Stanford University, CA, USA, March 2008.

[16] Rubin DL, Lewis SE, Mungall CJ, Misra S, Westerfield M, Ashburner M, Sim I, Hute CG, Solbrig H, Storey MA, Smith B, Day-Richter J, Noy NF, Musen MA. National Center for Biomedical Ontology: Advancing Biomedicine through Structured Organization of Scientific Knowledge. *OMICS A Journal of Integrative Biology*, 10(2): 185–198, June 2006.

[17] Musen MA, Shah NH, Noy NF, Dai B, Dorf M, Griffith NB, Buntrock J, Jonquet C, Montegut M, Rubin DL. BioPortal: Ontologies and Data Resources with the Click of a Mouse. In *American Medical Informatics Association Annual Symposium, Demonstrations, AMIA'08*, 1223–1224, Washington DC, USA, November 2008.

[18] Zou Q, Chu WW, Morioka C, Leazer GH, Kangarloo H. IndexFinder: A Method of Extracting Key Concepts from Clinical Texts for Indexing. In *American Medical Informatics Association Annual Symposium, AMIA'03*, 763–767, Washington DC, USA, November 2003.

[19] Reeve LH, Han H. CONANN: An Online Biomedical Concept Annotator. In: Cohen-Boulakia S, Tannen V. editors, *4th International Workshop Data Integration in the Life Sciences, DILS'07*, volume 4544 of *Lecture Notes in Computer Science*, pages 264–279, Philadelphia, PA, USA, June 2007. Springer-Verlag.

[20] Hersh WR, Greenes RA. SAPHIRE - an information retrieval system featuring concept matching, automatic indexing, probabilistic retrieval, and hierarchical relationships. *Computers and Biomedical Research* 1990 ; 23(5): 410–425.

[21] Baud R, Ruch P, Lovis C, Rassinoux AM. Recherche conceptuelle dans les textes médicaux. In : Fieschi M, Bouhaddou O, Beuscart R, Baud R, editors, *8e Journées Francophones d'informatique Médicale, JFIM'00*, volume 12 of *Informatique et Santé*, 205–216, Marseille, France, May 2000. Springer-Verlag.

[22] Can AB, Baykal N. MedicoPort: A medical search engine for all. *Computer Methods and Programs in Biomedicine* 2007 ; 86(1): 73–86.

[23] Ide NC, Loane RF, Demner-Fushman D. Essie: A Concept-based Search Engine for Structured Biomedical Text. *American Medical Informatics Association* 2007 ; 14(3): 253–263.

[24] Kamel-Boulos MN, Roudsari AV, Carson ER. HealthCyberMap: A Semantic Visual Browser of Medical Internet Resources Based on Clinical Codes and the Human Body Metaphor. *Health Information and Libraries*2002 ; 19(4): 189–200.

[25] Khelif K, Dieng-Kuntz R, Barbry P. An ontology-based approach to support text mining and information retrieval in the biological domain. *Universal Computer Science, Special Issue on Ontologies and their Applications* 2007; 13(12): 1881–1907.

[26] Cunningham H, Maynard D, Bontcheva K, Tablan V. GATE: A framework and graphical development environment for robust NLP tools and applications. In *40th Annual Meeting of the Association for Computational Linguistics, ACL'02*, Philadelphia, PA, USA, July 2002.

[27] Corby O, Dieng-Kuntz R, Faron-Zucker C, Gandon F. Searching the Semantic Web: Approximate Query Processing Based on Ontologies. *IEEE Intelligent Systems* 2006 ; 21(1): 20–27.

[28] Névéol A, Rogozan A, Darmoni S. Automatic indexing of online health resources for a French quality controlled gateway. *Information processing and Management* 2006 ; 42(3): 695–709.

[29] Handschuh S, Staab S. editors. *Annotation for the Semantic Web*, volume 96 of *Frontiers in Artificial Intelligence and Applications*. IOS Press, 2003.

Adresse de correspondance

Stanford Center for Biomedical Informatics Research,

Stanford University School of Medicine
Medical School Office Building, Room X-215
251 Campus Drive, Stanford, CA 94305-5479 USA

E-mail : {jonquet, musen, nigam}@stanford.edu

http://www.bioontology.org
http://obs.bioontology.org

Informatique et Santé, 2009 (17) : 163-170
© Springer – Verlag France

Les aspects organisationnels et logistiques de la télémédecine : expériences du projet CERTES au Mali

Abdel Kader Traoré[2], Cheick Oumar Bagayoko[1, 3], Mahamoudane Niang[1], Seydou Tidiane Traoré[1], Antoine Geissbühler[3],

[1]Centre d'Expertise et de Recherche en Télémédecine et E-santé (CERTES), Hôpital Mère Enfant le « Luxembourg », Bamako, Mali
[2]Centre National d'Appui à la lutte contre la Maladie, Bamako, Mali
[3] Service d'Informatique Médicale, Hôpitaux Universitaires de Genève, Suisse

Abstract

Les expériences de télémédecine en termes de formation médicale continue (FMC) et de téléexpertise prouvent qu'elle joue un rôle important dans l'amélioration de la qualité des soins surtout dans les pays vastes et en développement. Cependant la coordination organisationnelle et logistique de ces activités devient un enjeu majeur pour éviter la démotivation des professionnels de la santé s'agissant du délai de réponse à leurs demandes d'où la mise en place du Centre d'Expertise et de Recherche en Télémédecine et E-Santé (CERTES) sous la forme d'un projet de recherche sur deux ans.

La méthodologie utilisée est la valorisation du rôle des intermédiaires entre les professionnels demandeurs d'expertise et les experts qui semblent très surchargés pour répondre directement via les plates-formes de télémédecine.

Les résultats en cours démontrent que nos travaux pourraient permettre à terme de valider un modèle potentiel d'organisation logistique de la télémédecine. Ce modèle doit être validé par une évaluation plus complète à la fin du projet.

Keywords

Télémédecine ; FMC ; Télé-expertise ; CERTES.

1 Introduction

Il convient en premier lieu de définir la Télémédecine, qui ne fait pas aujourd'hui l'objet d'un consensus absolu. Cependant le point commun de toutes les définitions est qu'elle permet des pratiques médicales à distance grâce aux Technologies de l'Information et de la Communication (TIC). Ces pratiques sont multiples et variées, allant de la formation médicale continue, des téléconsultations, jusqu'aux applications complexes comme la télé-chirurgie. Les moyens technologiques utilisés sont aussi variés, avec des environnements asynchrones et à temps réel, tels que les plates-formes interactives directes et les visioconférences.

D'une manière générale, les applications de la télémédecine deviennent de plus en plus incontournables dans tous les pays du monde pour lutter contre la désertification médicale de certaines zones, en particulier rurales. La télémédecine constitue à ce titre un puissant vecteur d'équité d'accès aux soins et à la formation médicale continue [1-2].

Les pays du Sud particulièrement, de par leurs énormes géographies et leur nombre limité de personnel médical, éprouvent sans doute un besoin naturel de collaboration à distance dans le but d'équilibrer leurs systèmes sanitaires déficients. Le refus des professionnels de la santé d'aller servir à l'intérieur de ces pays, par peur d'isolement personnel et professionnel, reste une cause importante de ce déséquilibre.

Incontestablement, la télémédecine grâce aux nouvelles technologies peut jouer un rôle moteur dans ce cadre, en diminuant le sentiment d'isolement des professionnels de la santé et en leur offrant la possibilité de mise à jour de leurs connaissances par une formation médicale à distance [3].

Depuis les années 1960, le monde de la formation médicale a connu plusieurs applications informatiques allant des bases de données d'images, des outils de simulations, jusqu'aux systèmes experts mais dont la lourdeur de gestion et la convivialité n'ont pas été favorables à leur développement [4].

Devant la complexité et le coût des processus de soins, on assiste de plus en plus à une réelle dynamique de mise en réseau des professionnels de la santé, dans un but d'amélioration de la qualité des soins et d'optimisation des dépenses afférentes [5].

En France, on assiste à un développement très rapide des réseaux de téléformation dans les facultés et les hôpitaux. Un des exemples est celui du réseau des maternités du Nord-Pas-de-Calais avec des réunions multidisciplinaires hebdomadaires entre le CHU de Lille, référent, et plus de 20 maternités. Selon les données du ministère de la Santé, on constate une augmentation remarquable à la fois des acteurs en particulier des établissements de santé et des applications de télémédecine concernant le partage d'informations médicales dans le cadre de réseaux « ville- hôpital », la visioconférence et le partage des équipements dans les structures hospitalières [2]. En outre, selon le même rapport en 2004, on comptait 426 applications de télémédecine avec 1 117 établissements échangeant à distance des informations médicales multimédia. À cette panoplie, il faut ajouter la mise en place du Réseau de Télémédecine Régional Midi-Pyrénées en 1993 par le gouvernement français, après la création de l'institut européen de télémédecine à Toulouse [6].

En Amérique du Nord, plusieurs projets d'enseignement à distance depuis deux décennies ont été initiés dans le cadre de la formation médicale continue à distance [7- 9].

En 2001, l'Organisation mondiale de la santé (OMS) élabore une stratégie régionale de téléenseignement pour les pays du Pacifique occidental, dans le but de surmonter les contraintes géographiques et d'améliorer les performances du personnel de la santé dans les petites îles et les communautés rurales éloignées [10].

En Afrique Francophone depuis 2001 existe le Réseau de Télémédecine pour l'Afrique Francophone (RAFT) [11] qui relie actuellement 15 pays d'Afrique francophone et qui depuis quelque temps s'étend peu à peu aux pays anglophones et ceux arabophones [12].

Au Mali, depuis 2001, on assiste à une explosion des initiatives sous forme de projets pilotes dans le domaine de la télémédecine, principalement celui du Réseau Informatique Malien d'Information et de Communication Médicale appelé « Keneya Blown » (vestibule de la santé en langue vernaculaire bambara) [13].

Force est de reconnaître que si ces applications de télémédecine sont fortement appréciées par les professionnels de la santé, il manque aujourd'hui un modèle organisationnel et logistique adapté pour réduire les délais d'attente des demandes.

L'objectif de notre travail est donc d'expérimenter dans le cadre du projet CERTES un modèle organisationnel basé sur la notion d'intermédiaire actif qui fera le lien entre les demandeurs d'expertises et les experts via les plates-formes de télémédecine.

2 Matériel et méthodes

2.1 Matériels

2.1.1 Plate-forme de formation médicale continue à distance « Dudal » [14]

« Dudal » signifie en langue peuhl le foyer autour duquel les élèves coraniques se réunissaient autrefois pour réciter les leçons à leur maître coranique. Ce foyer dégage une lumière qui peut symboliser le savoir ou la science et il est alimenté par du bois cherché pendant la journée par tous les élèves coraniques, d'où également un sens de partage. Cette plate-forme développée dans le cadre du Réseau en Afrique Francophone la Télémédecine tire donc son nom de toutes ces significations.

Techniquement il s'agit d'un outil adapté aux faibles bandes passantes avec un débit minimum variant entre 16 et 30 kbit /s.

Il est complètement implémenté en langage Java ™, ce qui assure au système sa sécurité, sa transportabilité et l'indépendance de la plate-forme. L'accès au système se fait depuis une page Internet ou à travers un lien permettant de télécharger le logiciel et de l'exécuter sur la machine locale de l'utilisateur de manière automatique, grâce à la technologie Java Web start [15], qui permet en même temps la mise à jour permanente.

L'architecture de la plate-forme a été voulue double pour garantir une sécurité maximale vis-à-vis de la connexion Internet, dans le but d'assurer sa compatibilité avec les faibles bandes passantes des pays du Sud et garantir son fonctionnement instantanément par un relais en cas d'instabilité locale du réseau de communication au moment des cours. Il fonctionne du côté serveur, donc selon deux modes : le mode local où la machine locale sert de serveur à condition qu'elle possède une connexion avec une adresse IP publique pour permettre la diffusion web (figure 1) et le mode relais avec un serveur distant situé aux Hôpitaux Universitaires de Genève (figure 2).

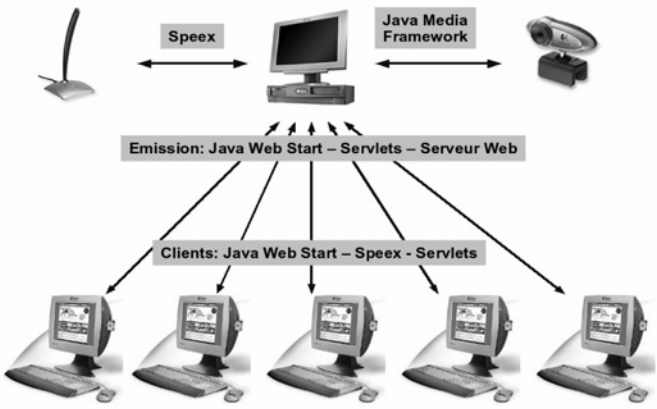


Figure 1 - Architecture « Dudal » : mode serveur local.

Figure 2 - Architecture « Dudal » : mode serveur relais.

Il faut aussi noter que la transportabilité physique de « Dudal » est très pratique par le nombre réduit de matériels informatiques nécessaires. Avec uniquement un ordinateur récent et une webcam, il est possible de produire et de suivre les cours partout avec un accès Internet bas débit (figure 3).

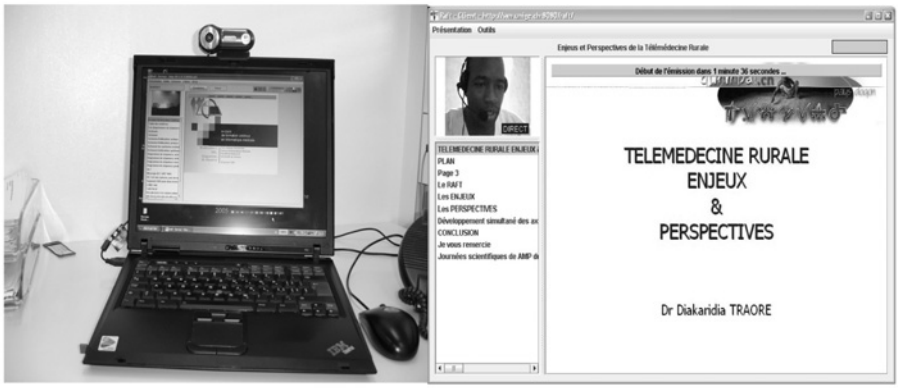


Figure 3 - Environnement de « Dudal », et son interface conférencier.

2.1.2 Plate-forme de téléconsultation asynchrone « iPath » [16]

C'est une plate-forme de télémédecine open source, notamment utilisée pour la téléconsultation asynchrone (figure 4). Cet environnement est surtout très important dans les domaines comme la neurochirurgie, où les cas sont décrits avec des images radiographiques ou de scanner au format « jpeg ». Une discussion est ainsi menée autour du cas en vue de poser soit le diagnostic, soit le meilleur protocole thérapeutique. La figure ci-après illustre un cas de téléconsultation asynchrone.

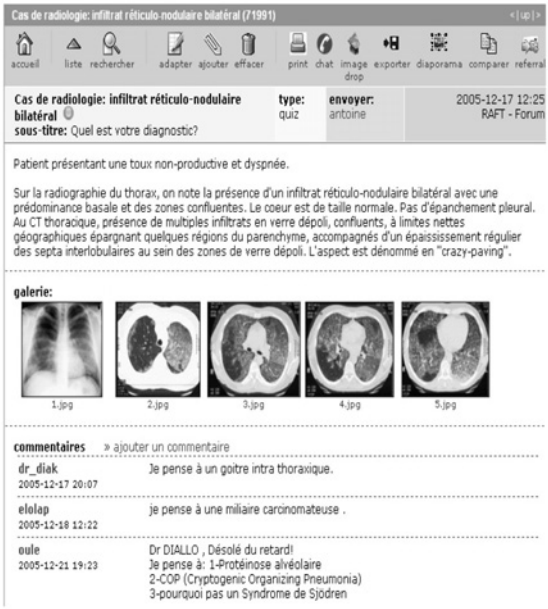


Figure 4 - Environnement « iPath » pour la téléconsultation.

2.2 Méthodes

L'organisation de notre travail est surtout basée sur l'équipe (« *dream team* ») de CERTES.

Il s'agit d'une équipe pluridisciplinaire constituée de médecins, d'informaticiens, des spécialistes en informatique médicale et des stagiaires internes en médecine.

La méthode est surtout basée sur la valorisation de la notion d'« intermédiaire actif ». Ce rôle est joué surtout par les stagiaires au sein de l'équipe qui font leurs thèses sur les différentes applications de la télémédecine. Ils jouent en effet l'interface entre les demandeurs des zones isolées dans le pays d'enseignement à distance et de second avis diagnostiques et/ou thérapeutiques et les experts le plus souvent concentrés dans les capitales et dont la charge du travail ne leur permet pas souvent de se connecter directement aux plates-formes utilisées.

Ces intermédiaires s'appuient sur les médecins opérateurs de la salle de contrôle qui les informent dès qu'une nouvelle demande arrive et les dirigent vers le médecin référent le plus approprié qu'ils soit au niveau local au Mali ou au niveau international. Le médecin référant sera ainsi contacté soit physiquement soit par d'autres moyens comme le téléphone en vue de garantir les délais d'attente.

3 Résultats

Les résultats en cours du projet sont très encourageants.

Concernant la production des contenus pour la formation médicale continue, du mois de

mai à septembre, quinze cours ont été produits par les experts de la capitale du Mali en direction des professionnels de la santé des zones les plus éloignées de la capitale et d'autres professionnels des pays d'Afrique francophone à travers le réseau RAFT.

Figure 5 - Auditeurs de FMC à distance dans un couloir du centre de santé de la commune IV de Bamako, Mali.

Il faut ajouter à cela une mobilisation sans précédent des professionnels de la santé, toutes catégories confondues, pour suivre les formations continues à distance de manière hebdomadaire (tous les jeudis matin) sur le réseau RAFT.

Dans le domaine de la téléconsultation asynchrone pour les avis médicaux secondaires diagnostiques et/ou thérapeutiques, de mai à septembre, 19 cas ont été traités avec satisfaction. Les demandes étaient adressées par les professionnels de la santé du Mali et d'autres pays d'Afrique francophone via la plate-forme « iPaht ». Les experts ayant répondu sont les médecins référents nationaux ou internationaux de CERTES. Ces cas concernaient surtout les domaines de la cardiologie, la neurochirurgie, la médecine interne, la dermatologie, la traumatologie, l'ORL, la chirurgie et la proctologie. Il est surtout important de noter que les délais d'attente étaient plus ou moins raisonnables variant de 23 minutes à 8 jours selon le degré d'urgence des cas.

Outre ces deux grandes applications de la télémédecine, le dynamisme des « intermédiaires actifs » a permis d'étendre les activités de télémédecine à d'autres centres de la capitale et à l'intérieur du Mali. Un exemple phare de ce dynamisme est l'accord de collaboration entre CERTES et la fédération nationale des centres de santé communautaires du Mali (FENASCOM). Ces centres constituent la médecine de première ligne donc, logiquement, les plus nécessiteux en termes de services de télémédecine.

4 Discussion - Conclusion

Si les résultats de nos travaux en cours semblent encourageants il existe tout de même quelques difficultés organisationnelles qui ont été fortement réduites par la mise en place

du mécanisme d' « intermédiaires actifs ». Il s'agit notamment de la non-disponibilité des experts pour produire des contenus de formation médicale continue et répondre aux demandes d'expertise à distance, et du retour d'information des praticiens demandeurs. Ce dernier aspect a fait que sur quatre-vingt-dix (90) demandes d'expertise, seulement 19 cas ont pu être traités avec satisfaction. En effet un cas traité avec satisfaction est celui dont la demande a été traitée par un expert avec un retour d'information du demandeur notifiant sa satisfaction. Il faut ajouter à cela le manque d'infrastructure qui fait que le local de réception des cours de formation continue est souvent réceptionné pour d'autres activités des structures de santé. La figure 5 illustre cette difficulté tout en démontrant l'enthousiasme et le besoin des professionnels de la santé qui suivent les activités jusque dans les couloirs des services.

Par ailleurs, aucune difficulté technique n'a été signalée, du fait que les plates-formes utilisées sont adaptées au contexte du terrain et surtout aux faibles bandes passantes.

En conclusion, cette étude a une nouvelle fois prouvé l'énorme potentialité qu'offrent les services de télémédecine pour pallier au problème de désertification médicale surtout en Afrique. Le temps n'est donc plus à la démonstration des avantages de la télémédecine mais plutôt à une réflexion profonde sur les enjeux liés au respect des délais d'attente des différents services. Dans ce sens, les expériences à terme de CERTES et son évaluation complète permettront de valider un modèle organisationnel et logistique de la télémédecine plus efficace.

Remerciements

Nous remercions l'Institut International de Communication et de Développement (IICD) et les Hôpitaux Universitaires de Genève pour leurs appuis financiers et techniques

Références

[1] Curran V R, Fleet L, Kirby F. Survey of Distance Learning Provision in Continuing Health Professional education in Canada. *Canadian Journal of university Continuing Education* 2003; 29 (2): 51-72.

[2] *Avis présenté par la Commission des affaires culturelles, familiales et sociales sur le projet de loi finances, Pour lutter contre les déserts médicaux : les promesses de la Télémédecine.* Poletti B, 2005; http://telemedecine.aphp.org/doc/senat/a1864-11.pdf

[3] Curran V R, Fleet L, Kirby F. Factor's influencing rural health care professional's access to continuing professional education. *Aust. J. Rural Health* 2006; 14 : 51-55.

[4] http://www.cidmef.u-bordeaux2.fr/wnantes/texte7.htm, dernière visite le 10 juin 2006.

[5] *Évaluation des réseaux de soins Bilan de l'existant et Cadre méthodologique, Mission d'évaluation des réseaux de soins, ANAES.* Le Boeuf D, 2001.

[6] http://www.telemedmidipyrenees.fr/, dernière visite le 25 novembre 2008

[7] Preston J, Brown F W. Using telemedicine to improve health care in distant areas. *Hospital community Psychiatry* 1992 ; 43 (1) : 25-32.

[8] Zollo S.A., Kienzle M.G., Henshaw Z, Crist L.G, Wakefield D.S. Tele-education in a Telemedicine Environment: Implications for rural Health Care and academic Medical Centers. *Journal of Medical Systems* 1999; 23 (2): 107 – 122.

[9] Kienzle BM, Curry D, Galvin J, Hoffman E, Holtum E, Shope L, Torner J, Wakefield D. Iowa's national laboratory for the study of rural telemedicine: a description of the work in progress. *Bull Med Libr Assoc* 1995; 83 (1).

[10] *Enseignement à distance. OMS Comité régional du Pacifique, 52e Session WPR/RC52/8*. Darussalam B, 10-14 septembre 2001.

[11] Bagayoko CO, Muller H, Geissbuhler A. Assessment of Internet-based Telemedicine in Africa (the RAFT project), 2006. Accept for publication for *Computerized Medical Imaging and Graphics*; 2004.

[12] http://raft.hcuge.ch/, dernière visite le 27 novembre 2008.

[13] www.keneya.net, dernière visite le 27 novembre 2008.

[14] www.dudal.net, dernière visite le 27 novembre 2008.

[15] http://java.sun.com/products/javawebstart/, dernière visite le 10 novembre 2008.

[16] http://telemed.ipath.ch/raft/, dernière visite le 20 novembre 2008.

Adresse de correspondance

Dr Cheick Oumar Bagayoko

Centre d'Expertise et de Recherche en Télémédecine et E-santé

Hôpital Mère Enfant le « Luxembourg », Bamako, Mali

E-mail : cob@hcuge.ch

Informatique et Santé, 2009 (17) : 171-182
© Springer – Verlag France

Évaluation du portfolio électronique et de son système d'accompagnement dédié à l'évaluation des pratiques professionnelles en technique et médecine transfusionnelles

Sophie Vessière[1], Stéphanie Jullien[1], Philippe Rouger[1], Pascal Staccini[2]

[1] *Institut National de la Transfusion Sanguine, Paris, France*
[2] *Laboratoire STIC Santé, UFR Médecine, Université de Nice Sophia Antipolis, Nice, France*

Abstract

In order to help physicians to meet the compulsory evaluation of professional practices as regards blood transfusion, the French National Blood Transfusion Institute and the University of Nice imagined a structured and accompanied eportfolio available on the Internet. This innovative system, faced up problem of the approval of an implement of new technology intended for a specific professional environment. This article offers to analyze conditions of approval of system and actions to implement so as to increase its performance and to optimize its use. The chosen methodology is that of questionary of contentment based on the evaluation of the performance of an information system offered by DeLone and McLean. The main results show that the eportfolio and its organizational system meet the physician's expectations. Indeed, the technological device and the provided information make the implementation of their EPP easier. Besides the notion of accompaniment by tutor has been shown as indispensable to the appropriation of this step.

Keywords

Electronic portfolio; Evaluation of professional practices; Blood transfusion; Information system; Evaluation.

1 Introduction

Assurer la qualité des soins et de la prise en charge est un objectif majeur. La diversité des soins et la complexité de leurs processus ainsi que la variabilité des pratiques constatées et la mise à jour continuelle des connaissances permettent de comprendre que l'atteinte de cet objectif doit passer par une démarche structurée d'évaluation et de formation. Il en est de même avec l'amélioration des technologies médicales et la nécessaire maîtrise des risques des pratiques. Connaissances et pratiques sont indissociables et contribuent conjointement aux compétences des praticiens. L'exercice professionnel quotidien amène le praticien à constater, dans sa pratique, de possibles carences de sa formation initiale, en particulier au regard de l'évolution rapide des connaissances, des techniques et des textes réglementaires. La mise en place de systèmes plus ou moins complexes et contraignants de mesure de la qualité des soins et de vérification des connaissances des professionnels est en cours, sous la double contrainte de l'amélioration de la qualité, liée à des influences et aux exigences sociétales (usagers, payeurs, assureurs) et de « tendance à périmer » des connaissances

académiques et techniques. Face à cela, la notion d'évaluation des pratiques professionnelles (EPP) apparaît. Elle est définie par la Haute Autorité de Santé (HAS) comme « l'analyse de la pratique professionnelle en référence à des recommandations et selon une méthode validée comportant la mise en œuvre et le suivi d'actions d'amélioration des pratiques. Ces pratiques concernent les activités diagnostiques, thérapeutiques ou préventives » [1]. Les modalités de satisfaction à l'obligation d'EPP sont définies par la HAS. Celle-ci a fixé la participation à au moins une action d'amélioration des pratiques pour une première période de 5 ans.

L'EPP s'inscrit dans la démarche globale de Formation Médicale Continue (FMC). À l'origine, la FMC est une obligation du code de déontologie médicale : « Tout médecin doit entretenir et perfectionner ses connaissances ; il doit prendre toutes les dispositions nécessaires pour participer à des actions de formation continue. Tout médecin participe à l'évaluation des pratiques professionnelles.» Cette démarche est inspirée du « *formative assessment* » des Anglo-Saxons. L'évaluation des pratiques se traduit par un double engagement des médecins : premièrement, ils s'engagent à fonder leur exercice clinique sur des recommandations et, deuxièmement, ils mesurent et analysent leurs pratiques en référence à celles-ci. Globalement, pour l'hospitalier, il s'agit essentiellement de formaliser ce que la plupart des médecins faisaient de longue date, tant du point de vue de la formation continue que de l'évaluation, par exemple lors des « staffs » ou de concertations pluridisciplinaires. Le but est de répondre le mieux possible à la mission hospitalière ; formation continue et EPP, individuelles et collectives, sont orientées vers cet objectif, elles ne peuvent s'envisager, dans leurs modalités pratiques, qu'intégrées dans le quotidien.

À ce titre, l'Institut National de la Transfusion Sanguine (INTS) est identifié comme organisme agréé EPP en technique et médecine transfusionnelles (TMT) et collabore avec le laboratoire STIC Santé de l'Université de Nice Sophia-Antipolis pour mettre en place le système qui permet aux praticiens de remplir leur obligation d'EPP en transfusion. Ce partenariat donne naissance au portfolio électronique structuré et tutoré : eEPP, qui agit pour la gestion et le suivi personnalisé des dossiers d'EPP en TMT. La conception de ce portfolio s'est inspirée des travaux nord-américains où le principe est plus développé qu'en Europe. D'après la Northwest Evaluation Association, un portfolio est « une collection finalisée et raisonnée de documents témoignant de la qualité et de la progression du travail d'un étudiant au travers de certaines de ses réalisations » [2]. Ce portfolio aide notamment les médecins à gérer leur EPP et les aide à assurer la traçabilité de leurs actions. Il se veut simple et ludique ; de plus, il est défini pour impacter au minimum l'activité professionnelle des praticiens et il doit être disponible sur l'ensemble du territoire afin de répondre à la demande en transfusion. À ce titre, la plate-forme est disponible grâce à une connexion Internet. Enfin, la démarche d'un praticien au sein de ce processus est accompagnée grâce à la présence d'un tuteur qui intervient pour un groupe de 10 praticiens au maximum, tant au niveau métier qu'au niveau humain.

Fin 2008, la plate-forme a clôturé sa phase pilote. Il est essentiel d'évaluer la satisfaction de ses utilisateurs, au niveau technique et organisationnel. L'évaluation du eportfolio sera la transition entre la phase de lancement et l'amorçage de son premier cycle de vie.

La mise en place d'un outil de mesure de la satisfaction a pour objectif de fournir les éléments d'un diagnostic précis des satisfactions et des insatisfactions des utilisateurs. Ce diagnostic se traduit logiquement par des actions concrètes pour maintenir ou accroître le niveau de satisfaction. Toute recherche de l'amélioration passe par le constat et l'analyse

de l'existant. L'enquête de satisfaction s'inscrit dans la logique d'amélioration continue de l'organisme selon les principes du Plan-Do-Check-Act, aussi nommé « Roue de Deming ». en place d'un système d'information (SI), il est à spécifier que de nombreuses études existent quant à la définition d'un modèle d'étude de ces systèmes et de l'évaluation de leur performance.

Ainsi le présent article propose et teste le modèle de William DeLone et Ephraim R. McLean pour évaluer le succès d'un portfolio électronique pour l'évaluation des pratiques professionnelles en technologie et médecine transfusionnelles auprès des praticiens utilisateurs et définit les perspectives d'amélioration du eportfolio pour être en adéquation avec les résultats de satisfaction obtenus.

2 Matériel et méthodes

2.1 Objectifs de l'enquête

La mesure de la satisfaction des usagers du dispositif « eportfolio pour l'EPP en TMT » vise à évaluer l'ensemble du système électronique et son environnement organisationnel, déployé depuis deux ans. Au travers de l'opinion des utilisateurs et de leur satisfaction sur les différents niveaux du processus, il a été jugé nécessaire d'identifier et d'apprécier les points à améliorer, dans le but de mettre en place le plan d'action adéquat à l'amélioration du eportfolio et de son système environnant, en adéquation avec les besoins exprimés.

2.2 Méthode de l'enquête

2.2.1 Choix du modèle d'évaluation

L'étude des principaux cadres et modèles conceptuels présentés dans la littérature permet d'identifier les principaux concepts proposés par les auteurs et les relations qui les unissent. Ainsi, les principaux concepts qui interviennent dans la relation entre les nouvelles technologies de l'information et de la communication (NTIC) et leur succès sont:

- les caractéristiques (techniques et économiques) des technologies de l'information;
- les caractéristiques des utilisateurs ;
- les caractéristiques de l'utilisation des technologies de l'information ;
- la satisfaction des utilisateurs ;
- les caractéristiques du processus relatifs aux technologies de l'information [3].

Parmi les modèles théoriques les plus répandus dans la littérature pour l'étude de l'adoption d'un SI innovant, deux modèles sont apparus comme incontournables : le modèle d'acceptation de la technologie (Technology Acceptance Model (TAM)) de Davis décrit en 1986 et le modèle DeLone et McLean établi en 1992. Ces modèles de recherche sont reconnus au sein de la communauté scientifique et largement commentés par les chercheurs en système d'information.

Le TAM étudie l'adoption en termes de facilité d'utilisation perçue et d'utilité perçue sur la base des intentions de comportement et les croyances relatives au système. Le TAM pose le principe que les perceptions qu'ont les utilisateurs de l'utilité et de la facilité d'utilisation agissent sur les attitudes d'utilisation des systèmes d'informations qui influencent à leur tour les intentions d'utilisation [4].

Une autre approche théorique considère que « l'utilisateur cherche d'abord à optimiser sa satisfaction et voit ainsi, dans l'évaluation subjective de la satisfaction durant l'usage, une cause déterminante du succès d'une technologie et donc de son utilisation ». Il s'agit du modèle décrit par DeLone et McLean. Au travers de ce modèle, deux nouvelles dimensions importantes sont nouvellement prises en compte pour l'acceptation des systèmes par rapport au Modèle TAM : l'acceptation des systèmes, à savoir la qualité du système et la qualité de l'information [5].

Ainsi, le modèle de DeLone et McLean (D & M) soutient que la réussite d'un système d'information est un phénomène dynamique et non un état statique [6]. Le modèle D & M identifie les paramètres qui contribuent à la réussite d'un système d'information. Ces facteurs sont au nombre de six :

- La qualité de l'information
- La qualité du système
- L'utilisation
- La satisfaction de l'utilisateur
- L'impact individuel
- L'impact organisationnel

L'ensemble de ces facteurs sont interdépendants temporellement et causalement [4], et permettent de mesurer la satisfaction des utilisateurs pour un système d'information mais aussi pour les bénéfices organisationnels adjacents. Ainsi la qualité du système et la qualité de l'information agissent conjointement sur l'utilisation mais aussi la satisfaction de l'usager pour un système. Parallèlement, l'utilisation du système agit directement sur la satisfaction individuelle de l'utilisateur et inversement. Aussi le nombre d'utilisation et la manière que l'utilisateur a d'interagir avec le SI, ont une incidence directe sur la satisfaction de ce dernier. La satisfaction du SI est donc interdépendante de la qualité du système d'information, de la qualité de l'information mais aussi de l'utilisation en elle-même. Par ailleurs ces interconnections décrites ont une incidence directe sur l'impact individuel du SI qui influencera l'impact organisationnel global (cf. figure 1) [7].

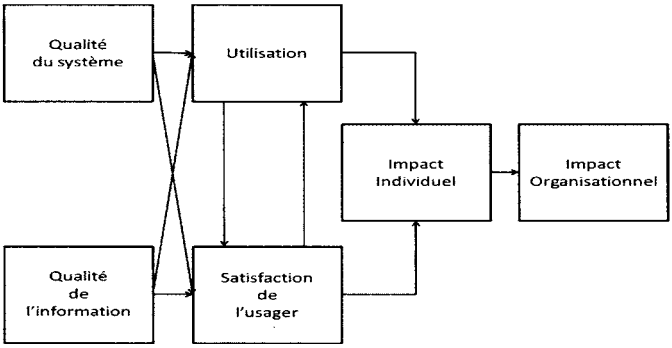


Figure 1 - Le modèle de DeLone et McLean pour la réussite d'un système d'information.

Le modèle D & M incorpore différentes thématiques d'études issues de l'instrument sémantique du modèle de Bailey et Pearson, mais aussi des études proposées par Chin en 1988 [8] et Davis en 1989. En effet, en 1983, 39 thématiques d'études de l'acceptation d'un système d'information ont émergé d'un examen de différentes études ciblées sur les

systèmes informatiques, les interfaces et les utilisateurs. Cet instrument sémantique et les études antérieures définissent les associations suivantes :

- Pour la qualité du système : les thématiques suivantes sont à étudier : la commodité d'accès, la flexibilité du système, l'intégration du système, le temps de réponse, le langage et l'apparition d'erreur.
- Pour la qualité de l'information : les thématiques suivantes sont proposées : la complétude, la précision, la fiabilité, la concision, le format et la pertinence [9].
- Pour l'impact individuel : il est à envisager que l'utilisation du système facilite l'accomplissement d'un travail, et la réalisation plus rapide d'une tâche.
- Pour la mesure de la satisfaction globale : elle ne s'étudie pas avec des thématiques mais directement avec des items qualificatifs (ex : système Inadéquat/Adéquat).
- Pour l'utilisation actuelle du système : elle se pondère avec l'utilisation journalière et la fréquence d'utilisation.

Pour chacune des thématiques, à l'exception de l'utilisation actuelle du système, quatre items d'évaluation sont associés. Ainsi le modèle D & M va évaluer une thématique avec, pour chaque niveau de description, deux qualificatifs opposés sémantiquement afin que l'utilisateur identifie sa satisfaction sur une échelle allant d'un terme à celui qui lui est totalement opposé dans l'interprétation (cf. figure 2).

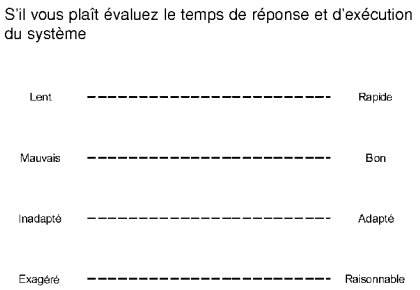


Figure 2 - Extrait du système de mesure proposé par Livari J [9].

2.2.2 Transposition du modèle de recherche D & M à l'évaluation du eportfolio pour l'EPP en TMT

Afin d'appliquer ce modèle d'évaluation global d'un SI lié aux NTIC à notre plate-forme électronique, nous sommes partis du postulat que, dans le modèle de D & M, la dimension « qualité du système » mesure l'efficacité technique du eportfolio ; dans un second temps, la dimension « qualité de l'information » apprécie son succès sémantique ; et enfin, les dimensions « d'utilisation, de satisfaction de l'utilisateur, et d'impact individuel », reflètent la mesure de la globalité du succès du portfolio électronique.

Nous nous sommes inspirés des thématiques et des items proposés par les auteurs pour chaque dimension. Chaque item se voit attribué une échelle allant de 0 à 5 (exemple : pour la qualité d'une information, le 0 correspond à plutôt « Mauvais » et 5 à plutôt « Bon »). En outre, nous nous sommes appuyés sur une analyse complète du processus et de la problématique de l'EPP pour adapter notre questionnaire à chaque dimension. Ainsi, l'évaluation de l'impact individuel s'est axée sur l'impact du dispositif global, ainsi que sur l'amélioration de leur pratique. Enfin les fréquences d'utilisation ont été proposées à partir d'une étude interne sur le temps moyen d'utilisation du portfolio chez les usagers.

2.2.3 Qualité du dispositif

La dimension de la qualité du dispositif se décompose en deux sous-dimensions d'étude : « le dispositif administratif et le dispositif technique ». Dans le second cas, nous répondons tout à fait au modèle D & M. Cependant, dans le processus, la partie administrative demeure importante car les praticiens interagissent au premier abord avec la plate-forme pour la gestion de leur dossier. Par ailleurs, cette première étape dispense aussi le premier contact avec la partie organisationnelle et humaine du dispositif, d'où l'intérêt de l'évaluer de manière à part entière dans la qualité du dispositif. La sous-dimension « dispositif technologique » s'évalue au travers des thématiques inspirées du modèle de Bailey et Pearson, à savoir : le temps de réponse du système, l'évaluation de l'interface utilisateur et les dysfonctionnements de la plate-forme.

2.2.4 Qualité de l'information

L'étude de la dimension de la « qualité de l'information » se définit grâce à quatre sous-dimensions directement liées au processus du eportfolio pour l'EPP en TMT:

- La qualité des échanges :
  - Elle vise à évaluer en première instance les réponses octroyées par le « secrétariat EPP ». En effet, ces personnes sont les interfaces humaines du dispositif pour les problématiques d'ordre logistique et bien entendu technique.
  - Dans un second temps, la qualité des échanges doit être évaluée pour la relation tuteur-praticien. En effet, la fonction de tuteur a été définie pour valider en première intention le processus d'EPP en TMT pour un praticien ; cependant, son rôle ne se limite pas à cette action. Il doit aussi orienter ses missions vers un rôle de pair, de conseil méthodologique et d'expertise métier mais surtout vers l'accompagnement et la motivation. Cette relation au même titre que celle établie par le « secrétariat EPP », intervient dans la rupture du processus d'apparition de sentiments d'éloignement et de délaissement que peut générer cette plate-forme.
  - Dans un troisième temps, la qualité des échanges est étudiée via les outils d'échanges mis en place sur la plate-forme électronique comme la boîte à outil, les rubriques d'informations ou encore les alertes mails automatiques réceptionnées.
- La qualité du référentiel de pratique : elle est évaluée grâce à l'appréciation des critères du référentiel et du système de cotation, ce document servant à l'une des étapes majeures dans le processus d'amélioration des pratiques. Son évaluation fait l'objet d'une thématique distincte.
- La qualité de la boîte à outils : ce dossier rassemble des documents méthodologiques pour mener à bien une démarche d'EPP mais aussi des documents orientés vers les métiers de la transfusion pour aider les praticiens à mettre en place leur action dans leur établissement respectif. L'intérêt de l'évaluation de cet entrepôt d'informations réside notamment dans son utilité et son adaptabilité à chaque métier.
- La qualité du manuel utilisateur : pour ce processus d'évaluation électronique à distance, il est mis en place un manuel utilisateur pour les praticiens. Ce manuel comporte à la fois des descriptifs techniques pour le eportfolio mais aussi pour des rappels sur les manipulations informatiques et les paramétrages à effectuer pour tirer parti au mieux des fonctionnalités de la plate-forme. Les praticiens étant autonomes et à distance dans la gestion de leurs interactions avec le eportfolio, le manuel demeure le support le plus adapté pour les aider au quotidien. Aussi son évaluation pourra être corrélée avec leur acceptation du système informatique.

2.2.5 Avis personnel

Cette notion regroupe à la fois la satisfaction personnelle pour le dispositif informatique et son système d'accompagnement mais aussi l'impact individuel de leur entrée dans le processus tant sur la facilitation de la réalisation de leur EPP via le système, que sur l'amélioration de la performance de leur EPP et l'amélioration de leur pratique au quotidien.

2.2.6 Fréquence d'utilisation

L'évaluation de la fréquence d'utilisation cible à la fois le nombre de connexions et aussi la durée moyenne de ces connections. Ces données seront à mettre en rapport avec la dimension de qualité du dispositif et de la satisfaction générale, et ainsi peut être définir s'il existe dans notre système une corrélation entre la fréquence d'utilisation et l'adaptation et le succès de la plate-forme.

2.2.7 Déploiement du questionnaire

Ce questionnaire a été diffusé à l'ensemble des praticiens effectuant leur démarche au sein du système EPP depuis une période supérieure à six mois. En effet, l'ensemble des groupes de praticiens débute leur EPP via le eportfolio lorsqu'au moins 8 personnes sont inscrites dans le même programme et lorsqu'un tuteur peut leur être attribué. Le questionnaire est distribué anonymement ; cependant, l'étude permet d'identifier la période depuis laquelle ils demeurent dans notre processus. Ceci est indispensable, notamment pour établir des corrélations entre le temps dans le processus, le temps d'adaptation et de manipulation du système et la satisfaction pour le dispositif technique et la structure organisationnelle attenant. Ainsi le déploiement du questionnaire s'est effectué auprès de cinq groupes de praticiens ayant adhéré à notre processus pour réaliser leur EPP dans le programme orienté « hémovigilance et sécurité transfusionnelle dans les établissements de santé ».

3 Résultats

L'étude diffusée pendant quelques mois a obtenu 62 % de retour au mois de juillet 2008. Les résultats sont représentés en pourcentage pour chacun des items (tableau 1). Les nombres présentés dans les colonnes de gauche représentent les pourcentages de non-réponse par item.

Dans la dimension «qualité du dispositif », le dispositif administratif semble être apprécié par les utilisateurs. Concernant le dispositif technologique, les résultats sont globalement moyens. Les utilisateurs jugent les items caractérisant le temps de réponse de manière mitigée. Les résultats pour cette sous-dimension seront donc à étudier de manière plus détaillée. Les résultats oscillent entre « 3 et 4 » pour l'interface utilisateur alors que les dysfonctionnements sont évalués quasi inexistants, anodins et insignifiants.

Pour la dimension « qualité de l'information », les résultats sont très satisfaisants. Cependant, il est à noter que les relations avec le tuteur sont considérées comme plutôt rares. Par ailleurs, les items « utilité et efficacité » des outils d'échanges demeurent majoritairement au niveau « 3 ». Les pourcentages, pour les items qualifiant la boîte à outils, sont essentiellement au niveau « 4 » à l'exception de la pertinence. Enfin, dans cette dimension, le manuel utilisateur obtient des résultats mitigés.

Pour la dimension « Avis personnel », les résultats sont de manière générale satisfaisants. Pour 53 % des praticiens questionnés, le système est « stimulant » ; 51 % optent pour une certaine « rigidité ». Concernant « l'impact individuel», 61 % sont « entièrement

d'accord » pour dire que l'utilisation du dispositif global facilite la réalisation de leur EPP et 50 % pensent que le dispositif améliore la performance de leur EPP. Cependant, uniquement 25 % jugent le dispositif comme améliorant leur pratique au quotidien.

Enfin, dans la dimension « Fréquence d'utilisation », il est à constater que 67 % des praticiens se connectent à la plate-forme une fois ou moins par mois et que le temps moyen d'utilisation par connexion ne dépasse pas 30 minutes pour 68 % des utilisateurs interrogés.

L'ultime élément d'étude du questionnaire repose dans la partie « Commentaires ». À propos du dispositif dans sa globalité, il est apparu qu'il s'agit d'une formule innovante et intéressante avec la garantie de l'agrément de l'INTS. L'efficacité et l'impact du système demeurent non perçues pour certains des praticiens car l'avancement dans le processus n'est pas suffisant pour pouvoir correctement en juger. Concernant l'outil eportfolio, il est apparu l'existence d'une sensation de ne pas toujours savoir « où le système mène le praticien », en contre-partie, l'utilisation de la plate-forme reste pour certains trop irrégulière pour l'évaluer. L'informatique ne remplace pas les partages d'expériences et les échanges, constructifs et stimulants, qui sont au demeurant estimés comme très enrichissants lors des sessions de rencontres organisées. Les problèmes techniques ou logistiques sont quasi inexistants. Enfin, le questionnaire doit être rediffusé, après quelques mois d'utilisation et des items qualifiant les thématiques sont à reformuler et à expliciter.

Lorsque l'on étudie les liens entre les différentes dimensions, il apparaît tout d'abord que les praticiens évaluant plutôt négativement le système informatique sont aussi ceux qui utilisent peu la plate-forme ou qui l'utilisent peu de fois mais avec des connexions très longues. Ainsi, il est nécessaire de comprendre grâce à d'autres études s'il s'avère que l'utilisation impacte directement sur l'évaluation du système ou si le système est jugé peu fonctionnel et interfère sur l'utilisation globale. De plus, il est démontré que la qualité du système agit dans notre dispositif davantage sur la satisfaction des participants que sur l'utilisation. Par ailleurs, le temps de connexion diminue lorsque le nombre de connexion augmente ; ceci doit être identifié car l'étude démontre par ailleurs que la satisfaction globale et l'impact individuel augmente positivement auprès des utilisateurs ayant un nombre de connexion plus élevé. En outre, il est à constater que le temps de connexion est d'environ 30 minutes pour une satisfaction intéressante du système. Aussi l'étude des cas de connexion longue permettra de savoir si le temps passé sur la plate-forme est lié à un attrait pour la technologie ou si, au contraire, il y a une non-appropriation du système. Ceci pourra permettre d'identifier des liens plus ciblés avec l'accompagnement ou la qualité du système. Enfin il existe un lien étroit entre l'utilisation et la satisfaction de l'utilisateur, par ailleurs ces deux paramètres influencent directement l'impact individuel de la plate-forme.

4 Discussion

L'application du modèle de DeLone et McLean à l'évaluation du dispositif a légitimé la classification des étapes du processus mis en exergue lors de l'analyse effectuée durant la phase pilote.

Tableau 1 - Résultats de l'enquête de satisfaction sur le dispositif électronique et organisationnel d'EPP en TMT, exprimés en pourcentage de réponses par items (juillet 2008).

I. QUALITE DU DISPOSITIF

A. Dispositif administratif

1. Comment évaluez-vous la gestion administrative de votre dossier ?

	0	1	2	3	4	5	
Lente	0	11	25	28	25	11	Rapide
Mauvaise	0	0	3	29	36	28	Bonne / 4
Inadaptée	0	0	0	24	43	29	Adaptée / 4
Inefficace	0	0	4	29	43	20	Efficace / 4

2. Comment évaluez-vous les modalités de contact avec le secrétariat ?

Inadaptées	0	0	0	21	36	43	Adaptées
Désagréables	0	0	0	14	43	39	Agréables / 4
Inefficaces	0	0	4	25	32	35	Efficaces / 4
Incommodes	3	0	4	18	36	35	Pratiques / 4

B. Dispositif technologique *(plate-forme de gestion à distance eEPP)*

1. Comment évaluez-vous le temps de réponse du système ?

Lent	4	11	14	25	29	15	Rapide / 4
Mauvais	0	0	7	39	29	25	Bon
Inadapté	0	0	4	43	32	14	Adapté / 7
Exagéré	0	0	7	39	32	18	Raisonnable / 4

2. Comment évaluez-vous l'interface utilisateur du système ?

Complexe	4	7	11	29	29	9	Simple / 4
Incompréhensible	0	7	18	43	24	4	Intuitive / 4
Difficile	0	4	11	36	38	7	Facile
Mal structurée	0	4	7	57	21	4	Bien organisée / 7

3. Comment jugez-vous les dysfonctionnements de la plate-forme EPP ?

Récurrents	0	7	11	14	43	21	Inexistants / 4
Différents	4	11	7	18	31	4	Identiques / 25
Sérieux	7	7	7	18	40	7	Anodins / 14
Handicapants	4	4	11	32	32	10	Insignifiants / 7

III. AVIS PERSONNEL

A. Satisfaction personnelle

1. Donnez votre sentiment global quant au dispositif EPP et à son système d'accompagnement par Internet.

	0	1	2	3	4	5	
Insatisfaisants	0	4	10	11	50	25	Satisfaisants
Inadéquats	0	0	18	11	46	20	Adéquats / 4
Décourageants	0	0	14	29	42	11	Stimulants / 4
Rigides	0	4	18	29	38	7	Flexibles / 4

B. Impact individuel *(Représenté en % par catégories)*

1. L'utilisation du dispositif dans sa globalité facilite la réalisation de mon EPP

	0	1	2	3	4	5	
Pas d'accord	0	0	18	21	29	32	Entièrement d'accord

2. L'utilisation du dispositif dans sa globalité améliore la performance de mon EPP.

Pas d'accord	0	0	21	25	36	14	Entièrement d'accord / 4

3. Le dispositif dans sa globalité améliore ma pratique au quotidien

Pas d'accord	0	11	25	39	18	7	Entièrement d'accord

IV. FREQUENCES D'UTILISATION

A. Fréquence d'utilisation

1. Quelle est votre fréquence d'utilisation de la plate-forme eEPP ?

Moins d'une fois par mois	21	Une fois par semaine	4
Une fois par mois	46	Plusieurs fois par semaine	0
Plusieurs fois par mois	29	Une fois par jour	0

B. Utilisation en connexion

1. Quelle est votre durée moyenne d'utilisation à chaque connexion ?

Moins de 10 minutes	0	Entre 1 heure et 2 heures	7
Moins d'une demi-heure	68	Entre 2 heures et 3 heures	0
Entre une demi-heure et 1 heure	25	Plus de 3 heures	0

II. QUALITE DE L'INFORMATION

A. Echanges

1. Comment évaluez-vous les réponses fournies par le secrétariat EPP, à vos questions / difficultés d'ordre technique?

	0	1	2	3	4	5	
Inappropriées	0	0	4	11	29	38	Appropriées / 18
Lentes	0	4	7	25	14	32	Rapides / 18
Vagues	0	0	7	18	25	32	Précises / 18
Inefficaces	0	0	4	25	18	39	Efficaces / 14

2. Comment évaluez-vous les réponses fournies par le secrétariat EPP à vos questions / difficultés d'ordre logistique ?

Rares	0	0	0	25	25	21	Fréquentes / 29
Inadaptées	0	0	0	11	29	39	Adaptées / 21
Indisponibles	0	0	0	11	32	36	Disponibles / 21
Insatisfaisantes	0	0	0	11	29	42	Satisfaisantes/18

3. Comment évaluez-vous les relations avec votre tuteur ?

Désagréables	0	0	0	7	21	68	Agréables / 4
Insatisfaisantes	0	4	0	4	39	53	Satisfaisantes
Rares	0	18	18	29	14	14	Fréquentes / 7
Inefficaces	0	4	4	13	36	36	Efficaces / 7

4. Comment évaluez-vous les outils d'échanges mis à disposition via la plate-forme (boîtes de commentaires, rubriques d'informations, alertes mails...) ?

Inadaptés	0	0	4	21	54	14	Efficaces / 7
Inutiles	0	0	4	39	29	21	Utiles / 7
Complexes	0	7	14	14	46	14	Simples / 4
Inefficaces	0	0	4	50	25	14	Efficaces / 7

B. Référentiel

1. Comment évaluez-vous les critères du référentiel de bonnes pratiques?

Peu pertinents	0	11	4	14	50	17	Très pertinent / 4
Incomplets	0	7	14	11	46	21	Complets
Inadéquats	0	7	11	14	46	18	Adéquats / 4
Imprécis	0	4	7	39	39	7	Précis / 4

2. Comment évaluez-vous le système de cotation?

Inadapté	4	7	14	39	28	4	Adapté / 4)
Imprécis	4	11	21	32	18	7	Précis / 7
Incomplet	4	7	21	25	29	7	Complet / 7
Inadéquat	4	4	21	39	18	7	Adéquat / 7

C. Boîte à outils

1. Comment évaluez-vous la boîte à outils à disposition sur la plate-forme ?

Inutile	0	0	0	25	50	25	Utile
Peu pertinente	0	0	7	50	28	11	Très pertinente/4
Inadaptée	0	0	7	32	46	11	Adaptée / 4
Obsolète	0	0	4	29	46	17	Actualisée / 4

D. Manuel de l'utilisateur

1. Comment évaluez-vous le manuel utilisateur ?

Inutile	0	0	11	18	50	21	Utile
Long	7	11	21	32	21	4	Court / 4
Austère	4	4	11	36	28	13	Convivial / 4
Inadapté	4	0	4	32	38	18	Adapté / 4

V. COMMENTAIRES

Merci de nous faire part de toutes remarques ou demandes que vous jugez utiles à l'amélioration de cette démarche.

Nous avons mis en place un système d'évaluation reproductible œuvrant pour l'amélioration de l'acceptation de notre système par ses utilisateurs. En effet le questionnaire analyse l'ensemble des processus, mais aussi les documents accompagnant le praticien et implémentant son portfolio individuel, et l'ensemble des interfaces indispensables au bon déroulement d'une EPP dans le dispositif, à savoir le tuteur qui a un rôle d'accompagnement et de pair et le secrétariat EPP qui demeure l'un des processus supports du praticien. Les résultats cette étude sur le eportfolio et de son dispositif organisationnel environnant démontrent des axes d'améliorations à définir et à prioriser.

En ce qui concerne les points à améliorer sur la rédaction des référentiels, ces derniers ont déjà été mentionnés avant le déploiement de l'enquête de satisfaction. Aussi des experts ont réalisé une version 2 du référentiel dernièrement. Il n'a été proposé qu'à un seul groupe qui n'a pas bénéficié de ce questionnaire de satisfaction, et sera à évaluer prochainement. En outre, certains points à modifier vont engendrer des développements informatiques qu'il faudra définir avec l'équipe informatique afin de les hiérarchiser dans le temps.

Par ailleurs, il est essentiel de comprendre que le principal atout de ce dispositif est paradoxalement l'un de ses points critiques : le tutorat. En effet, l'intervention de ces personnes octroie au processus sa crédibilité auprès des instances et du praticien réalisant son EPP. Ce facteur est « humain » dépendant, aussi a-t-il été testé un questionnaire d'évaluation du comportement d'une personne face à la gestion de groupe. Ce questionnaire a été en adéquation totale avec les points à améliorer pour certains tuteurs qui sont apparus lors d'entretiens avec les praticiens en démarche. Ce questionnaire pourrait ainsi être « la clé de voûte » du recrutement des tuteurs pour les prochains groupes.

En parallèle de cet aspect de gestion de groupe, il est apparu une lacune quant aux outils mis à disposition auprès des tuteurs pour le suivi d'un groupe. Des travaux sont en cours, avec des tuteurs volontaires, pour établir un ensemble d'outils de suivi. L'ensemble des résultats de satisfaction devront être, en outre, comparés ultérieurement avec les fréquences d'utilisation, la date d'insertion dans le système et les relations avec leur tuteur pour comprendre s'il s'agit davantage de problématiques liées à l'informatique (système trop rigide ou connaissance des NTIC inadaptée pour participer à une telle démarche autonome) ou de l'impression d'éloignement et d'isolement dus au système, qui impactent sur les résultats globaux de satisfaction. Des études plus approfondies sur les relations d'un praticien à l'informatique et aux NTIC, et des études ciblées sur le tutorat pourront nous aider à comprendre ce phénomène et ainsi apporter les mesures adéquates pour améliorer les relations avec ce système.

En effet, si les utilisateurs sont satisfaits des NTIC, ils utilisent pleinement ces technologies et vont donc être encore plus performants dans ce domaine. En outre, plusieurs chercheurs ont démontré que le succès des NTIC ne dépend pas seulement des gains d'efficacité et d'efficience et donc de l'augmentation attendue de la performance, mais plutôt de l'acceptation de la nouvelle technologie [10]. Par ailleurs, si les utilisateurs des NTIC ne parviennent pas à les admettre ou n'exploitent pas l'ensemble de ses capacités, il y aura pas ou peu d'impact sur la performance [11].

Dans cette logique, des simplifications des interfaces informatiques sont en cours d'élaboration. Le primo objectif est de regrouper au maximum les fonctionnalités sur

quelques onglets de navigation afin d'éviter la multiplication des possibilités de navigation pour des internautes qui ne sont pas obligatoirement familiers avec ce type de système. Afin d'optimiser l'acceptation de la plate-forme, il est nécessaire d'augmenter les temps et les fréquences de connexions des utilisateurs. Nous proposons de mettre en place des outils ludiques d'échanges en ligne tels que des forums ou une messagerie instantanée (clavardage).

L'adoption du système passe par des phases d'apprentissage qui pourraient être déployées via des outils utiles et adaptés pour le quotidien du praticien qui a une charge de travail non négligeable. Ceci va donc devoir intégrer le déploiement et la restructuration de la boîte à outils qui permettra aux praticiens d'interagir avec la plate-forme pour rechercher des éléments indispensables à l'amélioration de leur pratique au quotidien.

Enfin, même si la création de cette plate-forme d'évaluation à distance est le premier avantage pour l'accessibilité de ce système agréé, à n'importe quel praticien en France et dans les DOM-TOM, il est important de mentionner que l'essentiel du travail du praticien va se faire de façon autonome et solitaire. En effet, « l'instrumentation informatique ne vient qu'en support de ce dispositif pour améliorer la synchronisation des étapes, assurer leur traçabilité et faciliter la communication et les échanges entre les acteurs » [12]. Aussi, le tutorat et la présence du secrétariat EPP sont les seuls liens humains qui relient le praticien à la mise en place de l'EPP dans leur établissement.

Aussi, si nous avons pu constater que des actions sont déployées pour améliorer la relation « tutorale », les échanges avec le secrétariat EPP ne sont pas à négliger. En ce sens un suivi davantage programmé se met en place peu à peu. La possibilité d'échanges téléphoniques et/ou en face à face, lorsque cela est réalisable, est étudiée pour chaque praticien, pour chaque groupe mais aussi auprès des tuteurs, qui, demeurant le maillon essentiel de ce processus, ne doivent en aucun cas être délaissés au point de vue relationnel.

5 Conclusion

L'ensemble du dispositif mis en place conjointement par l'INTS et l'Université de Nice Sophia Antipolis a débuté sa phase de lancement il y a environ un an et demi. Ce portfolio imaginé en 2005, demeure un projet novateur dans le secteur médical. En effet, si des systèmes similaires se sont peu à peu mis en place dans le secteur médical et plus précisément pour la formation continue, seul ce projet vise à aider les praticiens à remplir leur obligation d'EPP en TMT via un organisme agréé, l'INTS. Cet agrément a de ce fait approuvé le système du portfolio électronique comme outil de validation et de traçabilité de cette évaluation. L'enquête et les résultats présentés ont clairement démontré que le système doit encore s'améliorer mais répond aux attentes principales de ses utilisateurs. Les actions mises en place pour répondre à la logique de l'amélioration continue de la qualité du système tendent à permettre au eportfolio et à son système environnant de répondre à des objectifs plus étendus d'EPP notamment par l'intermédiaire de la mise en place d'évaluation via des « staffs protocolés », initiée dans le secteur de la médecine du don. La simplification du processus et les développements de nouvelles fonctionnalités de communication via la plate-forme laissent entrevoir une adaptabilité du système déployé à de nouveaux métiers à travers l'hexagone.

Références

[1] Haute Autorité de Santé. ' L'évaluation des pratiques professionnelles dans le cadre de l'accréditation des établissements de santé. Juin 2005 [a]. http://www.has-sante.fr/portail/upload/docs/application/pdf/guide_epp_juin_2005.pdf. (Dernière consultation : 12/11/2008).

[2] Baron GL, Bruillard E. Les technologies de l'information et de la communication en Èducation aux USA : ÈlÈments d'analyse sur la diffusion d'innovations technologiques. *Revue Française de Pédagogie* 2003 ; n°145 : 37-49.

[3] Ktat S. L'impact des nouvelles technologies de l'information sur la performance des auditeurs : validations empiriques dans le contexte tunisien. Doctoriales du GDR TIC & Société 28-2. Janv. 2004; 2-5.

[4] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. Perception de l'utilité perçue facilité d'emploi, et l'acceptation par les utilisateurs des technologies de l'information. *MIS Quarterly*, 13(3) : 319-39.

[5] Brangier E, Hammes S. Comment mesurer la relation humain-technologies-organisation? Élaboration d'un questionnaire de mesure de la relation humain-technologie-organisation basée sur le modèle de la symbiose. *PISTES©* Vol. 9 No. 2– Recherche. Oct. 2007 ; 1-3.

[6] Sicotte C, Moreault MP, Paré G. La mesure du succès en matière de déploiement de Système d'information clinique. *Journée Francophones d'Informatique Médicale,* Lille.12-13 mai 2005 ; 2.

[7] DeLone WH, McLean ER. The DeLone and McLean model of Information Systems Success. *Journal of Management Information Systems* 2003; 19, (4): 9-30.

[8] Chin JP, Diehl VA, Norman KL. Development of an instrument measuring user satisfaction of the human-computer interface. CHI 88: *Proceedings of the SIGCHI conference on Human factors in computing systems.*213-18.

[9] Livari J. An empirical test of DeLone-McLean Model of information system success. *The DATA BASE for Advances in Information Systems.* Spring 2005 (Vol. 36, N°2) 9:14: 23.

[10] Agarwal R, Prasad J. Are individual differences germane to the acceptance of new technologies? *Decision Sciences* 1999; 30 (2): 361-91.

[11] Lucas HC, Spitler VK. Technology usage performance: a field study of broker workstations. *Decision Sciences* 1999; 30 (2): 1-21.

[12] Staccini P, Rouger P. Modeling and using a web-based and tutored portfolio to support certification of professional competence in transfusion medicine. *AMIA Annu Symp Proc* 2008 ; 697-701.

Adresse de correspondance

Sophie Vessière

Institut National de la Transfusion Sanguine

6, rue A. Cabanel 75015 aris

E-mail : svessiere@ints.fr

Chapitre 5

Aide à la décision en pratique clinique

Informatique et Santé, 2009 (17) : 185-190
© Springer – Verlag France

Prescription informatisée : de l'ordre à l'itinéraire clinique

Mathias Tschopp, Magali Despond, Damien Grauser, Jean-Christophe Staub, Christian Lovis

Unité d'Informatique Clinique, Service d'Informatique Médicale, Hôpitaux Universitaires de Genève, Suisse

Abstract

Computer-based physician order entry has become a key component of electronic health record systems and the topic of many scientific publications. The importance of a global prescription platform, that includes drugs but also the other aspects of patients care, is often undervalued. So is the ability to create order sets targeted at common clinical problems, which will translate common clinical practice a lot more efficiently that single basic orders. This logic can be pushed further by developing clinical pathways that will fully realize the potential of e-prescribing as far as security, quality and efficiency.

Keywords

Clinical Pathways; Order Entry.

1 Introduction

La National Library of Medicine définit un itinéraire clinique comme « *Schedules of medical and nursing procedures, including diagnostic tests, medications, and consultations designed to effect an efficient, coordinated program of treatment* », généralement traduit en « *critical pathways* » ou encore « *clinical pathway* » dans la littérature anglo-saxonne. Cette définition met l'accent sur l'organisation globale de la prise en charge d'un patient, en termes opérationnels. Un itinéraire clinique va au-delà des soins médicaux et infirmiers en englobant toutes les activités impliquées directement ou indirectement dans les soins.

De manière générale, on distingue trois niveaux de conception d'itinéraires cliniques :

1) les itinéraires cliniques standardisés, visant une population homogène de patients ;
2) les itinéraires cliniques spécifiques, s'adressant à des patients dont on ne peut anticiper a priori tous les aspects de la prise en charge, au niveau d'un groupe, mais qui, une fois leurs spécificités connues, suivront un itinéraire adapté ;
3) le *case management*, qui s'intéresse à l'optimisation de la concertation interdisciplinaire, dans la prise en charge de groupes de patients peu prévisibles.

Ce travail s'attache à décrire le premier type en relation avec la prescription informatisée. Les itinéraires cliniques standardisés sont utilisés pour un processus de soins bien prévisible au niveau d'un groupe de patients avec des caractéristiques comparables et connues. Un exemple en est l'intervention pour la cataracte, la période post-partum après un accouchement par voie basse ou encore une intervention pour une cholécystectomie laparoscopique. Dans ces itinéraires cliniques, le déroulement est prévisible, les interventions et objectifs sont standardisés au niveau de la population de patients. De tels itinéraires ont des impacts importants, tant sur la qualité que sur les coûts, notamment en contrôlant la variabilité de la prise en charge [1-3].

Un itinéraire clinique est donc construit sur les connaissances issues de la médecine factuelle, l'expérience clinique, les attentes des patients, les possibilités et les limites de l'organisation en particulier en termes opérationnels. Ce qui caractérise donc la mise en place d'un itinéraire clinique réside dans le fait que l'on ne commence pas avec un dossier patient vierge, mais que le déroulement escompté des soins et de leurs effets est déjà établi à l'avance, ce qui permet la planification des soins pour le patients, mais aussi des ressources dans l'hôpital. Une discussion plus complète sur ce sujet dépasse le cadre de cet article, en particulier concernant les aspects liés à l'évaluation, pour lesquels le lecteur pourra se référer aux publications de l'équipe de l'université de Louvain [4, 5].

2 Contexte

Les Hôpitaux Universitaires de Genève (HUG) regroupent les hôpitaux publics du canton de Genève répartis sur quatre grands campus et sur de nombreux sites ambulatoires. Ils couvrent tous les types de soins hospitaliers, primaires, secondaires et tertiaires ainsi qu'ambulatoires. Les HUG ont un peu plus de 2 000 lits pour plus de 45 000 hospitalisations et 850 000 consultations par an.

Le système d'information clinique (SIC) des HUG est une solution basée sur l'intégration de produits existants et le développement de solutions propres, en l'absence de solutions commerciales satisfaisantes. Ceci est le cas de la plate-forme de prescription. Les développements sont faits dans l'environnement Java, sur une architecture fondée sur des composants métiers, un middleware orienté messages, et l'utilisation massive de standards d'interopérabilité. Le dossier patient intégré (DPI), client opérationnel du SIC, est utilisé quotidiennement par plusieurs milliers d'utilisateurs. Chaque jour, ce sont plus de 65 000 images, plus de 2 500 documents, plus de 15 000 prescriptions qui sont enregistrées dans le système.

3 PresCo, la *Pres* cription *Co* nnectée

La plate-forme de prescription (PresCo) est conçue pour permettre de prescrire tous les types de soins nécessaires à la prise en charge des patients. La liste du tableau 1 illustre certains des types utilisés, de manière non exhaustive :

Tableau 1 - Types de prescription.

Attitude thérapeutique
Mobilisation
Régime
Surveillances
Soins infirmiers
Médicaments
Laboratoire
Radiologie
Consultations spécialisées

En outre, de nombreux outils d'aide à la décision ou de facilitateurs pour la prise de décision, sont implémentés. Le système permet d'avoir des rappels, par exemple la nécessité de documenter les mesures de précautions infectieuses ; des alertes, par exemple en cas d'interactions médicamenteuses, ou de règles complexes de prescription comme

pour les dosages pédiatriques. Le système supporte des regroupements d'ordres simples, comme par exemple l'INR et les anticoagulants ou plus complexes comme l'ensemble des ordres nécessaires à la prise en charge d'une patiente après accouchement par voie basse.

Tableau 2 - Fonctionnalités majeures.

Rappels (attitude, surveillance, etc.)
Alertes (interactions, traitement déjà fait, dosages, etc.)
Informations (laboratoire utile, description d'examens, indications, etc.)
Notes de prescriptions automatiques
Ordres groupés (admission médecine, post-accouchement, etc.)
Gestion de flux (itinéraires cliniques)

4 Itinéraires cliniques et dossier informatisé

L'architecture du SIC des HUG se prête particulièrement à la mise en place des itinéraires cliniques. Construit sur une approche par composants orientés services et un middleware piloté par des messages, les fonctionnalités du SIC peuvent être comparées à autant de musiciens pour lesquels l'itinéraire clinique va jouer le rôle de partition. À chaque itinéraire, sa partition. À cet ensemble, il a été décidé de mettre en place un moteur de flux pour jouer le rôle de chef d'orchestre, permettant ainsi à l'ensemble des fonctionnalités, documentation clinique, laboratoire et autres résultats d'investigations, ordres de toute sortes, etc. d'être organisés et pilotés de manière cohérente dans le cadre d'un processus complet de prise en charge, de l'admission à la sortie. Notre choix c'est porté sur *JBoss jBPM*, qui est un système de gestion de flux permettant de coordonner des applications et des services dans le cadre de processus. Il facilite la description et la gestion de flux d'information ainsi que la coordination entre actes, objets et acteurs. *jBPM* est un logiciel libre écrit en Java et s'intègre parfaitement dans l'architecture du système d'information clinique des HUG. En plus du moteur de flux, *jBPM* fournit des interfaces de développement simple et permet de poser graphiquement les différentes étapes qui composent un processus [6].

Un itinéraire clinique est constitué de phases qui se suivent ou se superposent. Le passage d'une phase à l'autre, qui correspond à des étapes de prises en charge, est dicté par des règles qui varient en fonction de l'itinéraire clinique. Ceci peut correspondre à un temps écoulé, à des résultats d'investigation, ou encore à l'état clinique du patient, par exemple. Il s'agit souvent d'une association de divers éléments. À une phase de l'itinéraire clinique va correspondre des éléments liés aux soins, aux investigations et également à la documentation de la situation, entre autres. De même, pour chaque phase comme pour l'entièreté de l'itinéraire, il y aura divers éléments dont on mesurera la variance. Il faut souligner que les itinéraires peuvent être complexes. Ainsi par exemple, l'itinéraire clinique développé aux HUG concernant l'embolie pulmonaire associe un itinéraire diagnostic et un itinéraire thérapeutique. Les deux itinéraires se chevauchent et s'influencent mutuellement, jusqu'à ce que le diagnostic soit définitivement retenu ou rejeté.

Sans trop entrer dans les détails de ces aspects, nous illustrons ici un itinéraire clinique développé dans le cadre de la prise en charge de l'insuffisance cardiaque gauche.

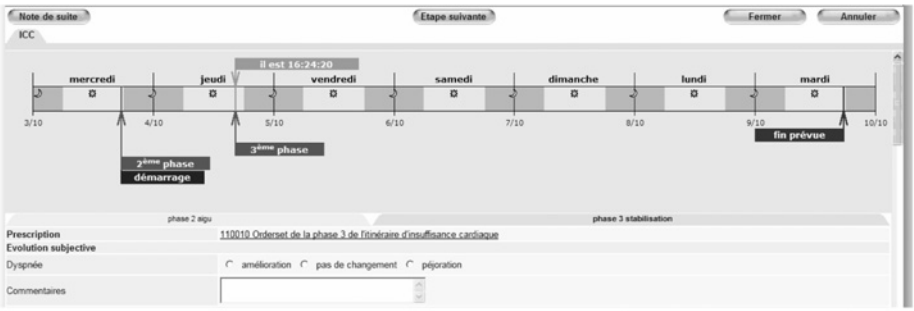


Figure 1 - Vision du flux et documentation médicale.

La figure 1 illustre la vision globale de l'itinéraire, son début ainsi que celui de chaque phase, la date de sortie prévue du patient. L'écran permet également de faire la documentation spécifique à chaque phase et de prescrire l'ensemble des ordres spécifiques à chaque phase.

La zone graphique temporelle permet au médecin de visualiser facilement comment se déroule la prise en charge et son adéquation avec ce qui était prévu. La documentation clinique de l'évolution du patient sera déterminante pour les règles de transition à une phase suivante de la prise en charge. Si les éléments cliniques relevés ne permettent pas de passer à une phase suivante, par exemple en cas d'aggravation de la dyspnée, alors les éléments graphiques passeront au rouge, indiquant ainsi que l'évolution ne se passe pas comme prévu. Ceci devrait permettre une réévaluation plus rapide de la situation, accompagnée d'une correction plus rapide de la prise en charge.

Figure 2 - Ordres groupés d'une phase de l'itinéraire.

La figure 2 illustre les ordres spécifiques à la prise en charge d'un patient dans une des phases. Il faut noter ici que l'ensemble des soins, surveillance et investigations peuvent être préparés. Tout en étant guidé, le prescripteur peut néanmoins adapter de manière précise l'ensemble des ordres.

Basé sur le modèle de l'itinéraire en cours, la documentation clinique est guidée et générée pendant la prise dans l'itinéraire (figure 3).

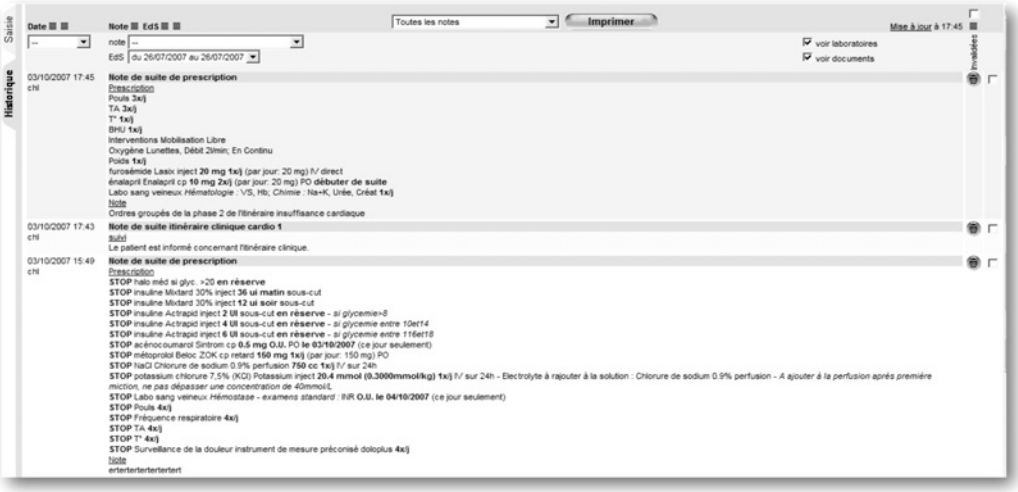


Figure 3 - Génération de la documentation.

5 Évaluation

La réalisation d'itinéraires cliniques est une entreprise complexe, engageant des partenaires de multiples disciplines, cliniques, paracliniques et techniques. Face à l'effort mis en œuvre, il est logique de se poser la question de l'évaluation, surtout lorsque la plate-forme utilisée est un système d'information permettant une totale traçabilité des interventions sur les patients par les soignants. Une discussion détaillée des problématiques d'évaluation de l'impact de l'implémentation des itinéraires cliniques sur des indicateurs, tels que la durée de séjour, les coûts hospitaliers et le devenir des patients, dépasse le cadre de cet article, de notre expérience naissante dans le domaine et peut être trouvée ailleurs. Il est cependant indéniable que, au fur et à mesure que l'offre en termes d'itinéraire s'étoffera dans notre institution, il nous faudra suivre de près différents éléments tels que le taux d'utilisation des itinéraires, dont l'utilisation, il faut le souligner, reste facultative.

Les différents indicateurs relatifs à la qualité de la prise en charge pourront être comparés, soit de manière historique, soit de manière prospective mais la validité des observations ainsi générées ne passe que par l'existence d'une offre suffisante, tant quantitative que qualitative, en termes d'itinéraire, afin de ne pas tomber dans un biais de sélection des cas les plus simples, se prêtant naturellement plus facilement à l'inclusion dans des itinéraires standardisés.

À noter que les indicateurs varient pour chaque itinéraire. Un certain nombre d'entre eux sont actuellement en mode d'évaluation pour l'itinéraire clinique de l'insuffisance cardiaque et un outil permettant de les suivre en temps réel est à disposition des médecins cadres des services concernés.

6 Conclusion

La mise en place d'itinéraires cliniques permet l'amélioration de l'efficience dans la prise en charge des patients, une optimisation de la qualité et de l'économicité, dans un souci aussi d'améliorer la sureté des soins [7, 8]. Cette amélioration passe principalement par la standardisation des pratiques, qui limite les variations inutiles dans la prise en charge des pathologies, souvent génératrices de coûts ; par une meilleure planification et coordination

des différents acteurs de santé, permettant d'éviter les journées d'hospitalisation inappropriées, et par une dimension didactique, intéressante dans un hôpital universitaire.

Le développement des itinéraires cliniques est un défi pour les multiples partenaires impliqués, médecins, soignants, patients, entres autres dans une vision de standardisation des processus de prise en charge, c'est un nouveau paradigme. En outre, le passage de la vision théorique d'une prise en charge à son opérationnalisation pratique et la description fine de chaque action à entreprendre, chaque résultat attendu, nécessitent également une profonde remise en question.

Pour les sciences de l'information, le défi est au moins aussi grand. Tant le besoin d'abstraction que le besoin de formalisation sont nécessaires à la description d'un itinéraire. La prise en compte des contraintes en ressources humaines, médico-techniques et logistiques sont indispensables. En ce qui concerne un système d'information clinique, sa capacité à être entièrement piloté par un moteur de flux, tant pour les ordres que pour les actes, la documentation, et l'ensemble des actions possibles est un élément critique et absent de tous les systèmes commerciaux à quelques rares exceptions près. En outre, la plate-forme de prescription, pierre de voûte d'un tel système, doit permettre la gestion de l'ensemble des actions requises pour la prise en charge d'un patient, de son admission à sa sortie. La planification et le suivi de ces actions doivent pouvoir être intégré dans un agenda patient unifié.

Références

[1] Vanhaecht K, Sermeus W, Tuerlinckx G, Witters I, Vandenneucker H, Bellemans J. Development of a clinical pathway for total knee arthroplasty and the effect on length of stay and in-hospital functional outcome. *Acta Orthop Belg* 2005; 71(4): 439-44.

[2] Coffey RJ, Richards JS, Remmert CS, LeRoy SS, Schoville RR, Baldwin PJ. An introduction to critical paths. *Qual Manag Health Care* 2005; 14(1): 46-55.

[3] Liesmons I. The added value of a process oriented hospital information system supporting the integrated patient care. *Stud Health Technol Inform* 2004; 110: 1-8.

[4] De Bleser L, Depreitere R, De Waele K, Vanhaecht K, Vlayen J, Sermeus W. Defining pathways. *J Nurs Manag* 2006; 14(7): 553-63.

[5] Vanhaecht K, De Witte K, Depreitere R, Sermeus W. Clinical pathway audit tools: a systematic review. J Nurs Manag 2006; 14(7): 529-37.

[6] JBoss. Java business process management jBPM. http://wwwjbosscom/fr/products/jbpm. 2007.

[7] Dy SM, Garg P, Nyberg D, Dawson PB, Pronovost PJ, Morlock L, et al. Critical pathway effectiveness: assessing the impact of patient, hospital care, and pathway characteristics using qualitative comparative analysis. *Health Serv Res* 2005; 40(2): 499-516.

[8] Kent P, Chalmers Y. A decade on: has the use of integrated care pathways made a difference in Lanarkshire? *J Nurs Manag* 2006; 14(7): 508-20.

Adresse de correspondance

Mathias Tschopp

Service d'Informatique Médicale, Unité d'Informatique Clinique

Hôpitaux Universitaires de Genève, Genève, Suisse

E-mail : mathias.tschopp@hcuge.ch

Informatique et Santé, 2009 (17) : 191-202
© Springer – Verlag France

Caractérisation de l'évolution d'une base de connaissances suite à la réactualisation de recommandations de pratiques cliniques : une étude de cas sur la prise en charge des tumeurs de la vessie

Brigitte Séroussi[1], Jacques Bouaud[2], Ambre Brizon[2], Thibault Culty[3], France Mentre[4], Vincent Ravery[3]

[1]Université Paris 6, UFR de Médecine, Paris, France
Hôpital Tenon, Dépt de Santé Publique, Assistance Publique – Hôpitaux de Paris, Paris, France
Université Paris 13, UFR SMBH, LIM&BIO, Bobigny, France ; APREC, Paris, France
[2]STIM, DSI/AP-HP, Paris, France ; INSERM, UMR_S 872, Eq. 20, Paris, France
[3]Hôpital Bichat, Service d'Urologie, Assistance Publique – Hôpitaux de Paris, Paris, France
[4]INSERM U738, Paris, France ; Université Paris 7, UFR de Médecine, Paris, France

Abstract

The constant evolution of medical knowledge requires clinical practice guidelines (CPGs) to be regularly updated. Clinical decision support systems based on computerized CPGs have thus to be revised as their knowledge sources evolve. We propose a formal characterization of recommendations evolution between two successive CPG versions. Each atomic recommendation is represented as a rule linking a clinical situation with a recommended action plan. Using subsumption functions on recommendation components, we identified 7 evolution patterns of recommendations. The method has been applied on the knowledge bases of the 2002 version of the French bladder cancer guidelines and its 2004 revision. Surprisingly, despite similar coverage, only 54% of the 2002 version was formally reused in 2004 and 36% of the 2004 version were considered new recommendations. Evaluation of CPG update on actual patient medical records from the Urology Department of the Bichat-Claude Bernard Hospital showed a significant fall of physician decisions compliance with guidelines as a side-effect of the CPG revision.

Keywords

Computer decision support systems ; Clinical practice guidelines updating; Clinical practice guidelines computerization; Bladder cancer

1 Introduction

Il est encore fréquent d'observer des variations de pratiques cliniques entre les médecins. En effet, il n'est pas rare de constater qu'un même patient ne se voit pas proposer la même prise en charge selon le praticien qu'il consulte, ce qui met en question la qualité des soins qu'il reçoit. Si certains traitements sont « seulement » sous-optimaux, de nombreuses erreurs médicales sont de plus en plus souvent rapportées dans la plupart des établissements de soins. Ainsi, la mise en œuvre de la « médecine factuelle » ou « médecine fondée sur les preuves » (ou encore *evidence-based medicine*) constitue

aujourd'hui un enjeu majeur dans les pays développés. Celle-ci consiste à utiliser, dans les prises de décision, de manière rigoureuse et judicieuse, les meilleures données scientifiques établies lors d'études cliniques valides. Dans ce contexte, les recommandations de pratiques cliniques (RPC) ou *guidelines* représentent des outils permettant d'améliorer la qualité des soins.

Initialement produites sous la forme de documents textuels, les RPC sont classiquement structurées sous la forme d'un catalogue de situations cliniques auxquelles sont associées les prises en charge recommandées. Si l'objectif est que les recommandations s'appuient sur les « preuves » scientifiques issues de l'état de l'art, en l'absence d'*evidence*, elles résultent le plus souvent d'accords professionnels.

De nombreuses études ont montré que la simple diffusion des RPC, sur support papier ou numérisées, n'avait que très peu d'impact sur les pratiques des médecins [1]. Au contraire, les systèmes d'aide à la décision médicale (SADM), permettant, au moment de la décision, la diffusion des recommandations dans une approche centrée patient, ont permis d'améliorer l'observance des RPC par les médecins [2, 3]. Dérivés des anciens systèmes experts, les SADM reposent actuellement sur des modèles formels construits à partir des RPC textuelles. Pourtant, la traduction des RPC textuelles en bases de connaissances (BC) structurées est aujourd'hui encore une tâche complexe qui mobilise de nombreuses équipes [4, 5, 6].

Par ailleurs, du fait de l'accélération de la production de nouvelles connaissances médicales, l'état de l'art est une notion relative qui évolue au cours du temps. Ainsi les RPC sont des documents impliqués dans un cycle de vie et elles doivent être régulièrement actualisées. Ce processus est pris en charge par les instances à l'origine de leur production avec une fréquence qui peut être fixée a priori (de l'ordre de 3 ans) ou bien lorsque de nouveaux résultats scientifiques, remettant en question les pratiques cliniques recommandées jusqu'alors, sont publiés (cf. le « NICE guidelines manual » [7]). En France, la Haute Autorité de Santé (HAS) est responsable de l'élaboration, de la diffusion et de la réactualisation des RPC. La méthode de réactualisation des RPC généralement utilisée capitalise rarement sur les versions précédentes des documents, et les versions réactualisées sont souvent de nouvelles RPC totalement réécrites. En conséquence, lors de la réactualisation des RPC, les BC structurant le contenu sémantique de la version précédente ne peuvent pas être réutilisées et de nouvelles BC doivent être développées pour modéliser la nouvelle version des RPC. Ainsi, la réactualisation des BC au fur et à mesure de la publication des versions des RPC reste un problème à résoudre.

Nous avons étudié la réactualisation d'une BC au cours de deux révisions successives de RPC. Le travail a porté sur les RPC pour la prise en charge des tumeurs de la vessie développées par l'Association Française d'Urologie (AFU) dans leurs versions de 2002 et 2004. Pour chaque document, nous avons construit une représentation formelle des connaissances sous la forme d'un arbre de décision utilisé comme BC du SADM UroDoc. L'objectif est de comparer les deux BC construites afin de proposer une caractérisation formelle des types d'évolution résultant de la réactualisation et d'en évaluer la distribution sur le cas particulier de la prise en charge des tumeurs de vessie. Par ailleurs, une étude sur dossiers (service d'urologie, hôpital Bichat-Claude Bernard) nous a permis d'évaluer l'impact de l'actualisation des RPC sur les taux de conformité des décisions aux RPC de 2002 et 2004.

2 Contexte

2.1 La prise en charge des tumeurs de la vessie

Avec plus de 10 000 cas par an, le cancer de la vessie est, en France, le 4e cancer masculin et le 8e cancer féminin. Les tumeurs de la vessie couvrent une grande variété de pathologies allant des tumeurs superficielles relativement bénignes aux adénocarcinomes au pronostic péjoratif pour lesquels des traitements radicaux sont nécessaires.

Le traitement des tumeurs superficielles de la vessie repose principalement sur la résection transurétrale de vessie (RTUV). Des instillations adjuvantes de mitomycine C ou de BCG peuvent être indiquées afin de prévenir les récidives ou les métastases chez les patients présentant des facteurs de risque. Une surveillance stricte et prolongée est dans tous les cas indispensable. Le traitement standard des patients présentant des cancers invasifs est la cystectomie. Pourtant, les taux de guérison suite à la chirurgie ne sont que de 10 à 30 % dans les cancers de vessie localement avancés. Actuellement, de nombreux profils cliniques restent sans schéma de prise en charge recommandé ce qui explique pourquoi la prise en charge des cancers de la vessie est un domaine de recherche très actif.

Fondée en 1896, l'Association Française d'Urologie est une instance nationale qui représente la majorité des urologues français. Le comité d'oncologie est constitué d'un groupe d'experts responsables de la publication des RPC sur la prise en charge des cancers urologiques et de leur réactualisation afin de diffuser à l'ensemble des professionnels les nouveaux standards de prise en charge, et d'améliorer ainsi la qualité des soins et les taux de survie. Dans le cas particulier du cancer de la vessie, les RPC ont été régulièrement réactualisées. L'objet de cette étude a porté sur l'évolution des RPC sur la prise en charge des cancers de la vessie entre les versions de 2002 et de 2004.

2.2 La modélisation des RPC pour les SADM

2.2.1 Les approches classiques

Les SADM reposent sur des BC structurées qui implémentent une modélisation formelle des RPC textuelles dont la traduction reste une tâche complexe.

La plupart des outils permettant la modélisation des RPC considèrent que les textes des RPC sont les documents sources à partir desquels le travail de traduction sous la forme de BC doit être réalisé. Il existe d'une part des approches *centrées sur les documents* (telles que GEM, DeGel [6], Stepper [5]) qui mettent en œuvre un processus graduel d'annotations du document permettant sa réécriture progressive en BC. La plupart du temps, ces outils de marquage utilisent des DTD XML pour associer des descriptions textuelles à des modèles semi-structurés des RPC. La relation entre le modèle semi-formel des portions de texte balisées et le texte d'origine est ainsi préservée.

D'autre part, dans les approches *centrées sur les modèles*, un expert du domaine est sollicité pour construire le modèle formel de la RPC à partir du texte d'origine. Ainsi, c'est lui qui choisit le point de vue interprétatif permettant de lever les ambiguïtés du texte et les modalités d'explicitation de l'implicite, voire de complétion des incomplétudes.

Dans les deux types d'approches, la dernière étape consiste à traduire le modèle formel de la RPC dans un formalisme approprié (GLIF, Asbru, PROforma, etc.) pour construire la BC. La BC finale est alors intégrée dans un SADM. Dans la grande majorité des cas, ces SADM sont exécutés automatiquement sur les données patient disponibles (dossier médical informatisé, système d'information hospitalier) et produisent des alertes (ou

rappels) à destination du professionnel de santé lui rappelant les bonnes pratiques recommandées adaptées au patient qu'il devrait, sauf cas particulier, mettre en œuvre.

2.2.2 L'aide à la décision documentaire

Le succès des SADM basés sur les alertes n'est pas systématique [2]. La pertinence et l'exactitude des conseils générés automatiquement sont en effet souvent approximatives : d'après Judge *et al.* [8], seule une petite moitié des alertes proposées par les systèmes d'aide à la prescription médicamenteuse serait appropriée ce qui pourrait conduire les praticiens à ne plus tenir compte des alertes dans 49 % à 96 % des cas [9]. De plus, le manque de flexibilité des approches automatisées est souvent critiqué : si la réduction de la singularité d'un patient à un ensemble de *data* est une étape nécessaire pour documenter le dossier patient informatisé, le raisonnement médical impliqué dans le processus décisionnel nécessite une interprétation flexible et contextuelle de notions.

C'est pourquoi, d'autres approches, dans lesquelles les connaissances des RPC sont structurées de manière à permettre à un utilisateur de retrouver des recommandations centrées patient plus facilement que dans les textes, tout en gardant la flexibilité de l'interprétation des notions, ont été proposées pour aider les praticiens [10, 11]. Ainsi, la BC est bien sûr structurée, par exemple sous la forme d'un arbre de décision, mais elle n'est pas automatiquement exécutée et, au contraire, parcourue de manière hypertextuelle. Le médecin joue alors le rôle d'un médiateur fournissant les informations caractérisant le patient sans qu'elles aient besoin d'être codées et interprétant dans le contexte de son patient les informations qui lui sont proposées en retour. À chaque étape du parcours, un critère décisionnel est affiché avec la liste des modalités possibles (exhaustives et exclusives). L'utilisateur choisit la valeur qui instancie au mieux le critère décisionnel et contrôle ainsi la caractérisation du profil patient jusqu'à obtenir les meilleurs traitements centrés-patient recommandés.

Le système OncoDoc [11] appliqué à la prise en charge thérapeutique du cancer du sein a été développé selon ces principes, le mode guidé d'ASTI [12], appliqué à la prise en charge de l'hypertension artérielle, du diabète, des dyslipidémies, de la fibrillation auriculaire, etc., également. Nous avons développé un système comparable, UroDoc [13], pour la diffusion des RPC sur la prise en charge des cancers de la vessie de l'AFU.

2.3 Le problème de l'actualisation des RPC

Lors d'une mise à jour des RPC, le processus d'élaboration des nouvelles recommandations est souvent entièrement reconduit sans nécessairement capitaliser sur les versions précédentes, ce qui aboutit à la publication de textes totalement nouveaux, sans relation avec les versions précédentes. De récentes recommandation anglo-saxonnes pour la mise à jour des RPC suggèrent, sous certaines conditions, de ne modifier que les parties concernées et de conserver à l'identique les parties qui n'ont pas changé [7, 14].

Après l'actualisation d'un texte de RPC, la mise à jour des BC qui s'appuyaient sur la version précédente du texte doit également être effectuée. Dans les approches centrées sur les modèles, selon la quantité de modifications, la mise à jour peut être simple ou compliquée allant jusqu'à la reconstruction complète d'une nouvelle BC. C'est une tâche complexe qui requiert du temps de développement. Toutefois, il existe des éditeurs de BC permettant la gestion des versions et facilitant la réutilisation de composants existant dans les versions antérieures [15, 16].

Dans les approches centrées sur les documents, lorsqu'une RPC est réactualisée, l'actualisation de la BC qui lui est associée est censée être guidée par les évolutions du document. Cependant, détecter ces évolutions n'est pas trivial. Même en mettant en œuvre

des techniques de traitement du langage naturel, l'identification des portions de textes correspondant à un contenu sémantique nouveau est une tâche difficile. Par conséquent, il est fréquent que la BC modélisant la nouvelle version des RPC soit totalement reconstruite. Aussi, les BC utilisées dans les SADM devraient être réactualisées plus fréquemment, en continu, au fur et à mesure de la publication des preuves. Pour pallier ce problème, des auteurs [17, 18] ont proposé la notion de RPC « vivantes » ou *living guidelines* afin de faciliter les mises à jour. En effet, le texte de RPC vivantes est construit de façon incrémentale à partir des versions précédentes de façon à mettre en évidence les informations qui ont été changées : les modifications de chaque révision sont annotées, datées et mises en relief. Dans ce cas, les différences entre deux versions successives d'un document sont facilement identifiables. Au niveau de la réactualisation des BC, les portions inchangées du texte correspondent à des parties déjà modélisées dans la BC et sont donc conservées telles quelles. Seules les nouvelles portions du texte doivent être modélisées. Actuellement, les RPC françaises, entre autres, ne sont pas élaborées de cette manière.

3 Matériel

Pour établir, sur le plan formel, les différences entre les deux versions des RPC, il est nécessaire de disposer des BC correspondantes. Pour construire ces BC, nous avons étudié les deux textes des recommandations de prise en charge des tumeurs de vessie dans l'ordre chronologique de leur élaboration, celui de 2002 puis celui de 2004. Pour chaque RPC, le travail de modélisation a été conduit de façon indépendante ; le même processus de construction a été appliqué afin de garantir la comparabilité des deux formalisations et ne pas biaiser l'évaluation des différences.

En suivant les principes d'encodage des RPC proposés par Shiffman [4], nous avons tout d'abord procédé à l'étape d'atomisation visant à extraire les concepts élémentaires du texte écrit en langage naturel (*hématurie microscopique*, *grade*, *instillations endovésicales*, etc.). Puis, nous avons réalisé une étape de désabstraction afin d'abaisser le niveau de généralité utilisé dans les RPC pour décrire les variables de décision et d'action au niveau nécessaire à leur opérationnalisation en pratique clinique (« *récidive précoce* » a été traduit en « *récidive à un plus haut stade ou au même stade que la tumeur initiale dans un délai de 3 mois* »). Enfin, l'étape de désambiguisation a permis d'établir l'axe sémantique d'interprétation à adopter pour les formules ambiguës (« *suspectes* » dans « *des biopsies seront réalisées dans les aires suspectes* »). L'étape la plus importante est la vérification de la complétude qui permet d'assurer que les critères décisionnels implicites et les modalités implicites de critères décisionnels explicites ont été explicités sous la forme de variables de décision et d'action identifiées, et qu'il y a une attitude thérapeutique (que ce soit une recommandation ou un accord professionnel selon la disponibilité des preuves) pour toute combinaison logique de critères, ce qui garantit que l'ensemble de toutes les situations cliniques décrit par l'expansion de l'arbre de décision est couvert. Ainsi, deux arbres de décision différents ont été construits, un pour les RPC de 2002 et un pour les RPC de 2004. Ces arbres de décision sont les BC du système UroDoc qui se décline en conséquence en deux versions : UroDoc-2002 et UroDoc-2004.

4 Méthode

Nous proposons une méthode permettant de comparer deux BC construites à partir de deux versions successives d'une même RPC. Au-delà de simplement mettre en relief les

différences existant entre les deux BC, la méthode fournit une caractérisation formelle des dimensions de l'évolution des connaissances des RPC.

4.1 Notations

Une fois modélisées, les RPC peuvent être considérées comme équivalentes à un ensemble de recommandations atomiques, notées $\{R_i\}_i$. Classiquement, chaque recommandation R_i est constituée d'un couple (S_i, P_i) et correspond à une règle de production telle que $S_i \Rightarrow P_i$, signifiant qu'un plan d'actions P_i (prescriptions, examens, surveillance, etc.) est recommandé dans la situation clinique S_i (caractérisant un profil patient). Une situation clinique S_i est définie par une conjonction de critères considérés comme des variables de décision : $S_i = \{c_{i,n}\}_n$. Un plan d'actions P_i est défini par une séquence ordonnée d'actions : $P_i = (a_{i,m})_m$.

4.2 Définitions

Deux recommandations R_i et R_j sont identiques si et seulement si leurs situations cliniques sont identiques et leurs plans d'actions sont identiques. Deux situations cliniques sont identiques si leurs ensembles de variables de décision sont identiques ; deux plans sont identiques si leurs séquences d'actions sont identiques (même ordre).

$$(R_i = R_j) \Leftrightarrow (S_i = S_j \wedge P_i = P_j)$$

Utiliser l'identité comme fonction de comparaison conduit à deux statuts, « identique » ou « différent », sans tenir compte d'aucune similitude. Nous avons utilisé des méthodes de subsomption pour pouvoir comparer les recommandations à des niveaux plus abstraits.

Une première transformation possible est l'abstraction « ontologique ». Sans considérer une véritable hiérarchie ontologique, le principe consiste à décider qu'il existe deux niveaux : le niveau « haut » dont les concepts ne sont pas abstraits, et le niveau « bas », a priori celui de la représentation, dont les notions sont généralisées dans les concepts du niveau « haut ». L'élaboration du niveau « haut » a été manuellement effectuée et une fonction, notée *abst* a été construite entre des éléments de la représentation et des concepts abstraits qui les subsument. L'abstraction est vue ici comme un moyen de considérer apparentées des notions différentes tout en évitant le problème de la sur-généralisation,

L'abstraction s'applique sur les variables de décision et sur les actions. Par exemple, *abst*(gemcitabine) = *abst*(cysplatine) = chimiothérapie ; ainsi la gemcitabine et le cysplatine, molécules distinctes, sont considérées comme des chimiothérapies, mais non comme des radiothérapies ou des chirurgies. Dans le cas où, lors de l'évolution des RPC, une notion disparaît sans avoir de notion proche dans la nouvelle version, on considère que son abstraction est nulle. Par extension, les situations cliniques et les plans d'actions peuvent également être abstraits en substituant chacun des composants par son abstraction, les valeurs nulles étant éliminées.

On considère que deux plans d'actions, qui sont des séquences, ont le même contenu s'ils sont constitués du même ensemble d'actions, donc indépendamment de l'ordre. On utilise donc la fonction booléenne *samecontent* pour comparer deux plans d'actions. Elle correspond en pratique à l'égalité au sens ensembliste :

$$samecontent(P_i, P_j) = (P_i \subseteq P_j) \wedge (P_j \subseteq P_i)$$

Les situations cliniques peuvent être comparées en utilisant une fonction de subsomption basée sur la structure, notée *subsumes*, qui pour deux arguments S_i et S_j retourne vrai quand S_j est plus spécifique que S_i. Une manière simple d'implémenter une telle fonction est d'utiliser l'inclusion :

$$subsumes(S_i, S_j) = (S_i \subseteq S_j)$$

Ces fonctions de comparaison de contenu et de subsomption sont étendues en les combinant avec la fonction d'abstraction. Nous introduisons les deux fonctions *Similar* et *Derives* définies comme suit :

$$Similar(P_i, P_j) = samecontent(abst(P_i), abst(P_j))$$
$$Derives(S_i, S_j) = subsumes(abst(S_i), abst(S_j))$$

4.3 Patrons d'évolution entre deux recommandations

On se place dans un contexte où l'on dispose de deux ensembles de recommandations $\{R_i\}_i^v$ et $\{R_j\}_j^{v+1}$, correspondant respectivement aux versions v et $v + 1$ d'une BC.

On considère 2 recommandations, R_i^v et R_j^{v+1} issues de chacun des 2 ensembles et l'on se place dans l'hypothèse où R_j^{v+1} est une évolution de R_i^v. Sous cette hypothèse, les deux recommandations devraient partager une certaine similarité. On distingue 7 « patrons d'évolution » possibles caractérisant l'évolution de R_i^v à R_j^{v+1} lors de la mise à jour de la RPC. Ces patrons sont identifiés par l'algorithme de comparaison décrit en figure 1 :

1. Si $S_i^v = S_j^{v+1}$
2. alors si $P_i^v = P_j^{v+1}$
3. alors *Identité*
4. sinon si *Similar*(P_i^v, P_j^{v+1})
5. alors *Révision du plan d'actions*
6. sinon *Nouveau plan d'actions*
7. sinon si *Derives*(S_i^v, S_j^{v+1})
8. alors si $P_i^v = P_j^{v+1}$
9. alors *Révision de la situation clinique*
10. sinon si *Similar*(P_i^v, P_j^{v+1})
11. alors *Révision de la recommandation*
12. sinon *Nouvelle pratique*
13. sinon *Pas de correspondance*

Figure 1 - Algorithme d'identification des patrons d'évolution de R_i^v à R_j^{v+1}.

1. **Identité** : la recommandation R_i^v reste inchangée dans la nouvelle version $v + 1$.

   P. ex. $R_i^{2002} = R_j^{2004} = (\{$*cystectomie antérieure, pas de tumeur, TUR récente, T1, grade 1, récidive, récidive précoce, récidive après instillation de BCG, cystectomie contre-indiquée*$\}$, (*instillation de mitomycine, surveillance haut risque*)).

2. **Révision du plan d'actions** : les 2 situations cliniques sont identiques mais le plan d'actions a été révisé tout en étant similaire au précédent.

   P. ex. $S_i^{2002} = S_j^{2004} = \{$*résection incomplète, tumeur invasive, non métastatique, N2, chimiothérapie non contre-indiquée*$\}$ avec P_i^{2002} = (***chimiothérapie***) et P_j^{2004} = (***chimiothérapie adjuvante***).

3. **Nouveau plan d'actions** : les 2 situations cliniques sont identiques, mais leurs plans d'actions ne sont pas similaires, c.-à-d. R_j^{v+1} recommande un nouveau plan d'actions pour la même situation.

   P. ex. $S_i^{2002} = S_j^{2004} = \{$*résection incomplète, tumeur invasive, non métastatique, N2, T4b, chimiothérapie non contre-indiquée*$\}$ avec P_i^{2002} = (***radiothérapie***) et P_j^{2004} = (***surveillance***). Ici, les plans d'actions sont différents même après abstraction.

4. **Révision de la situation clinique** : la situation clinique S_j^{v+1} a été dérivée de S_i^v, c.-à-d. R_j^{v+1} est plus restrictive que R_i^v, mais le même plan d'actions est recommandé.

   P. ex. S_i^{2002} = {*résection incomplète, tumeur superficielle, RTUV réalisée, cystectomie non contre-indiquée*} et S_j^{2004} = {*résection incomplète, tumeur superficielle, RTUV réalisée,* ***Ta, Grade 1****, cystectomie non contre-indiquée*} avec $P_i^{2002} = P_j^{2004}$ = (*nouvelle RTUV*). Ici, 2 critères (Ta et grade 1) ont été ajoutés, sans modification des plans d'actions.
5. **Révision de la recommandation** : dans ce cas, les 2 composants de la recommandation ont subi des transformations, S_j^{v+1} est dérivée de S_i^v et leurs plans d'actions sont similaires.

   P. ex. S_i^{2002} = {*cystectomie antérieure, tumeur, tumeur résiduelle, tumeur actuelle superficielle, seconde reprise chirurgicale, cystectomie non contre-indiquée, urétrectomie non indiquée,* ***Ta, Grade 3***} et S_j^{2004} = {*cystectomie antérieure, tumeur, tumeur résiduelle, tumeur actuelle superficielle, seconde reprise chirurgicale,* ***Ta, Grade 3****, cystectomie non contre-indiquée, urétrectomie non indiquée,* ***sexe masculin***} avec P_i^{2002} =(*curage ilio-obturateur, cystectomie, entérocystoplastie*) et P_j^{2004} = (*prostato-cystectomie, curage, entérocystoplastie*). Ici, 2 critères (*Ta* et *grade 3*) ont été déplacés et un nouveau critère (*sexe masculin*) a été ajouté. Les 2 plans d'actions sont similaires, ils ont le même contenu (*samecontent*) après abstraction, en considérant que *abst(curage ilio-obturateur) = curage*, et que *abst(prostato-cystectomie)= cystectomie*.
6. **Nouvelle pratique** : S_j^{v+1} est dérivée de S_i^v, mais leurs plans d'actions ne sont pas similaires.

   P. ex. S_i^{2002} = {*cystectomie antérieure, tumeur, tumeur résiduelle, tumeur actuelle superficielle, seconde reprise chirurgicale, cystectomie contre-indiquée, urétrectomie indiquée, Tis,* ***CIS diffus***} et S_j^{2004} = {*cystectomie antérieure, tumeur, tumeur résiduelle, tumeur actuelle superficielle, seconde reprise chirurgicale, Tis, cystectomie contre-indiquée, urétrectomie indiquée,* ***sexe féminin***} avec P_i^{2002} =(*RTUV*) et P_j^{2004} = (*pelvectomie, curage, Bricker*). On note ici que *CIS diffus* est un critère devenu obsolète en 2004. S_j^{2004} dérive de S_i^{2002} après avoir éliminé *CIS diffus* et en ajoutant *sexe féminin*. Les 2 plans d'actions sont totalement différents.
7. **Pas de correspondance** : R_i^v et R_j^{v+1} n'ont rien en commun.

4.4 Comparaison de deux ensembles de recommandations

La méthode décrite ci-dessus permet une comparaison des recommandations deux à deux. Pour comparer deux ensembles de recommandations $\{R_i\}_i^v$ et $\{R_j\}_j^{v+1}$, chaque R_j^{v+1} est comparée à chaque R_i^v pour déterminer si un patron d'évolution est pertinent pour ce couple de recommandations. À la fin du processus, les recommandations R_j^{v+1} qui restent classées *Pas de correspondance* sont considérées comme de *Nouvelles recommandations* de la version $v + 1$. De la même manière, les recommandations R_i^v sans correspondance avec aucune des recommandations de l'ensemble $v + 1$ sont considérées comme des *Recommandations obsolètes* de la version v.

5 Résultats

5.1 Comparaison des bases de connaissances

Pour chacune des deux versions des RPC, on dispose de la BC correspondante. Chaque BC a été expansée en un ensemble de recommandations. Puis l'algorithme d'identification des

évolutions a été implémenté et les résultats ont été automatiquement produits par l'analyse comparative informatisée des BC correspondant aux deux versions. Le résultat d'une première comparaison quantitative des BC figure dans le tableau 1.

Dans la version 2004, de nombreuses variables de décision ont été introduites (+70 %). Il en a été de même pour les actions (+80 %). Certaines variables de décision et actions de 2002 ne sont en revanche plus utilisées en 2004. L'étape d'abstraction réduit cet effet, mais ne l'annule pas, indiquant ainsi l'émergence de nouveaux concepts.

Tableau 1 - Recommandations, variables de décision et actions en 2002 et 2004.

	Spécifiques de 2002	**Communes**	**Spécifiques de 2004**
Recommandations	526	51	1 030
Variables de décision	6	33	22
Actions	8	27	22

Lorsque l'on considère les deux arbres de décision construits, on trouve 366 situations cliniques (c'est à dire chemins) en 2002 et 584 en 2004. La longueur moyenne des chemins est de 10 en 2002 et de 11 en 2004. Quant aux recommandations, il y en a 577 en 2002, et 1 081 en 2004. Le tableau 2 donne la distribution des patrons d'évolution identifiés lors de la comparaison de $\{R_i\}_i^{2002}$ et $\{R_j\}_j^{2004}$ issues des 2 BC.

Tableau 2 - Distribution des patrons d'évolution entre les recommandations 2002 et 2004.

Patrons	**(n)**	**%**	
Recommandations obsolètes (2002)	206	36,0 %	*(Réf. 2002 ; total = 577)*
Identité	51	4,3 %	*(Réf. 2004 ; total = 1 081)*
Révision du plan d'actions	8	0,7 %	
Nouveau plan d'actions	39	3,6 %	
Révision de la situation clinique	80	7,4 %	
Révision de la recommandation	170	15,7 %	
Nouvelle pratique	339	31,4 %	
Nouvelle recommandation	394	36,4 %	

On constate que la version 2004 de la BC ne couvre strictement que très partiellement celle de 2002 avec seulement 51 recommandations identiques. Celles-ci représentent 8,8 % des recommandations de 2002 et 4,3 % de celle de 2004. Un tiers de la BC de 2002 est considérée obsolète. Un tiers de celle de 2004 est constituée de nouvelles recommandations, indiquant que les deux tiers représentent des transformations de recommandations de 2002.

5.2 Effet de la réactualisation sur les pratiques

Du point de vue pratique, nous avons analysé un échantillon tiré au sort de dossiers de patients atteints d'une tumeur de la vessie et suivis en 2004 dans le service d'urologie de

l'hôpital Bichat-Claude Bernard, à Paris. Pour chacune des 97 décisions thérapeutiques, nous avons comparé les traitements effectivement administrés aux patients à ceux préconisés par les recommandations de 2002 et de 2004. Il est à noter que l'analyse sur dossiers a été réalisée de façon rétrospective, les décisions de prise en charge consignées dans les dossiers ayant été réalisées indépendamment de l'utilisation du SADM UroDoc.

Nous avons observé des taux de conformité des pratiques cliniques de 67 % lorsqu'elles étaient comparées aux RPC de 2002, et de 46 % lorsque ces mêmes pratiques étaient comparées aux RPC de 2004 (tableau 3). Une analyse plus fine nous a permis de constater que la chute significative du taux de conformité des pratiques aux recommandations de 2004 ne provenait pas d'une « mauvaise » prise en charge des patients avec une tumeur invasive ou un envahissement ganglionnaire, non détecté en 2002 (puisqu'il n'y avait pas de recommandations pour ces profils) mais de la spécialisation d'un plan d'actions qui représentait 49 % des décisions, transformant une instillation endovésicale *optionnelle* en 2002, et donc conduisant à une pratique conforme même si elle n'était pas réalisée, en *recommandée* en 2004 et donc conduisant à une pratique non conforme dans le même cas.

Tableau 3 - Taux de conformité des pratiques vis à vis des RPC de 2002 et 2004.

n = 97	**Non conformes**	**Explicables**	**Taux de conformité**
2002	32	4	67 %
2004	52	5	46 %

6 Discussion et conclusion

Quand on considère l'évolution des RPC sur la prise en charge des tumeurs de la vessie entre 2002 et 2004 et qu'on analyse les deux ensembles de recommandations $\{R_i\}_i^{2002}$ et $\{R_j\}_j^{2004}$, ce qui frappe en premier lieu, c'est l'augmentation du nombre de situations cliniques pour lesquelles il existe des recommandations en 2004. Cette conclusion apparaissait déjà dès l'analyse comparative des textes au niveau du nombre de pages allouées aux différentes parties des documents où on s'apercevait que la prise en charge des tumeurs invasives et des tumeurs avec envahissement ganglionnaire était plus développée en 2004 qu'en 2002. Le document passe de 10 à 25 pages, hors annexes et bibliographie.

Du point de vue formel, en adoptant la méthode de comparaison proposée dans ce travail, le nombre de recommandations a pratiquement doublé, alors que seulement 8,8 % des recommandations de 2002 ont été réutilisées telles quelles en 2004. En tenant compte des dérivations possibles, c'est 54,0 % des recommandations de 2002 qui sont, directement ou indirectement réutilisées. S'il n'est pas surprenant que de nouvelles recommandations soient apparues en 2004, représentant 36,4 % de la BC, il est en revanche très étonnant que plus d'un tiers des recommandations de 2002 n'ait pas d'équivalent dérivé alors même que le guide 2004 est censé couvrir (et étendre) celui de 2002.

Les démarches de modélisation et de capitalisation des connaissances ont toujours considéré l'évolution des connaissances comme un réel problème. Des outils incluant des modules de gestion des versions fournissent des informations sur les évolutions des BC. Les classes de changement sont typiquement : « obsolète », « identique », « nouveau ». La présence de fonctions de similarité ou de suivi de l'édition permet d'ajouter une classe « modifié ». Ces outils s'appliquent aux ontologies, comme PromptDIFF [15, 19] ou à des

descriptions plus complètes comme des cas dans le contexte d'un SADM en oncologie [20]. Dans notre approche, la classe « modifié » se divise en 5 sous-classes pertinentes dans le contexte de l'évolution de RPC (plus précisément de règles) : « révision du plan d'actions », « nouveau plan d'actions », « révision de la situation clinique », « révision de la recommandation » et « nouvelle pratique ».

L'expérience relatée dans ce papier s'est focalisée à un niveau plus élevé de représentation, celui des RPC qui sont des règles. Toutefois, sans l'avoir étudié très en détail, il est certain que l'évolution des ontologies sous-jacentes à chaque RPC (qui couvrent le même domaine d'expertise) est en partie responsable du peu de recouvrement entre les versions 2002 et 2004, en plus de l'évolution des stratégies de prise en charge, les deux étant nécessairement liées. Les résultats obtenus nous conduisent à questionner deux notions souvent mobilisées : (i) la réutilisation en tant que méthode d'actualisation de BC, et (ii) la possibilité d'une interopérabilité sémantique et/ou ontologique dans une perspective chronologique.

Références

[1] Matillon Y & Durieux P. *L'évaluation médicale, du concept à la pratique.* Flammarion, Paris, 2000.

[2] Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. *JAMIA* 1999; 6(2): 104-14.

[3] Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyenne J, Sam J, Haynes RB. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. *JAMA* 2005; 293(10): 1223-38.

[4] Shiffman RN, Michel G, Essaihi A, Thornquist E. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc* 2004; 11(5): 418-26.

[5] Svatek V, Ruzicka M. Step-by-step formalisation of medical guideline content. *Int J Med Inf* 2003; 70(2-3): 329-35.

[6] Shahar Y, Young O, Shalom E, Galperin M, Mayaffit A, Moskovitch R & Hessing A. A framework for a distributed, hybrid, multiple-ontology clinical-guideline library, and automated guideline-support tools. J Biomed Inform 2004; 37(5): 325-44.

[7] National Institute for Health and Clinical Excellence. *The guidelines manual.* National Institute for Health and Clinical Excellence, London, April 2008. Available from: www.nice.org.uk.

[8] Judge J, Field TS, DeFlorio M, Laprino J, Auger J, Rochon P, Bates DW, Gurwitz JH. Prescribers' response to alerts during medication ordering in the long term care setting. *J Am Med Inform Assoc* 2006; 13(4): 385-90.

[9] van der Sijs H, Aarts J, Vulto A & Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006; 13(2): 138-47.

[10] Bouaud J, Séroussi B, Antoine EC. OncoDoc : une approche documentaire de l'aide à la décision. *Document Numérique* 1999; 3(3-4): 61-80. Numéro spécial Gestion des documents et gestion des connaissances.

[11] Séroussi B, Bouaud J, Antoine EC. OncoDoc, a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artif Intell Med* 2001; 22(1): 43-64.

[12] Séroussi B, Bouaud J, Dréau H, Falcoff H, Riou C, Joubert M, Simon C, Simon G, Venot A. ASTI, a guideline-based drug-ordering system for primary care. In: Patel VL, Rogers R, Haux R, eds, *Medinfo*, 2001; 528-32.

[13] Brizon A, Culty T, Séroussi B, Bouaud J, Ravery V. UroDoc : un système de mise en œuvre des recommandations pour la prise en charge des tumeurs urothéliales par navigation hypertextuelle. In : *Actes du 99e Congrès Français d'Urologie,* Paris, France. 16–19 novembre, 2005.

[14] Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developers' handbook. Scottish Intercollegiate Guidelines Network, Edimburg, January 2008. Available from: www.sign.ac.uk.

[15] Noy NF, Musen MA. Ontology versioning as an element of an ontology-evolution framework. *IEEE Intelligent Systems* 2003; 19(4): 6-13.

[16] Peleg M, Kantor R. Approaches for guideline versioning using GLIF. In: Musen M, ed, *Actes AMIA 2003*, Washington, DC. AMIA, novembre, 2003; p. 509-13.

[17] Seyfang A, Martinez-Salvador B, Serban R, Wittenberg J, Miksch S, Marcos M, ten Teije A, Rosenbrand K. Maintaining formal models of living guidelines efficiently. In: Bellazzi R, Abu-Hanna A, Hunter J, eds, *AIME 2007* (vol 4594) *Lecture Notes in Computer Science*, Amsterdam, The Netherlands. Springer, 2007; p. 441-3.

[18] Kaiser K, Miksch S. Versioning computer-interpretable guidelines: Semi-automatic modeling of 'living guidelines' using an information extraction method. *Artif Intell Med 2008*. À paraître.

[19] Noy NF & Musen MA. The PROMPT suite: Interactive tools for ontology merging and mapping. *Int J Hum Comput Studies* 2003; 59(6): 983-1024.

[20] d'Aquin M, Bouthier C, Brachais S, Lieber J, Napoli A. Knowledge editing and maintenance tools for a semantic portal in oncology. *Int J Hum Comput Studies* 2005; 62: 619–38.

Adresse de correspondance

Brigitte Séroussi
Département de Santé Publique
Hôpital Tenon
4, rue de la Chine
75970 Paris Cedex 20

E-mail : brigitte.seroussi@tnn.aphp.fr

Informatique et Santé, 2009 (17) : 203-212
© Springer – Verlag France

Gérer et améliorer les décisions des réunions de concertation pluridisciplinaires en oncologie : un cadre conceptuel utilisant des outils informatisés d'aide à la décision

Jean-Charles Dufour[ab], Fabrice Barlesi[c], Dominique Fieschi[a], Jean-Philippe Torre[b], Olivier Chinot[d], Marius Fieschi[ab]

[a] *LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille, France*
[b] *SSPIM, Hôpital de la Timone, AP-HM, Marseille, France*
[c] *Service d'Oncologie Thoracique, Fédération des Maladies Respiratoires, Hôpital Sainte-Marguerite, AP-HM, Marseille, France*
[d] *Unité de Neuro-Oncologie, Hôpital de la Timone, AP-HM, Marseille, France*

Abstract

Medical decision making and health care management is more and more a multidisciplinary approach especially in the field of oncology. Multidisciplinary meetings have to be supported in order to improve decisions taken and patients care. To meet this objective, supporting the various decisions making processes and facilitating the use of the best medical knowledge is essential. We have identified and explained six decision making processes to assist multidisciplinary committees in oncology. We are offering a conceptual framework aiming at pointing the means to be used for decision making in order to improve this process. Support for structuring and carefully presenting each patient's case, support for the selection of the most appropriate treatment using both computable guidelines and computable clinical trials is the core of our proposition.

Keywords

Professional Staff Committees; Decision Support Techniques; Clinical Practice Guidelines; Clinical Protocols; Clinical Trials; Medical Oncology.

1 Introduction

La prise de décision médicale tout comme les processus de prise en charge des patients sont de plus en plus des approches multidisciplinaires en particulier dans les cas complexes, s'agissant de pathologies malignes ou, plus généralement encore, lorsque le pronostic vital ou fonctionnel est une préoccupation de premier plan.

La prise en charge en oncologie est un exemple type de la multidisciplinarité nécessaire à l'optimisation des soins. Dans de nombreux pays, des comités de concertation multidisciplinaires en oncologie ont été mis en place afin de discuter des cas des patients, afin d'élaborer pour chacun d'eux un plan de soins optimal prenant en compte les connaissances et les techniques des différents spécialistes du domaine (chirurgiens, radio-oncologue, anatomopathologiste, chimiothérapeute…). En France, un plan national a été lancé pour imposer la consultation obligatoire de Comités de Concertation Pluridisciplinaires (CCP) chargés de préconiser la décision optimale concernant la mise en

œuvre du traitement oncologique adapté au cas de chaque patient. Ces comités ont recours à des Réunions de Concertation Pluridisciplinaires (RCP). Ils s'appuient sur des référentiels, des guides de bonnes pratiques cliniques validés, sur des protocoles standardisés afin de proposer aux patients des soins conformes aux meilleures pratiques. L'usage et le respect des guides de bonnes pratiques et des standards de prise en charge sont des critères de qualité explicitement mentionnés dans le plan cancer national français et doivent être régulièrement évalués par les CCP.

Il y a déjà une quinzaine d'années, Hammond *et al.* [1] avaient suggéré que, dans le domaine de la cancérologie tout particulièrement, le nombre et la diversité des guides de bonnes pratiques et des protocoles d'essais cliniques potentiellement utilisables, ainsi que la quantité de données cliniques qu'il est nécessaire de prendre en compte pour mettre en place un traitement adapté au cas de chaque patient pourraient bénéficier d'une aide apportée par des outils informatiques spécialisés [2-3]. De nos jours, parce que la masse de données cliniques et de connaissances est encore plus importante, l'usage de systèmes informatisés d'aide à la décision est obligatoire. L'amélioration du recrutement des patients dans les essais cliniques doit en bénéficier également. Les systèmes d'aide à la décision, basés sur les guides de bonnes pratiques informatisés notamment, peuvent améliorer l'efficience des CCP. Nous présentons dans cet article un cadre conceptuel ayant pour objectif de rendre doublement efficace ce dispositif en termes d'amélioration des prises en charge des patients et en termes de production et d'enrichissement des connaissances médicales utiles à cette amélioration.

2 Organisation des Comités de Concertation Pluridisciplinaires

Habituellement, plusieurs CCP interviennent dans un établissement de soins (généralement 2, il existe un CCP par domaine oncologique : neurologie, hématologie, gynécologie...). Ces CCP suivent des processus similaires et ont souvent une organisation et une planification des tâches comparables. L'organisation typique d'un CCP en oncologie a été analysée et détaillée par Castel *et al.* [4]. Les principes opérationnels de fonctionnement des CCP et le *processus*[1] des cas de patients sur lesquels ils doivent se prononcer peuvent se résumer et se schématiser ainsi :

- tous les cas de patients atteints de cancer pour lesquels la mise en route ou la modification d'un traitement oncologique est requise doivent être enregistrés par le secrétariat du CCP ;
- le CCP organise régulièrement des RCP à l'issue desquelles, pour chaque cas analysé, un avis concernant la conduite à tenir thérapeutique est donné ;
- chaque cas enregistré donne lieu à la constitution d'un document de synthèse clinique qui sert de base d'information lors des RCP pour déterminer la conduite à tenir. Il convient de préciser que les cas répondant à une procédure standard de prise en charge, ayant fait l'objet d'un accord pluridisciplinaire, traduite par un référentiel de pratique clinique validé et actualisé, ne sont pas obligatoirement présentés et discutés lors des RCP. En effet, dans certains CCP, un expert coordinateur en relation directe avec les membres du comité pluridisciplinaire intervient pour proposer une conduite à tenir adaptée et conforme aux référentiels approuvés par le CCP. L'intervention de l'expert permet ainsi de diminuer la charge de travail lors des RCP qui peuvent alors consacrer plus de temps aux cas complexes et/ou atypiques ;

[1] En anglais, le *workflow*.

- tous les cas (qu'ils soient discutés lors des RCP ou résolus par l'expert) font l'objet de préconisations thérapeutiques écrites et étayées par des référentiels de pratique et/ou des arguments précisant les motifs de non-conformité aux référentiels existants ;
- enfin, les préconisations thérapeutiques des RCP sont consignées et le médecin traitant est informé de la conduite à tenir dans le cas du patient qu'il prend en charge.

Les CCP doivent organiser et contrôler la qualité de l'ensemble du processus de décision et vérifier l'application effective des avis thérapeutiques qu'ils ont donnés. L'utilisation et l'étude de la conformité des décisions avec les référentiels de bonnes pratiques et la compliance des soins véritablement dispensés sont donc des points essentiels. La mise en place de procédures de contrôle, de suivi et d'amélioration est cruciale pour vérifier l'adéquation des avis donnés avec les guides de bonnes pratiques approuvés par la CCP. Le taux d'adéquation des avis avec les standards de pratique est d'ailleurs un des indicateurs qualité du rapport d'activité règlementairement demandé aux CCP.

Cette description du *processus* typique d'un CCP doit être complétée lorsqu'il s'agit de CCP impliqués dans des essais thérapeutiques. En effet, dans certains établissements de soins, les CCP veillent également à proposer aux patients des alternatives aux traitements standards en leur indiquant l'existence d'un essai thérapeutique potentiellement prometteur dont ils pourraient bénéficier. L'amélioration de l'inclusion des patients dans les essais thérapeutiques en cours est une préoccupation importante qui complète l'offre de soin et les alternatives standards qui peuvent être proposées aux patients.

3 Cadre conceptuel : objectifs et caractéristiques

L'analyse du *workflow* des CCP permet de dégager plusieurs processus dans lesquels la mise en œuvre de systèmes d'aide à la décision peut être efficace pour optimiser l'organisation et améliorer les prises en charges des patients (figure 1) :

1. **aide à la structuration du document de synthèse présentant le cas du patient**. Le contenu de ce document est important pour le processus de décision puisque c'est principalement sur la base des informations qu'il contient que les choix thérapeutiques seront établis. Il doit donc contenir les critères décisionnels pertinents et rigoureusement renseignés afin de permettre la prise de décision. Celle-ci est centrée sur des critères explicites [5]. La nature de ces critères varie évidemment en fonction du type de la tumeur (volume, ganglions, récepteurs hormonaux, délétion chromosomiques, etc.), mais un certain nombre d'items communs doivent figurer dans tous les documents de ce type (recommandation de l'INCa). Dans le cadre de prises de décisions médicales collaboratives, Christensen *et al.* [6] ont montré l'importance du formatage et de la présentation des informations objectivant le problème à résoudre dans un tel document. Classiquement, ce document prend la forme d'une fiche papier ou électronique qui est complétée en amont des RCP. L'objectif principal d'un système d'aide à la structuration du cas du patient est de colliger, de présenter de manière adaptée et de permettre le partage entre tous les membres de la réunion des informations pertinentes sur le patient en vue de prendre une décision éclairée. Les informations pertinentes décrivant le cas du patient doivent être issues : 1) des critères décisionnels présents dans les guides de bonnes pratiques et les protocoles standardisés ; 2) des cas déjà résolus ; et 3) des critères d'inclusions/exclusions des essais thérapeutiques en cours.

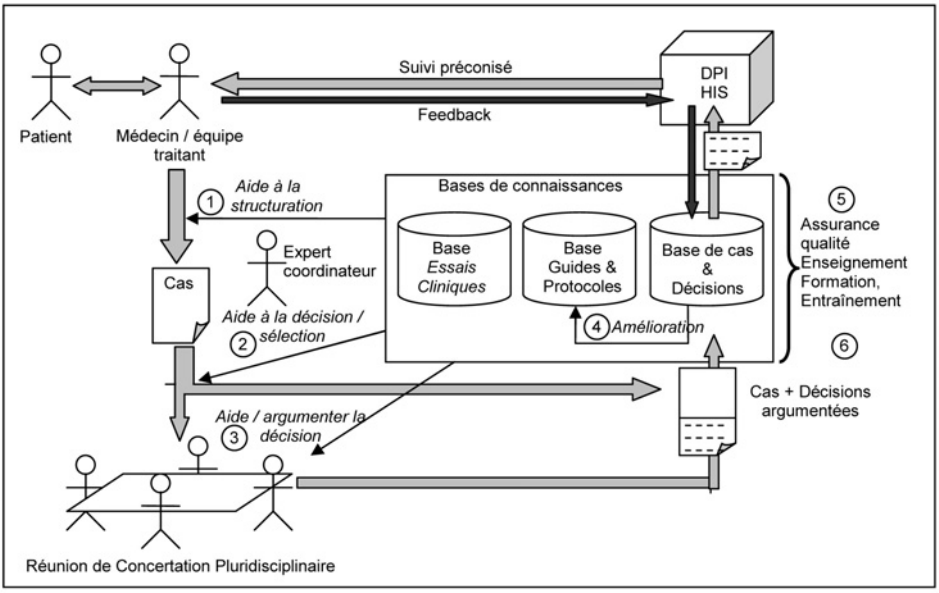


Figure 1 - Organisation et processus des CCP – Place des systèmes d'aide à la décision.

2. **aide à la sélection des cas qui doivent impérativement être discutés lors des RCP**. Pour cela il est nécessaire de mettre en place un système d'aide à la décision basé sur les guides de bonnes pratiques approuvés par le CCP. L'objectif principal de ce système est de proposer des recommandations applicables au cas du patient. Par exemple, lorsqu'un tel système propose une conduite à tenir standard sans alternative possible, l'expert peut alors entériner la proposition du système et décider de ne pas soumettre le cas à la RCP pour discussion.
3. **aide pour prendre et argumenter les décisions lors des RCP**. Un tel système doit utiliser l'ensemble des bases de connaissances disponibles afin de proposer des alternatives thérapeutiques issues : 1) des guides de bonnes pratiques et des protocoles standards ; 2) des essais cliniques dans lesquels le patient peut éventuellement être inclus ; 3) de la base de cas afin de rechercher des cas similaires pour lesquels une réponse thérapeutique favorable a été constatée. L'utilisation d'une base de cas suppose que les informations appropriées concernant l'application effective de la thérapeutique préconisée par la CCP et son effet chez le patient soient disponibles et colligés. Rossille *et al.* [7] ont proposé un système ayant pour objectif d'apporter une aide lors des RCP pour résoudre les cas « hors standard » (cas pour lesquels il n'existe pas de guide de bonnes pratiques adapté). Techniquement, ce système utilise à la fois un dispositif à base de règles d'inférences et un mécanisme de raisonnement à base de cas. Un tel système pourrait être amélioré en intégrant une base d'essais cliniques informatisés afin de proposer pour chaque cas de patient analysé une sélection des essais dans lesquels le patient pourrait être inclus. Il faut noter que l'utilisation d'un système d'aide à la décision complet aurait également pour intérêt de servir de support aux discussions lors des RCP et permettrait ainsi d'étayer et d'argumenter la décision sur la base des propositions du système.
4. **amélioration des guides de bonnes pratiques et adaptation locale**. Sauvagnac *et al.* [4] ont montré que les réunions multidisciplinaires sont particulièrement indiquées pour la promotion et l'adaptation des protocoles et des guides de bonnes pratiques dans

le but d'en faire des outils pragmatiques, facilement applicables et tenant compte des particularité locales (équipement médicaux disponibles, contexte social et culturel, etc.). À cette fin, nous devons garder à l'esprit que ces référentiels de soins doivent être explicitement exprimés de manière à être clairement compréhensibles et sans ambiguïté pour tous les spécialistes. L'informatisation des guides de bonnes pratiques est une solution qui amène une explicitation rigoureuse des connaissances que contiennent ces guides. Les guides de bonnes pratiques deviennent opérationnels et peuvent être vérifiés et renforcés. De même l'utilisation des critères de décision peut être précisée lors de leur informatisation et de leur utilisation dans les RCP. Ainsi, un cercle vertueux peut se constituer : les guides de bonnes pratiques améliorent les décisions de comités multidisciplinaires ; les comités multidisciplinaires contribuent à améliorer la qualité des guides de bonnes pratiques.

L'amélioration des guides de bonnes pratiques et leurs éventuelles spécialisations/adaptations locales doivent utiliser les connaissances issues des décisions des RCP et les informations apportées par le suivi du patient (efficacité du traitement, disponibilité des thérapeutiques, effets indésirables, qualité et durée de vie...). Ainsi, l'analyse des cas pour lesquels une conduite à tenir conforme aux recommandations d'un guide de bonnes pratiques a été retenue lors de la RCP d'une part et l'analyse des actions (et de leurs effets) qui ont été réellement mises en œuvre pour ces patients d'autre part, sont extrêmement utiles pour améliorer le fonctionnement des structures qui soignent ces patients et juger de l'adaptation des guides et des recommandations à la réalité clinique.

5. **facilitation du contrôle qualité**. Il est indispensable de disposer d'outils adaptés pour analyser de façon constructive les cas pour lesquels le comité a opté pour une décision en dehors des standards. La constitution de bases de cas, de tables de décisions et de guides de bonnes pratiques informatisés facilite le contrôle qualité des décisions. En effet, un système basé sur des guides de bonnes pratiques informatisés peut être utilisé à la fois: 1) de manière prospective afin d'aider l'expert ou le comité dans la recherche de la solution standard la plus adaptée au cas du patient ; 2) de manière rétrospective afin de repérer puis d'analyser les cas pour lesquels une solution non standard a été recommandée par le CCP.
6. **enseignement et formation**. La constitution et la mise à disposition des bases de connaissances peuvent permettre l'élaboration d'outils de formation et d'entraînement pour les médecins. En effet, ces bases peuvent être utilisées pour alimenter des environnements d'apprentissages et permettre aux praticiens en formation d'examiner les règles de décision, de vérifier leur domaine d'application et d'en cerner les limites.

4 Implémentations pilotes

Nous avons mis en œuvre certaines des fonctionnalités décrites précédemment dans des CCP de notre établissement. Ces implémentations pilotes sont basées sur la réutilisation et la spécialisation de notre plate-forme de développement, de diffusion et d'utilisation de guides de bonnes pratiques dans un processus de décision, décrite par ailleurs [8].

4.1 Informatisation de guides de bonnes pratiques (domaine de la neuro-oncologie)

Nous avons travaillé avec le CCP de neuro-oncologie afin de mettre en place un système informatisé utilisant des guides de bonnes pratiques « exécutables » (c'est-à-dire inférés automatiquement en fonction des données du patient) pour faciliter la recherche des traitements éprouvés applicables au cas de chaque patient. Dans ce CCP, les cas qui

doivent être examinés lors des RCP sont saisis via un formulaire informatisé et sont enregistrés dans une base de données dans le système d'information hospitalier (SIH).

Nous avons développé un prototype focalisé sur le traitement des gliomes. Des guides de bonnes pratiques (publiés sous forme textuelle par une des organisations faisant autorité dans le domaine de la neuro-oncologie [9]) ont été informatisés en utilisant l'éditeur spécialisé de notre plate-forme de développement. Un composant logiciel a été élaboré pour coupler la base de données du SIH et les guides de bonnes pratiques informatisés. Ce composant est basé sur les fonctionnalités du XSLT qui permettent de transformer les données extraites du SIH en un document XML ayant la structure attendue pour être correctement interprété par le moteur d'inférence de notre plate-forme. Le résultat produit par le système est, pour chaque cas, un document synthétique et structuré sur lequel les données du patient sont résumées sur la partie gauche et les recommandations adaptées au profil du patient sont transcrites sur le partie droite (figure 2). Ce document est destiné à être imprimé afin d'être utilisé par l'expert du CCP et par le groupe multidisciplinaire lors de la RCP pour faciliter la prise de décision.

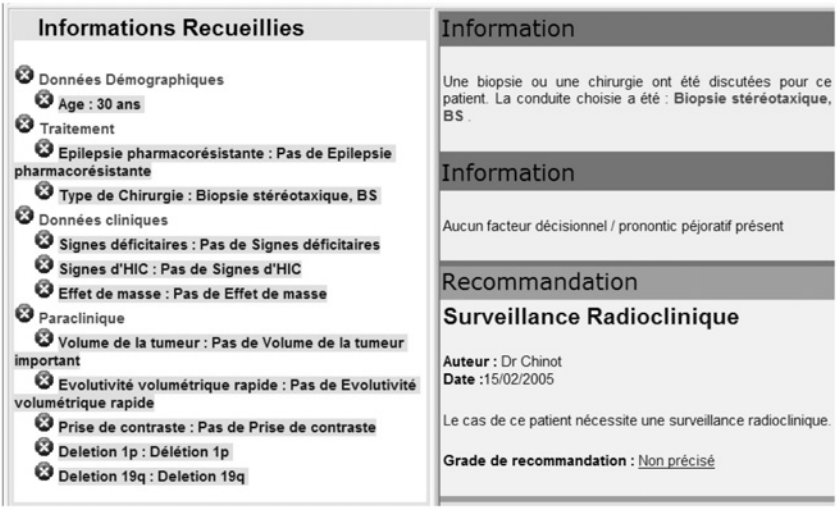


Figure 2 - Résultat produit par le système d'inférence automatisée des guides de bonnes pratiques.

L'implémentation pilote a porté sur quatre guides de bonnes pratiques concernant la prise en charge des différents grades du glioblastome.

4.2 Essais cliniques (domaine de l'oncologie thoracique)

Le système a été développé en collaboration avec le CCP d'oncologie thoracique. Il a pour objectif d'améliorer le taux d'inclusion des patients dans les essais thérapeutiques en cours.

Pour atteindre cet objectif, nous avons réalisé des développements additionnels afin d'enrichir et de modifier les fonctionnalités de notre système d'aide à la décision et d'en faire un outil de gestion et d'alerte concernant les essais thérapeutiques en cours pour lesquels les patients peuvent être éligibles. Cela revient à déterminer les critères d'inclusion et d'exclusion des différents essais thérapeutiques et à les traduire dans le formalisme que nous utilisons dans notre système d'aide à la décision. Les critères d'inclusion correspondent aux « *rule-in* » et les critères d'exclusion aux « *rule-out* » d'un

système classique de guide de bonnes pratiques exécutables. Là encore nous utilisons un composant logiciel pour coupler les essais thérapeutiques informatisés avec la base de données patient utilisée par le CCP (il s'agit d'une base de données au format MS Access®). Cette application, installée sur un ordinateur portable, est destinée à être utilisée en temps réel lors des RCP. L'écran présentant les données du patient et les protocoles dans lesquels le patient pourrait être inclus est projeté aux participants de la RCP (figure 3). Pour chaque patient, le système mentionne s'il peut être inclus ou pas dans les essais thérapeutiques (en haut de l'écran), le nombre et les motifs de non-inclusion dans chaque essai (en bas et à droite de l'écran), et les principales données du patient (en bas et à gauche de l'écran).

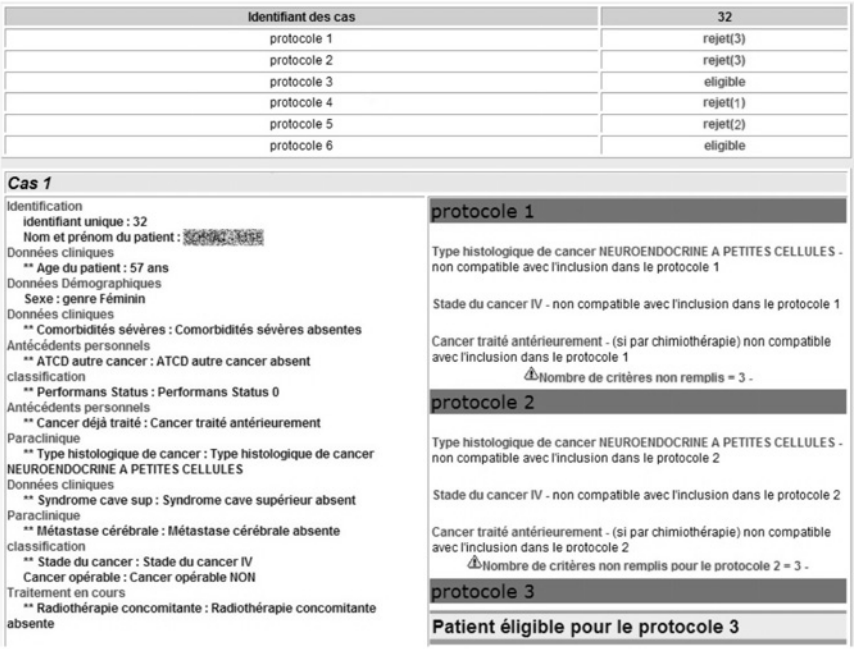


Figure 3 - Copie d'écran du système automatisé d'aide à la sélection des essais cliniques.

5 Discussion

Dans le domaine de l'oncologie, les CCP fournissent un cadre organisationnel et permettent de formaliser en partie les prises de décision. Leur objectif ultime est d'améliorer la qualité des soins et de réduire les disparités injustifiables dans les prises en charge. Pour atteindre cet objectif, il est indispensable de soutenir au mieux les différents processus décisionnels et de faciliter l'utilisation des connaissances validées. Nous proposons un cadre conceptuel afin de caractériser les moyens d'aide à la décision utilisables pour améliorer le processus décisionnel des CCP. L'aide à la structuration et à la présentation minutieuse du cas du patient, l'aide à la sélection du traitement le plus approprié en utilisant à la fois des guides de bonnes pratiques informatisés et la détection automatique des essais cliniques en cours pour lesquels le patient est éligible sont au cœur de notre proposition.

L'utilisation de systèmes d'aide à la décision basés sur des guides de bonnes pratiques et/ou des bases de cas validés dans les CCP permet d'éviter des dérives dans le choix des options thérapeutiques. Lors des RCP, les guides de bonnes pratiques peuvent ainsi jouer le rôle de garde-fous. Ils permettent le questionnement qui élimine une part de la subjectivité dans le choix des traitements conformes aux données de la science et adaptés aux patients. Par exemple, l'utilisation de guide de bonnes pratiques peut contribuer à relativiser des avis et des jugements d'experts qui pèsent parfois plus qu'il ne faudrait sur la décision finale. En effet, le poids du savoir d'un expert lors d'une RCP peut être délétère s'il n'est justifié que par une autorité empreinte de règles propres à cet expert [5]. Il faut également éviter que le jeu social, toujours présent dans les débats lors de RCP, devienne prééminent et que la discussion ne devienne qu'une « bataille d'experts » s'éloignant des connaissances standards, validées et partagées [10]. À ce titre, cet outil peut contribuer à éviter ces dérives en recentrant la discussion sur les données tangibles, validées et explicites. Finalement, l'utilisation plus rigoureuse des données issues de la littérature et des guides de bonnes pratiques doit être facilitée car elle contribue à renforcer et à étayer les décisions prises lors des RCP en les basant sur les données actuelles et validées [11].

Un autre argument justifie la mise en œuvre d'un tel outil dans les CCP : la nécessité d'optimiser le processus de gestion des cas de patients pour lesquels un avis est requis. Bien que cette utilisation n'ait pas fait l'objet d'une évaluation quantifiée, suite à un usage suffisamment long et intensif de notre outil, c'est ce qu'il ressort de cette expérience. Il est en effet crucial d'optimiser les temps de traitement des dossiers tout particulièrement dans le domaine de l'oncologie où tout retard de prise en charge thérapeutique est synonyme de perte de chance pour le patient. Il y a des cas pour lesquels la discussion lors de la RCP est pleinement justifiée car aucun référentiel de pratique n'est rigoureusement applicable ou bien le désir d'un patient peut amener à revoir une attitude protocolaire, en pesant les risques engendrés par une dérogation à la conduite recommandée. Mais de nombreux cas, soumis à discussion lors de la RCP, relèvent de protocoles standards [5]. Ils peuvent être identifiés, comme nous le proposons, par un système d'aide à la décision basé sur l'informatisation de ces protocoles. Il s'agit donc de repérer les dossiers standards pour accélérer leur traitement, et pour consacrer plus de temps aux dossiers complexes.

L'utilisation des guides de bonnes pratiques informatisés dans le cadre des CCP est également bénéfique pour l'évolution et l'adaptation des guidelines eux-mêmes. En effet, les guidelines nationaux ou internationaux sont rarement « prêts à l'emploi » et nécessitent un travail des médecins pour les rendre plus facilement utilisables (leur informatisation aide à ce travail), pour les actualiser en fonction des dernières publications ou pour préciser une stratégie. Ils doivent bien souvent être adaptés localement avant d'être opérationnels. Les RCP sont propices à l'amélioration des guides de bonnes pratiques car elles peuvent les faire évoluer non plus simplement à partir des nouvelles connaissances disponibles, mais aussi à partir des usages effectifs qui en sont faits [5]. Pour que cette amélioration soit possible, le simple enregistrement de la décision prise à l'issue des RCP ne suffit pas. Il convient d'enregistrer et d'analyser les caractéristiques des cas traités et surtout les éléments qui permettent de comprendre les raisons de la décision. Ainsi, la trace des délibérations doit être riche et ne peut se limiter à la décision finalement retenue, elle doit inclure ses justifications, les autres hypothèses envisagées, les raisons ayant conduit à les éliminer, les savoirs non partagés par tous (c'est-à-dire surspécialisés) qui ont été portés à la connaissance de tous à l'occasion de la discussion. Tous ces éléments doivent enrichir la base de cas si l'on souhaite par son analyse mettre à jour de nouvelles connaissances qui viendront enrichir les standards de prise en charge. Il s'agit de mettre en place un cercle vertueux. Cela suppose de formaliser les pratiques des CCP et de disposer d'outils adaptés

qui bénéficieront à la fois à la prise de décision de ces comités et à la construction de la connaissance.

Références

[1] Hammond P, Harris A, Das SK, and Wyatt JC. Safety and decision support in oncology. *Methods Inf Med* 1994; 33(4): 371-81.

[2] Dufour JC, Bouvenot J, Ambrosic P, Fieschi D, Fieschi M. Textual Guidelines versus Computable Guidelines: A Comparative Study in the Framework of the PRESGUID Project in Order to Appreciate the Impact of Guideline Format on Physician Compliance. *Proc AMIA* 2006 ; 219-23

[3] Séroussi B, Bouaud J, Gligorov J, Uzan S. Gestion de réunions de concertation pluridisciplinaire et prise en charge thérapeutique du cancer du sein: impact sur la qualité des décisions. *Actes des XIIe Journées Francophones d'Informatique Médicale*. 2007. Bamako, Mali.

[4] Castel P, Blay JY, Meeus P, Sunyach MP, Ranchere-Vince D, Thiesse P, Bergeron C, Marec-Berard P, Lurkin A, Ray-Coquard I. Organization and impact of the multidisciplinary committee in oncology. *Bull Cancer* 2004; 91(10): 799-804.

[5] Sauvagnac C, Beckendorf V, Lesur A, Luporsi E, Stines, J, Falzon, Bey P. Decision committees in oncology: from standard to actual cases. *Bull Cancer* 1999; 86(9): 767-72.

[6] Christensen C, Larson JR. Collaborative medical decision making. *Med Decis Making* 1993; 13(4): 339-46.

[7] Rossille D, Laurent JF, Burgun A. Modelling a decision-support system for oncology using rule-based and case-based reasoning methodologies. *Int J Med Inform* 2005; 74(2-4): 299-306.

[8] Dufour JC, Fieschi M, Fieschi D, Giorgi R, Gouvernet J. A platform to develop and to improve effectiveness of online computable guidelines. *Stud Health Technol Inform* 2003; 95: 800-5.

[9] Standards Options and Recommendations (SOR) on the Fédération Nationale de Lutte Contre le Cancer (FNCLCC) web site : http://www.fnclcc.fr

[10] Castel P, Merle I. [When standards become a resource for doctors: The case of oncology]. *Sociol Trav* 2002; (44): 337-55.

[11] Séroussi B, Bouaud J, Gligorov J, Uzan S. Supporting multidisciplinary staff meetings for guideline-based breast cancer management: a study with OncoDoc2. *Proc AMIA* 2007 ; 11: 656-60.

Adresse de correspondance

Dr Jean-Charles Dufour
LERTIM - Université de la Méditerranée - Faculté de Médecine –
27, boulevard Jean Moulin 1
3385 Marseille Cedex 5, France
E-mail : jean-charles.dufour@univmed.fr
Site Web : http://lertim.org

Informatique et Santé, 2009 (17) : 213-224
© Springer – Verlag France

Mise en œuvre d'un processus de conception objet pour l'implémentation de règles de décisions destinées à la validation pharmaceutique

Abdelali Boussadi [a], Cédric Bousquet [b], Patrice Degoulet [a]

[a] *Laboratoire de recherche en ingénierie des connaissances e-santé, (INSERM UMR_S 872, Eq 20), Université Paris Descartes, Faculté de médecine, Paris, France*
[b] *Université de Saint Etienne, Département de Santé Publique et Information Médicale, Saint-Étienne, France*

Abstract

The most important part of alarm systems is the rules set that allows triggering alerts. We propose a generic rule design framework which can support the design and the implementation of alerts for drug prescriptions or their pharmaceutical validation. Our approach takes into account two important criteria, which represent two real challenges for the system designers. First, the developer needs to identify and model both users' requirements and business processes in which the alarm system works. Second, the developer needs to model the knowledge associated with the decision rules with an appropriate language. We show how the Unified Process helps to model the business process for pharmaceutical validation. Knowledge representation benefited from use of SBVR, a language dedicated to rule representation within an object oriented framework. After a successful implementation of a prototype we plan to integrate such system within the hospital information system of the Georges Pompidou hospital.

Keywords

Software Engineering; Business Rules; Unified Process; SBVR; Decision Making; Computer-Assisted; Decision Support Systems; Expert Systems; Alert Systems; Medical Order Entry Systems/organization & administration; Medical Records Systems; Medication Errors/prevention & control; Pharmacy Service.

1 Introduction

De nombreuses publications évaluant les erreurs de prescription justifient l'utilisation de systèmes d'aide à la décision au moment de la prescription [1-3]. L'efficacité des systèmes d'aide à la décision dépend de plusieurs facteurs et notamment de la participation du médecin prescripteur. Van der Sijs *et al.* [3] montrent que les médecins prescripteurs passent outre les alertes des systèmes d'aides à la décision dans 49 % à 96 % des cas. La présence d'un autre garde-fou est donc nécessaire.

Dans une revue clinique par Moyen *et al.* [4], les auteurs citent les principales interventions qui peuvent améliorer la sécurité du circuit du médicament. Dans cette étude [5], les systèmes d'aide à la décision sont, entre autres, associés à l'intervention des

pharmaciens valideurs. En France et depuis 1991, la réglementation française a rendu obligatoire l'analyse et la validation des prescriptions médicamenteuses par les pharmaciens dans les hôpitaux [5, 6]. Les pharmaciens sont efficaces dans la prévention des événements médicamenteux indésirables [7], la prévention des erreurs induites par les systèmes de prescriptions informatisées [8] et la réduction des coûts induits par ces erreurs [9].

À l'Hôpital Européen Georges Pompidou (HEGP), la prescription est informatisée à l'aide du système « Dx-C@re® » [10]. L'ensemble des prescriptions d'examen de biologie ou de radiologie est informatisé mais ce n'est pas encore le cas pour les médicaments. L'HEGP dispose d'un système de contrôle des prescriptions avec déclenchement d'alertes pour les médecins mais ce système ne dispose pas de toutes les fonctionnalités nécessaires à une sécurisation complète des prescriptions. Actuellement, 300 lits sont informatisés pour la prescription des médicaments et l'objectif est d'atteindre 650 lits en 2009. Afin de conserver une sécurité maximale, il n'est pas envisagé de prescription informatisée sans validation pharmaceutique. L'étape de validation pharmaceutique devient donc un facteur limitant de la montée en charge, puisqu'il n'est pas envisagé d'augmenter le nombre de pharmaciens pour cette tâche (*i.e.* quatre pharmaciens à temps plein). L'introduction d'un système d'aide à la décision (SAD) pour l'aide à la validation pharmaceutique semble être la solution idéale pour faciliter le travail des pharmaciens et de fait améliorer la sécurité du circuit du médicament.

La majorité des études publiées sur les systèmes d'aide à la décision concernent des associations entre des systèmes d'aide à la décision et des systèmes de prescriptions informatisées. DoseChecker [11] est l'un des rares systèmes experts introduit à l'hôpital coté pharmacien pour le contrôle des dosages de prescriptions de médicaments. Dans cette étude et dans la majorité des études publiées sur les systèmes d'aide à la décision, les auteurs se concentrent plus sur l'évaluation de l'impact du système et ne donnent que très peu d'informations sur la méthodologie qui a conduit à la conception et la mise en œuvre du système.

Les processus de conception et d'implémentation de systèmes informatiques sont à la fois des processus sociaux et techniques [12]. Les concepteurs et les développeurs du système doivent pouvoir comprendre et modéliser les processus métier (aspect social, notions de workflows) et les relier aux aspects techniques de la conception (notion de gestion de processus de développement) [12, 13]. Le système implémenté doit ainsi pouvoir répondre aux besoins de l'utilisateur final mais ce dernier doit aussi être prêt à s'adapter aux changements métier initiés par l'introduction du système dans l'organisation [12]. Dans un système d'aide à la décision à base de règles, les règles métier qui déclenchent les alertes sont le composant le plus important du système. Nous proposons un processus générique de conception de système d'alerte permettant d'encadrer la spécification des règles métier implémentées au sein de ce type de système. Notre approche prend en compte deux critères importants qui représentent deux réels défis pour les concepteurs de systèmes à base de règles [12, 13].

Premièrement, les concepteurs doivent identifier et modéliser les besoins utilisateurs et les processus métier de l'activité métier pour laquelle le système d'alerte va être mis en œuvre. Il existe de nombreux formalismes pour la modélisation des besoins utilisateurs et pour la modélisation des processus métier. Le plus répandu est UML (*Unified Modeling Language*). L'utilisation d'UML pour la conception de systèmes informatiques a déjà fait l'objet de nombreuses publications [14, 15]. Dans ces études, il n'est pas spécifié si l'utilisation d'UML était encadrée par un processus de conception qui couvre tout le cycle de vie d'un système informatique (de l'étude de faisabilité à la maintenance). Les auteurs

d'UML insistent sur le fait d'encadrer l'utilisation d'UML comme langage de conception par un processus de conception et recommandent l'utilisation du Processus Unifié (*Unified Process*, UP) [16].

Deuxièmement, les concepteurs doivent modéliser les connaissances nécessaires à l'écriture des règles métier avec un formalisme approprié. Plusieurs formalismes ont été proposés pour la modélisation des guides de bonnes pratiques cliniques (GBPC) [17]. Ces formalismes ne sont pas adaptés au contexte de notre étude car ils sont dédiés à la modélisation de connaissances provenant des guides de bonnes pratiques cliniques. Leur mise en œuvre dans des systèmes d'alerte est susceptible d'entraîner une complexité inutile lors de la modélisation des règles métier. La syntaxe Arden présente de nombreux inconvénients. Le modèle de données n'est pas orienté objet, il est atomique et ne permet pas l'expression de relations temporelles [18]. GELLO (*object-oriented query and expression language*) comble les lacunes de la syntaxe Arden, cependant il est basé sur un langage destiné aux informaticiens, ce langage présente une grande complexité pour l'utilisateur final dans l'écriture et la lecture des règles métier produites [18]. Idéalement, le formalisme choisi doit s'articuler avec le processus de conception utilisé et doit permettre d'écrire des règles métier faciles à lire et à écrire pour l'utilisateur final. Ceci permettra d'incorporer l'utilisateur final dans le cycle de conception du système. La participation de l'utilisateur final au processus de conception permet de concevoir un système qui répondra mieux aux besoins de l'utilisateur et permettra à l'utilisateur final de comprendre comment l'introduction du système changera les processus métier de son activité. C'est le cas du formalisme SBVR : *Semantics of Business Vocabulary and Business Rules* [19].

Notre hypothèse de recherche s'articule autour de deux points principaux : (1) le Processus Unifié et la notation UML sont adaptables pour permettre la spécification de règles métier à partir de la modélisation des processus métier concernés par l'introduction du système, (2) SBVR permet la modélisation des connaissances nécessaires pour l'écriture des règles métier. Nous avons évalué l'efficacité de notre méthode en implémentant les règles produites sous forme d'alerte gérées par un système de gestion de règles métier (*Business Rule Management System*, BRMS) issu du commerce, en l'occurrence : ILOG JRules® et nous avons évalué la valeur ajoutée de la combinaison Processus Unifié, UML, SBVR dans le contexte applicatif de la pharmacie centrale de l'HEGP.

2 Matériel et méthode

Notre méthode s'articule autour des différentes phases et activités du Processus Unifié. Nous proposons et justifions une adaptation du Processus Unifié à notre contexte applicatif. Le Processus Unifié est un processus de développement logiciel qui encadre la conception et le développement de systèmes informatiques. Le Processus Unifié gère le cycle de vie de conception suivant différentes phases : création, élaboration, construction et transition. Chaque phase est composée d'un ensemble d'activités (expression et analyse des besoins, conception...) dont l'importance dépend du contexte applicatif et de la phase dans laquelle le projet se situe. Nous décrivons dans ce qui suit uniquement les adaptations que nous avons apportées aux phases d'élaboration et construction.

2.1 Phase d'élaboration : expression et analyse des besoins

Le principal objectif de l'expression des besoins est « l'élaboration d'un modèle du système à construire » [16]. L'emploi des cas d'utilisation système constitue un excellent moyen de procéder à la création de ce modèle, car les besoins fonctionnels sont

naturellement structurés comme des cas d'utilisation. La spécification des règles métier doit être associée à l'activité d'expression des besoins [20, 21]. Les auteurs du Processus Unifié ne précisent aucune étape dans l'activité d'expression des besoins pour spécifier les règles métier. Nous préconisons donc d'introduire la spécification des règles métier à ce niveau du processus de développement *i.e.* à la phase d'élaboration et à l'activité d'expression des besoins.

La dernière étape de l'activité de formulation des besoins de la phase d'élaboration est la structuration du modèle de cas d'utilisation système. Cette activité permet de faire ressortir les relations d'extension « Extend » et d'inclusion « Include » ainsi que des relations de généralisation entre cas d'utilisation. Cette activité permet aussi de mettre en évidence des cas d'utilisation dits « abstraits ». Ces cas d'utilisation prennent en charge des enchaînements d'actions qui se répètent dans d'autres cas d'utilisation dits « concret ». L'objectif étant de réorganiser le modèle de cas d'utilisation pour qu'il n'y ait pas de répétition dans les descriptions. Les enchaînements d'activités suivants permettent de mettre en lumière les premières ébauches de règles métier et ce en partant de la spécification des cas d'utilisation système et métier faite dans les activités précédentes.

La spécification des règles métier et leur introduction dans le cycle de développement ne sont pas décrites dans le cycle de conception du Processus Unifié. Il est donc nécessaire d'introduire une nouvelle étape qui complète l'activité de formulation des besoins dans la phase d'élaboration [20, 21]. Cette étape commence par identifier quelles sont les règles métier qui auront le plus d'impact sur le travail des pharmaciens dans le contexte de la validation pharmaceutique.

Pour cela, nous avons interrogé la base de données de Dx-C@re® afin de connaître la proportion des médicaments concernés par les alertes par rapport au nombre total de lignes de prescription et ce pour chaque département concerné par la validation pharmaceutique. Après avoir déterminé les règles métier les plus importantes dans l'activité de validation pharmaceutique, nous avons établi des ébauches de règles métier appelées modèle de règles écrit désignées par BRx (*Business Rule numéro x*) [20].

Ces modèles de règles écrits doivent orienter la description des cas d'utilisation qui leur correspond. Suivant ces modèles de règles, nous avons rédigé des règles du type « si / alors » sous forme de tables de décision que nous avons validées avec les pharmaciens. Ceci a permis de lister les molécules, les marqueurs biologiques et les situations cliniques pour lesquels ces règles déclencheront des alertes. L'adaptation du Processus Unifié par l'introduction de cette activité va se répercuter sur toutes les activités des différentes phases du Processus Unifié qui suivront cette activité.

2.2 Phase de construction : conception, implémentation et tests

Le passage du projet de la phase d'élaboration à la phase de construction marque un changement d'optique : « les phases de création et d'élaboration regroupent des activités de recherche, la phase de construction s'apparente, elle, au travail de développement. » [16]. À ce stade, les besoins des utilisateurs et l'architecture du système sont stables. On commence par identifier les nœuds et les configurations réseaux de l'architecture de notre système et on les représente en utilisant le diagramme de déploiement d'UML.

Plusieurs solutions sont possibles pour relier nos règles métier aux modèles UML au cours de l'activité de conception pour garantir une homogénéité des documents de spécification ainsi que pour avoir une base formelle pour l'expression de nos règles. Dans notre cas, nous avons décidé de respecter la « traçabilité » avec les cas d'utilisation précédemment identifiés. Nous avons donc associé pour chaque cas d'utilisation une « réalisation – conception de cas d'utilisation ». Cela passe par l'identification de classes de conception

formalisées par le diagramme de classes d'UML. Le diagramme de classes de conception associé aux modèles de règles écrits est dit : modèle objet métier (*Business Object Model*, BOM).

Avant de pouvoir instancier nos règles métier à partir du modèle objet métier, nous avons décidé de compléter la structure de données de notre modèle objet métier en le mappant sur un modèle générique pour la prescription [22]. Ceci nous a permis de gagner du temps par rapport au nombre d'itérations qui étaient normalement nécessaires pour compléter la structure du modèle objet métier. Ceci nous a également permis de générer des règles métier plus en adéquation avec les attentes des pharmaciens en termes de vocabulaire et de sémantique métier. Le modèle de prescription générique de Séné [22] est basé sur un formalisme orienté objet différent d'UML, appelé formalisme de Coad-Yourdon. Nous avons donc dû définir des règles de correspondances entre les éléments graphiques du formalisme de Coad-Yourdon et ceux d'UML afin de réécrire le modèle de Séné suivant la syntaxe d'UML. Pour effectuer la transformation du modèle de Séné vers la version UML, nous avons considéré le modèle de Séné comme un patron de conception (*design pattern*) [23]. Nous avons également procédé à quelques modifications sur la structure de données du modèle de Séné pour qu'il soit compatible avec les contraintes de notre application [23].

Après avoir identifié les relations pertinentes entre classes, on instancie une règle métier en utilisant les règles de correspondance entre la syntaxe SBVR et le modèle objet métier [24, 25]. Par exemple, le nom d'une classe du modèle objet métier correspond à un « nom de concept » dans la syntaxe SBVR et le nom d'une relation entre classes dans le modèle objet métier correspond à une relation de type : <Role1>Verbe<Role2> appelée « Fact type » dans la syntaxe SBVR. On répète cette procédure pour chaque médicament appartenant à la liste des médicaments établie avec les pharmaciens pour chaque modèle de règles écrit BRx en prenant en compte la dénomination commune internationale des médicaments. On obtient ainsi toutes les instances de règles.

Nous avons généré un test unitaire pour chaque modèle de règles implémenté dans JRules® en utilisant l'outil de génération automatique fourni par JRules®. Nous avons simulé l'exécution de notre système dans des conditions réelles avec un programme permettant d'injecter des données en entrées (le nom du médicament prescrit, l'âge du patient, le taux de potassium...) à nos règles. Le test unitaire est réussi si le système fait appel à la règle correspondante qui déclenche l'alerte adéquate pour des conditions anormales telles qu'un taux de potassium anormal, un débit de filtration glomérulaire anormal, etc.

3 Résultats

Les résultats que nous présentons suivent les enchaînements d'activité et les phases précédemment définies. Pour illustrer la « traçabilité » entre les différents artéfacts produits dans cette étude (diagrammes UML et règles métier), nous déroulons toutes les étapes de notre méthode pour un cas d'utilisation système : « Valider les prescriptions ».

3.1 Modèle métier et modèle système

Après avoir interviewé les pharmaciens en charge de la validation pharmaceutique, nous avons identifié 7 acteurs métier et deux cas d'utilisation métier. Le cas d'utilisation métier « Valider les prescriptions » comprend un scénario principal et 14 scénarios alternatifs.

3.1.1 Spécification des règles métier

Le choix des lignes de prescription que le système doit contrôler a été justifié par l'analyse des prescriptions des 300 lits informatisés à l'HEGP. Les pharmaciens ont sélectionné les classes de médicaments qui peuvent faire l'objet d'alerte, *i.e.* les médicaments dont l'adaptation posologique est parfois nécessaire suivant les paramètres biologiques du patient. Nous avons ensuite interrogé la base de données de Dx-C@re® afin de connaître la proportion de ces médicaments par rapport au nombre total de prescriptions faites dans Dx-C@re®. Les médicaments listés dans cette étape sont le point de départ pour l'écriture de nos règles métier, tel que chaque règle métier doit être liée à un cas d'utilisation système. Suivant les résultats de l'interrogation du système d'information de l'HEGP, les modèles de règles écrits choisis sont :

- BR1: « Adaptation de posologie pour les médicaments responsables de l'hyperkaliémie en fonction du taux de potassium du patient » ;
- BR2 : « Adaptation de posologie de médicaments en fonction du débit de filtration glomérulaire (DFG) » ;
- BR3 : « Adapatation de posologie de médicaments en fonction de l'International Normalized Ratio (INR) et / ou l'activité anti-Xa et la situation clinique du patient ».

Les lignes de prescription correspondant aux modèles de règles BRx représentent 5,74 % (8 584 lignes de prescriptions) du nombre total de prescriptions dans les services validées par les pharmaciens au cours de l'année 2006. Par rapport à la charge de travail journalière des pharmaciens valideurs, la pré-validation de 5,74 % des lignes de prescriptions est considérée comme satisfaisante par les pharmaciens.

Tableau 1 - Table de décision représentant une règle de type « Si Alors » correspondant au modèle de règle écrit BR3.

Colonnes conditions : contrôle de l'INR et ou de l'activité AntiXa				**Colonne actions**
Classe thérapeutique (médicaments)	**Mode d'injection**	**Situation clinique**	**Paramètre biologique**	
Héparine Calcique	2 à 3 injections par jour	Traitement curatif thrombose veineuse profonde	TCA entre 1,5 et 2 fois le témoin ou anti Xa entre 0,15 et 0,3 IU/mL	Le système valide la prescription
Héparine Calcique	2 à 3 injections par jour	Traitement curatif thrombose veineuse profonde	Anormal : soit le TCA est anormal, soit l'activité Anti Xa soit les deux	Le système alerte les pharmaciens

En partant des trois modèles de règles écrits BRx et de la description détaillée des cas d'utilisation systèmes qui y sont liés, nous avons écrit des règles de type « si / alors » sous forme de table de décision avec des colonnes conditions et des colonnes actions (tableau 1). Chaque entrée du tableau est un médicament appartenant à la liste de médicaments établie avec les pharmaciens. Les items des colonnes sont directement pris des descriptions de cas d'utilisation systèmes correspondants.

3.1.2 Expression des besoins des pharmaciens

En partant de la modélisation métier basée sur le modèle de cas d'utilisation métier et en prenant en considération les évaluations des pharmaciens durant les différentes itérations,

nous avons identifié trois cas d'utilisation système et trois acteurs systèmes liés au cas d'utilisation métier « Valider les prescriptions » et aux trois modèles de règles écrits BRx :

- Cas d'utilisation : « Valider la posologie des médicaments prescrits responsables de l'hyperkaliémie en fonction du taux de potassium du patient », qui inclut un scénario principal et 9 scénarios alternatifs.
- Cas d'utilisation : « Valider les prescriptions des patients insuffisants rénaux traités avec des médicaments qui doivent être adaptés au Débit de Filtration Glomérulaire DFG », qui inclut un scénario principal et 8 scénarios alternatifs.
- Cas d'utilisation : « Valider la posologie des anticoagulants prescrits en fonction de l'International Normalized Ratio (INR) et/ou de l'activité Anti-Xa et de la situation clinique du patient », qui inclut un scénario principal et 8 scénarios alternatifs.

Les modèles de règles écrits doivent guider la description du ou des cas d'utilisation systèmes qui leur correspondent. La structuration des trois cas d'utilisation concrets liés aux trois modèles de règles BRx donne lieu à un nouveau cas d'utilisation dit Abstrait (figure 1). Nous utilisons le cas d'utilisation abstrait « Valider les prescriptions » pour illustrer les artéfacts UML produits dans les étapes suivantes.

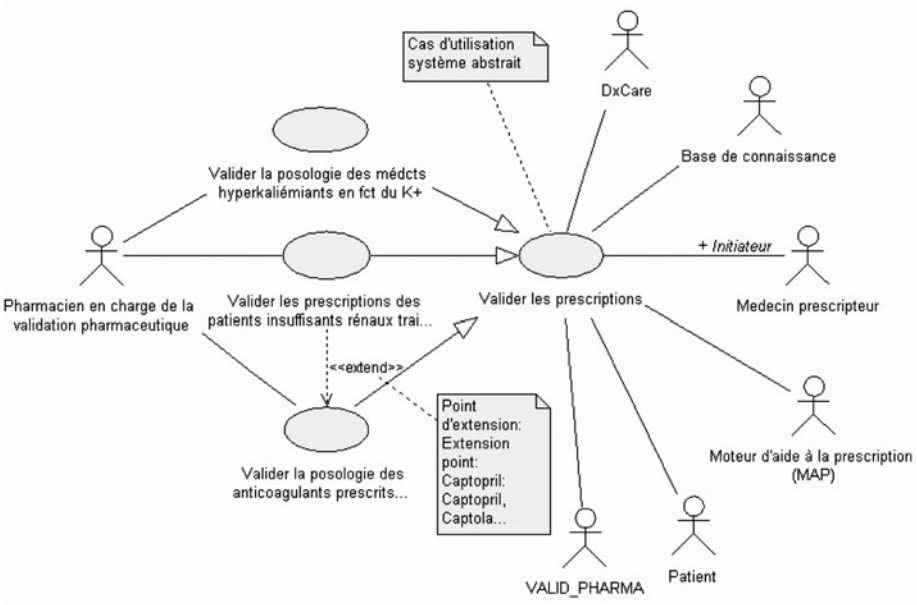


Figure 1 - Structuration du modèle de cas d'utilisation système.

3.2 Modèle de conception et instanciation des règles

L'analyse des cas d'utilisation liés aux modèles de règles écrits BRx nous a permis d'instancier les premières ébauches de classes de conception : médicament, message, prescription, mode d'injection, patient, données de validation. Ces classes nous ont permis de construire une première version de notre diagramme de classes. Nous avons ensuite consolidé la structure du diagramme de classes grâce au modèle de Séné [22]. Nous avons étudié en détail les classes de ce modèle et changé le nom de certaines classes afin de se conformer à la sémantique et au vocabulaire métier de notre contexte d'application. Par

exemple, les pharmaciens ont préféré gardé le nom de classe « Médicament » pour la classe « Préparation Médicamenteuse » proposée initialement. Nous avons également modifié certaines relations entre les classes [23]. Par exemple, la relation entre la classe « Administration Prescription » (PrescriptionAdministration) et la classe « Evénement » suivant les règles de transformation que nous avons établie du formalisme de Coad-Yourdon vers UML, est une relation de généralisation/spécification. Cependant nous avons préféré avoir une relation d'association entre ces deux classes car une classe administration de médicament peut avoir 0, un ou plusieurs événements déclencheurs. Même chose pour la relation entre la classe « Administration Prescription » (PrescriptionAdministration) et la classe « Phase posologique » (DosageRegimenPhase class).

si le nomSpécialité du Médicament de la Prescription est Heparine sodique et le verbeAdministration de la Prescription est Perfusion en Continue et la situationClinique de la DonnéeValidation du Patient concerné par la Prescription est Traitement curative thrombose veineuse profonde et (le TCA du RésultatBiologie du Patient est entre 1.5 fois le témoin et 2 fois le témoin ou l'activité anti-Xa du RésultatBiologie du Patient est entre 0.15 et 0.3) alors l'avisPharmacien du Message affiché pour la Prescription est prescription valide

Figure 2 - Instance de règle SBVR correspondant au modèle de règle écrit BR3.

La figure 2 illustre l'instanciation d'une règle SBVR en utilisant les règles de correspondances entre les éléments graphiques de notre modèle objet métier et les éléments syntaxiques de SBVR. En réitérant cette procédure pour chaque modèle de règles écrit BRx et pour chaque médicament appartenant à la liste des médicaments établie avec les pharmaciens dans la phase précédente, nous avons instancié :

- 247 règles pour le modèle de règles BR1 ;
- 156 règles pour le modèle de règles BR2 ;
- 24 règles pour le modèle de règles BR3.

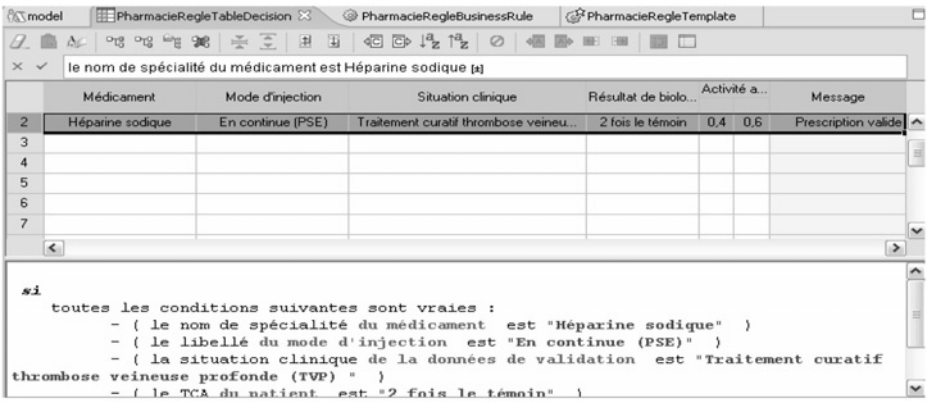


Figure 3 - Exemple de table de décision implémentant une règle du modèle de règle BR3.

3.3 Implémentation et tests

Nous avons implémenté chaque classe de notre diagramme de classes sous forme de classes Java sous Eclipse et généré le modèle objet métier correspondant grâce à ILOG JRules®. Le modèle objet métier ainsi implémenté nous permet d'écrire les règles correspondant à nos modèles de règles métier BR1, BR2, BR3. La figure 3 illustre l'implémentation de la règle métier de la figure 2 dans ILOG JRules® sous forme de table de décision.

Les tests unitaires effectués ont été concluants. Par exemple, nous avons injecté les données suivantes à la règle métier de la figure 3 :

- L'attribut « nomSpécialité » de la classe « Médicament » reçoit la valeur « Heparine sodique ».
- L'attribut « verbeAdministration » de la classe « Prescription » reçoit la valeur « Perfusion en Continue ».
- L'attribut « situationClinique » de la classe « DonnéeValidation » reçoit la valeur « Traitement curative thrombose veineuse profonde' ».
- L'attribut « TCA » de la classe « RésultatBiologie » reçoit la valeur « 1,6 ».
- L'attribut « activité AntiXa » de la classe « RésultatBiologie » reçoit la valeur '0,2'.

JRules® déclenche la règle correspondante et assigne à l'attribut « avisPharmacien » de la classe « Message » la valeur « Prescription valide » que le système affiche à l'utilisateur.

4 Discussion et conclusion

L'objectif de ce travail est de proposer un *framework* générique pour la conception et l'implémentation d'alertes sous forme de règles métier implémentées dans un système d'aide à la décision pour l'aide à la validation pharmaceutique.

L'adaptation du Processus Unifié que nous proposons consiste à introduire une nouvelle activité : la spécification des règles métier. Cela nous a permis de conserver la « traçabilité » entre les différentes phases et activités de conception du Processus Unifié mais aussi entre les différents artéfacts UML produits dans chaque étape de conception. Les cas d'utilisation sont toujours le fil d'Ariane du processus de conception (comme c'est le cas dans la mise en œuvre traditionnelle du Processus Unifié) mais ils sont maintenant liés à des modèles de règles métier qui guident leur description dans le contexte applicatif courant.

Le choix de SBVR comme formalisme de spécification des règles métier nous a permis d'écrire des règles métier qui respectent le vocabulaire métier des pharmaciens. De plus SBVR permet d'instancier des règles en partant d'un diagramme de classes de conception. Cela implique qu'il n'y a pas de dichotomie entre l'adaptation du Processus Unifié que nous proposons et l'utilisation de SBVR dans le cadre de cette adaptation.

Le choix d'inclure un système de gestion de règles métiers dans l'architecture de notre système, en l'occurrence ILOG JRules® semble être stratégique. Une étude comparative entre différents système de gestion de règles métiers prenant en compte différents critères : la capacité du système à s'intégrer avec des systèmes externes, la performance et la scalabilité du système, l'environnement d'édition des règles... a conclu au fait que ILOG JRules® est le système qui réunit le plus de critères décisifs [26]. JRules® nous a permis de reproduire la même simplicité de vocabulaire pour l'implémentation des règles que celles écrites avec SBVR dans l'étape de conception. L'utilisation d'un système de gestion de règles métiers dans l'architecture d'un système d'alerte dans le cadre applicatif de la

validation pharmaceutique est nécessaire tant le nombre de règles produites peut être important. La maintenance de ces règles est très fréquente et ILOG JRules® permet d'inclure les pharmaciens dans cette activité en utilisant son module de développement collaboratif Ilog Rule Team Server.

Les règles que nous avons implémentées dans notre système sont destinées au contrôle d'adaptation de posologie en fonction de résultats de biologie. L'extension du système à d'autres catégories de règles notamment celles qui traitent des contre-indications (interactions médicament – maladie) est une perspective importante [27].

Afin de réaliser des tests plus précis, les règles implémentées dans ILOG JRules® devraient être intégrées dans le SIH de l'HEGP pour exploiter les données issues de Dx-C@re®. Un développement informatique supplémentaire est souhaitable pour coupler le système à base de règles avec une base de connaissance du commerce. Pour anticiper l'extension du système d'alerte à des catégories de règles qui traitent des interactions médicamenteuses, nous avons inclus une base de connaissances sur les médicaments en tant qu'acteur dans la description de nos cas d'utilisation système. De nombreuses sources dans la littérature insistent sur la nécessité d'inclure une base de connaissances médicamenteuse issue du commerce pour le contrôle des interactions, des posologies et des dosages et réduire de ce fait les risques pour les patients. Cependant ces mêmes sources insistent aussi sur l'adaptation de celles-ci à la réalité du contexte local de chaque organisme de santé [27, 28]. À l'HEGP, les pharmaciens peuvent effectuer un contrôle des interactions médicamenteuses des prescriptions. Le système de prescription informatisé Dx-C@re® s'appuie sur la base de connaissances VIDAL® pour effectuer le contrôle en question. Les règles implémentées au sein de cette base de connaissances pour mettre en œuvre le contrôle des interactions sont de simples tables de mapping entre la liste des médicaments transmise par Dx-C@re® et les données présentes dans VIDAL®. La majorité des messages affichés par ce contrôle sont considérés comme peu précis par les pharmaciens. Les messages d'alerte affichés sont très souvent de haut niveau et non justifiés. Ceci peut s'expliquer par la non-prise de risque intentionnelle du fournisseur. La non-pertinence des messages d'alerte et leur manque de précision se traduit généralement par un phénomène de perte de confiance des praticiens vis-à-vis des systèmes d'aide à la décision connu sous le nom de « Wolf Cry Syndrome » [29]. Dans le cas de Dx-C@re®, les messages affichés sont très souvent redondants. Par exemple : le contrôle d'interaction entre deux médicaments A et B conduit à l'affichage de deux alertes par Dx-C@re®: une fois pour l'interaction entre A avec B, puis une deuxième fois pour l'interaction de B avec A. La modification de la logique métier qui effectue les contrôles d'interactions dans Dx-C@re® nécessiterait de modifier le code source de cette application ce qui rend l'établissement dépendant de son fournisseur pour ces modifications. Les règles implémentées au moyen d'un système de gestion de règles métier peuvent être adaptées facilement au contexte local d'un établissement en modifiant uniquement la base de connaissances. Il s'agit donc de la solution la plus appropriée pour les établissements qui ont choisi d'implémenter leurs règles en fonction de leurs besoins.

Remerciements

Nous remercions la société ILOG pour nous avoir fourni une version d'évaluation de ILOG JRules®. Nous remercions également le Dr Catherine Duclos du laboratoire Lim & Bio ainsi que les Drs Philippe Guillaume et Thibaut Caruba de la pharmacie centrale de l'HEGP.

Références

[1] Moore N, Lecointre D, Noblet C, Mabille M. Frequency and cost of serious adverse drug reactions in a department of general medicine. *Br J Clin Pharmacol* 1998; 45(3): 301-8.

[2] Oliven A, Michalake I, Zalman D, Dorman E, Yeshurun D, Odeh M. Prevention of prescription errors by computerized, on-line surveillance of drug order entry. *Int J Med Inform* 2005; 74(5): 377-86.

[3] Van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006; 13(2): 138-47.

[4] Moyen E, Camiré E, Stelfox HT. Clinical review: medication errors in critical care. *Crit Care* 2008; 12(2): 208. Epub 2008 Mar 12.

[5] Order of 9 August 1991, implementing Article R. 5203 Code of Public Health.

[6] Order of 31 March 1999 relating To "prescription, to drug dispensation and administration subject to the regulation of poisonous substances in health institutions, To Hospitals labour unions and To socio-medical establishments with a pharmacy for internal use referred to in Article L. 595-1 of the Public Health Code, 1999.

[7] Vernardet S, Bossaert S, Livrozet A, Pont E, Charpiat B. Pharmaceutical analysis and validation of hospital prescriptions, a 5-year study. *Presse Med* 2005; 34(14): 990-6.

[8] Fair MA, Pane F. Pharmacist interventions in electronic drug orders entered by prescribers. *Am J Health Syst Pharm* 2004; 15; 61(12): 1286-8.

[9] Dooley MJ, Allen KM, Doecke CJ, Galbraith KJ, Taylor GR, Bright J, Carey DL. A prospective multicentre study of pharmacist initiated changes to drug therapy and patient management in acute care government funded hospitals. *Br J Clin Pharmacol* 2004; 57(4): 513-21.

[10] Degoulet P, Marin L, Lavril M, Le Bozec C, Delbecke E, Meaux JJ, Rose L. The HEGP component based clinical information system. *Int J Med Inform* 2003; 69(2-3): 115-26.

[11] McMullin ST, Reichley RM, Kahn MG, Dunagan WC, Bailey TC. Automated system for identifying potential dosage problems at a large university hospital. *Am J Health Syst Pharm* 1997; 54(5): 545-9.

[12] Aarts J, Doorewaard H, Berg M. Understanding implementation: the case of a computerized physician order entry system in a large Dutch university medical center. *J Am Med Inform Assoc* 2004; 11(3): 207-16.

[13] Peleg M, Tu S. Decision support, knowledge representation and management in medicine. *Methods Inf Med* 2006; 45 Suppl 1: 72-80.

[14] Knape T, Hedermana L, Wadea VP, Gargana M, Harrisa C, Rahmanb Y. A UML Approach to Process Modelling of Clinical Practice Guidelines for Enactment, *Stud Health Technol Inform* 2003; 95: 635-40.

[15] LeBozec C, Jaulent MC, Zapletal E, Degoulet P. Unified modeling language and design of a case-based retrieval system in medical imaging. *Proc AMIA Symp* 1998; 887-91.

[16] In The Unified Software Development Process, eds Ivar Jacobson, Grady Booch, James Rumbaugh, *Addison-Wesley Professional*, 1999.

[17] Peleg M *et al.* Comparing computer-interpretable guideline models: a case-study approach. *J Am Med Inform Assoc* 2003; 10(1): 52-68.

[18] Sordo M, Boxwala AA, Ogunyemi O, Greenes RA. Description and status update on GELLO: a proposed standardized object-oriented expression language for clinical decision support. *Medinfo* 2004; 11(Pt 1): 164-8.

[19] "Anonyme" Semantics of Business Vocabulary and Business Rules Specification, Document de specification de l'OMG, 2006. http://www.omg.org/docs/formal/08-01-02.pdf last accessed on 2008/05/02.

[20] Ambler S. How to record the transactional logic behind UML object models http://www.ddj.com/dept/architect/184414624?cid=Ambysoft last accessed on 2008/05/15;

[21] Nijpels R. Business Rules in Requirements Analysis http://www.brcommunity.com/b259.php#fn1 Access with login, last accessed on 2008/05/15;

[22] Séné B, Venot A, de Zegher I, Milstein C, Errore S, de Rosis F, Strauch G. A general model of drug prescription. *Methods Inf Med* 1995; 34(4): 310-7.

[23] Gamma E, Johnson R, Helm R. Design patterns: Elements of reusable object-oriented software, *Addison-Wesley Professional* 1995.

[24] "Anonyme" Business Semantics of Business Rules, Business Rules Team Response to RFP: SBVR Submission bei/2005-08-01 Version 8. http://www.omg.org/docs/bei/05-08-01.pdf last accessed on 2008/05/02

[25] Schacher M. Moving from Semaman Row 2 to Zachman Row 3 Business Rules from an SBVR and an xUML Perspective, *KnowGravity* 2006.

[26] Goldberg HS, Vashevko M, Pastilnik A, Smith K, Plaks N, Blumenfeld BM. Evaluation of a commercial rule engine as a basis for a clinical decision support service. *AMIA Annu Symp Proc* 2006: 294-8.

[27] Kuperman GJ, Bobb A, Payne TH, Avery AJ, Gandhi TK, Burns G, Classen DC, Bates DW. Medication-related clinical decision support in computerized provider order entry systems: a review. *J Am Med Inform Assoc* 2007; 14(1): 29-40.

[28] Kuperman GJ, Reichley RM, Bailey TC. Using commercial knowledge bases for clinical decision support: opportunities, hurdles, and recommendations. *J Am Med Inform Assoc* 2006; 13(4): 369-71.

[29] Magnus D, Rodgers S, Avery AJ. GPs' views on computerized drug interaction alerts: questionnaire survey. *J Clin Pharm Ther* 2002;27(5): 377-82.

Adresse de correspondance

Abdelali Boussadi
INSERM U872, eq 20
Faculté de Médecine Paris 5
15, rue de l'École de Médecine
75006 Paris, France

E-mail : boussadiabdelali@yahoo.fr

Chapitre 6

Données médicales et usages en santé publique

Informatique et Santé, 2009 (17) : 227-238
© Springer – Verlag France

Projet REUSE : utilisation du dossier patient informatisé (DPI) en recherche clinique

Naji El Fadly[1,3], Noël Lucas[2], Pierre-Yves Lastic[4], Philippe Verplancke[5], Christel Daniel[1,2]

[1]*INSERM UMRS 872, équipe 20, Université René Descartes, Paris, France*
[2]*Hôpital Européen Georges Pompidou, Département d'Informatique Hospitalière, Paris, France*
[3]*Thales Service, Vélizy, France,* [4]*Sanofi-aventis & CDISC Board of Directors, Bagneux, France,*
[5]*XClinical GmbH, Munich, Allemagne*

Abstract

Integrating clinical research data entry with patient care data entry is a challenging issue. At the G. Pompidou European Hospital (HEGP), pulmonary function testing reports are captured twice, first in the Electronic Health Record (EHR) and then in Clinical Data Management System (CDMS). Statuary requirements and data standards are different for EHR and clinical research. So, this work has for object to suggest an integration solution between EHR and CDMS. This object can be divided in two steps: the first consists in modeling clinical research workflow in a university hospital environment in order to investigate the possible use of her in this context. The second step consist in conceiving, implementing and assessing an integration solution between EHR and CDMS taking into consideration standardization efforts and statuary requirements. The REUSE demonstrator showed that automatic transfer of data reduces errors risk, and enables to preserve collected data during clinical trials in the EHR that may be useful. We proposed the integration profile "Retrieve and Integrate Form for Data Capture". Beyond syntactic interoperability, a perspective is to implement this integration profile in the REUSE demonstrator in order to address the issue of semantic interoperability between both domains.

Keywords

HL7; CDA; CDISC; ODM; IHE profile; RFD; Electronic Health Record; Clinical Data Management System; Biomedical Research; Clinical trial; Pulmonary function testing.

1 Introduction

Dans un contexte d'accélération du processus d'informatisation des hôpitaux et de mise en place de dossiers médicaux personnels, la collecte de données cliniques présente un intérêt important dans un cadre de recherche clinique. Parfois, le monitoring des données de recherche peut également permettre la correction d'erreur dans le DPI. Les données d'un Dossier Patient Informatisé (DPI) peuvent être utiles dans un contexte de recherche clinique à plusieurs titres. Ces données peuvent contribuer notamment à améliorer le taux d'inclusion des patients dans des essais cliniques, à optimiser le recueil de données

cliniques et à améliorer la sécurité du patient. Par ailleurs, les DPI peuvent bénéficier des résultats de la recherche clinique en intégrant des systèmes ayant la charge de comparer les pratiques de soins et les résultats de la recherche.

Actuellement, les systèmes d'information de soin et de recherche clinique sont totalement déconnectés y compris dans les Centres Hospitalo-Universitaires (CHU) où pourtant on trouve les implantations les plus abouties de Dossier Patient Informatisé (DPI). Dans l'état actuel du fonctionnement des CHU, les cliniciens sont confrontés à un double recueil de données d'abord au sein du DPI puis ensuite dans un Système de Gestion de l'Essai Clinique (SGDC). Des études ont mis en évidence que 30 % à 50 % des items du dossier de soins étaient pertinents dans un contexte de recherche clinique [1]. La saisie multiple des mêmes données au sein d'applications différentes est une perte de temps au cours du processus de mise en œuvre d'un essai clinique et également à l'origine d'erreurs aboutissant à un recueil de données incohérentes entre applications.

Dans ce contexte, il est possible, et intéressant, d'améliorer l'efficacité du recueil de données d'essais cliniques dans les établissements de santé disposant de DPI. Les facteurs limitant la mise en œuvre de solutions opérationnelles d'interopérabilité entre les DPI implantés dans les hôpitaux et les SGDC des unités de recherche clinique sont de deux ordres : (i) d'une part d'ordre organisationnel, puisque les processus de soins et de recherche clinique étant différents, le traitement informatisé des données est réalisé de manière différente dans ces deux contextes. Il doit de plus se conformer à des impératifs réglementaires différents dans les deux cas ; (ii) d'autre part d'ordre technologique, puisqu'il existe des standards différents de codage de l'information dans le domaine des systèmes d'information de santé (standards HL7) [2] et de la recherche clinique (standards CDISC) [3].

L'objectif de notre travail est de proposer une solution d'intégration entre DPI et SGDC et de l'évaluer dans le contexte hospitalier de l'HEGP. Cet objectif se décompose en deux sous-objectifs. La première étape consiste à modéliser le déroulement des activités de recherche clinique dans un environnement hospitalo-universitaire afin d'investiguer l'utilité du DPI pour le processus de recherche clinique. La deuxième étape consiste, en s'appuyant sur cette analyse de processus, à concevoir, mettre en œuvre et évaluer une solution d'intégration entre les DPI et les SGDC tenant compte des efforts de standardisation internationaux, des contraintes réglementaires internationales et nationales et des contraintes techniques locales.

Dans cet article, nous présentons tout d'abord l'état de l'art concernant les contraintes réglementaires de mise en œuvre de SGDC, les efforts d'harmonisation des standards internationaux des systèmes d'information de santé (HL7) et de recherche clinique (CDISC) et une solution d'intégration DPI/SGDC explorée au niveau international dans le cadre de l'initiative IHE (*Integrating the Healthcare Entreprise*). Puis, nous exposons dans un premier temps la méthodologie que nous avons utilisée pour le développement de la solution intégrée DPI/SGDC de l'HEGP. Cette méthodologie s'appuyant sur les principes du Processus Unifié (*Unified Process*) de l'OMG comporte des étapes de modélisation de processus, d'analyse et de conception et d'implémentation et de test [4]. Enfin, nous présentons les modèles d'analyse et de conception du démonstrateur REUSE (*Retrieve in EHR Useful data and knowledge for Secondary Exploitation*) mis en œuvre à l'HEGP, discutons les résultats obtenus et présentons les perspectives d'expérimentation dans un contexte d'essai clinique multicentrique.

2 État de l'art

Dans un contexte de promotion institutionnelle, un essai clinique peut être décomposé en trois grandes phases : la phase de soumission et d'appel d'offre consistant à définir un protocole de recherche et à le soumettre au promoteur ; la phase d'instruction consistant à préparer tous les outils pour gérer les données des cahiers d'observations et à achever les tâches réglementaires et financières et la phase de mise en œuvre où sont gérés les consentements des patients, la collecte des données de l'essai clinique et les analyses statistiques.

Des travaux récents ont investigué le rôle que pouvait jouer le DPI dans un contexte de recherche clinique lors de chacune de ces phases et également en amont et en aval de l'essai clinique [1, 5]. Le tableau 1 résume une partie des résultats de ces travaux.

Tableau 1 - Le rôle éventuel du DPI lors de chacune des étapes de l'étude clinique.

Etape de l'étude		Rôle du dossier patient informatisé
Formulation de l'hypothèse de recherche		• Faciliter l'accès aux données cliniques pour aider à générer des hypothèses de recherche
Instruction	Configuration de l'étude	• Interroger la base de données du DPI pour connaître le nombre de patients concernés par l'étude
Instruction	Recrutement et inscription	• Mettre en œuvre des paramètres d'identification de patients éligibles • Informer les professionnels de santé de l'éligibilité de patients dont ils ont la charge • Contacter les patients par l'intermédiaire du professionnel de santé
Mise en œuvre	Gestion des données de l'étude	• Intégrer la saisie des données spécifiques à l'étude dans le cadre des activités de soins • Mettre en œuvre des règles et des alertes afin de respecter les exigences de la collecte des données pour l'étude
Mise en œuvre	Soumission des données de l'étude	• Fournir les données de l'étude aux agences réglementaires (par exemple) en utilisant les standards
Mise en œuvre	Rapport d'événements indésirables	• Fournir les rapports des effets indésirables

Mise en œuvre	Soumission des conclusions de l'étude	• Fournir les conclusions de l'étude (promoteur)
		  • Aux agences réglementaires en utilisant les standards
		  • Aux investigateurs et aux patients

Intégration des résultats de l'étude aux pratiques cliniques		• Mettre en œuvre des conclusions de l'étude dans le DPI (documentation, protocoles, alertes) • Évaluer la cohérence entre les résultats de l'étude et la pratique actuelle • Surveiller l'impact de l'aide à la décision sur l'évolution des pratiques et sur l'état de santé

L'utilisation du DPI dans un contexte de recherche clinique nécessite d'examiner attentivement le contexte réglementaire et technique de la mise en œuvre des SGDC qui est différent de celui des systèmes d'information dans le domaine du soin.

En ce qui concerne le respect de la confidentialité et de la sécurité des données, comme tout projet de recueil d'informations individuelles, la gestion informatisée des données d'un essai clinique doit être soumise à la Commission Nationale de l'Informatique et des Libertés (CNIL) qui veille au respect de la loi Informatique et Liberté, instituée par la loi

78-17 du 6 janvier 1978. Un des objectifs de cette loi est que les personnes concernées soient informées du fait que leurs données personnelles vont être traitées. Cette obligation d'information du patient a été rappelée dans la recommandation n° 58 du CCNE du 12 juin 1998 et plus récemment dans la loi du 4 mars 2002 [7]. Enfin, le Conseil National de l'Ordre des Médecins (CNOM) [8] joue un rôle important lors des essais cliniques, en particulier en ce qui concerne le contrôle des conventions conclues entre les médecins et les promoteurs [9]. En principe, les recherches dans le domaine de la santé mettant en œuvre des traitements de données nominatives ne peuvent se faire qu'après avis d'un comité consultatif institué auprès du ministre de la Recherche, préalablement à tout dépôt de dossier auprès de la CNIL (CCTIRS). Cependant, le législateur a introduit une simplification des démarches de déclaration, ainsi lorsque le responsable du traitement des données s'engage à respecter une méthodologie de référence (MR001) et n'a plus à déclarer l'ensemble de ses recherches. En cas de changement de finalité du traitement, c'est-à-dire d'utilisation pour la recherche de données recueillies dans le cadre des soins, le patient doit être informé de leur transmission et de leur utilisation pour la recherche. La MR001 ne s'applique plus et l'avis du comité consultatif ainsi qu'une demande d'autorisation spécifique à la CNIL sont requis [10-12].

En ce qui concerne les aspects techniques de la gestion informatisée des essais cliniques, l'organisation de standardisation CDISC (*Clinical Data Interchange Standards Consortium*), a été mise en place en 1997 par la FDA (*Food and Drug Administration*) [13], afin de développer des standards indépendants des plates-formes pour l'acquisition, l'échange, la soumission et l'archivage des données des essais cliniques. Parmi ces standards figure le modèle de données opérationnel ODM (*Operational Data Model*) qui est distribué par CDISC depuis 2001 [14]. Ce modèle définit de façon très précise les différents types d'information en rapport avec un essai clinique : les méta-données de l'essai clinique (définition d'un protocole et des items), les données de référence de l'essai clinique (par exemple les valeurs normales des examens de laboratoire), les données administratives (droits d'accès et profils utilisateur) et cliniques de l'essai, ainsi que le suivi des modifications des données (*audit trail*). En ce qui concerne le DPI, les efforts de standardisation de la structuration et de la communication des informations médicales sont conduits au sein d'organismes différents de CDISC, européens (CEN TC 251) [15] et internationaux (HL7, *Health Level* 7). HL7 propose notamment un Modèle d'Information de Référence (RIM) sur lequel sont basés les modèles de messages et également les modèles de partage de documents cliniques (CDA, *Clinical Document Architecture*) [16].

Afin de faciliter l'intégration des systèmes d'information de soin et de recherche clinique, des efforts d'harmonisation ont été initiés entre HL7 et CDISC au sein du comité technique RCRIM (*Regulated Clinical Research Information Management*) d'HL7. Le NCI (*National Cancer Institute*) [17] et la FDA participent activement à ces efforts. Dans le cadre de ces efforts, le projet BRIDG a pour objectif la réalisation d'un modèle d'information de référence couvrant l'ensemble du domaine de la recherche clinique et permettant de développer une sémantique commune aux différents acteurs (hôpitaux, promoteurs d'essais cliniques, autorités de santé, etc.) [18- 20].

Enfin, nous avons identifié des initiatives proposant des solutions d'intégration de systèmes d'information de santé et de recherche clinique ayant fait l'objet de démonstrateurs [21, 22]. En particulier, dans le cadre de l'initiative IHE (*Integrating Healthcare Enterprise*), le profil d'intégration RFD (*Retrieve Form for Data-capture*) [23] a pour objectif de fournir une méthode de collecte de données au sein d'une application utilisée (par exemple un DPI) pour satisfaire les exigences d'un système externe (par exemple un SGDC). La méthodologie d'IHE consiste en ce que des groupes de travail

(utilisateurs et industriels) spécifient des « profils d'intégration », qui sont des situations réelles décrivant des échanges d'informations appelés « transactions », de différents composants d'un système d'information de santé distribué, appelé « acteurs ». IHE fournit des guides de mise en œuvre de ces « transactions », en utilisant des standards informatiques établis tels que HL7 ou DICOM. Le profil RFD spécifie une solution d'intégration de DPI et de SGDC. Si, moyennant certains accords contractuels, un organisme externe a besoin de collecter des données cliniques dans le cadre d'un essai clinique, le profil RFD permettra à l'utilisateur du DPI (médecin investigateur, par exemple) d'accéder au cahier d'observation clinique électronique (eCRF, *electronic Case Report Form*) de l'essai clinique sans quitter le DPI, de saisir les informations requises afin qu'elles soient collectées par l'organisme externe. Dans ce profil RFD, un composant du SGDC de l'organisme externe fournit des formulaires de saisie de données dans un format défini par le profil. Le profil minimise l'effort que l'application d'affichage (le DPI) doit faire en fournissant des formulaires entièrement fonctionnels qui transportent avec eux les instructions nécessaires au composant de saisie de données («Form Filler») du DPI. Le profil assure le lien entre le formulaire d'affichage et le système gérant la collecte des données (« *Form Receiver* ») afin de traiter les situations de saisie d'une série de formulaires, de saisie partielle de formulaires ou de retour à des formulaires partiellement remplis dans les sessions antérieures. RFD soutient également l'archivage d'une copie du formulaire dûment rempli.

Par ailleurs, une extension du profil RFD propose une solution de pré-remplissage des données du formulaire par des données déjà présentes dans le DPI. En effet, lorsque le profil RFD est complété par le profil de capture de données de recherche clinique (CRD, *Clinical Research Data Capture*), ce dernier décrit le contenu et le format à utiliser pendant la transaction de pré-remplissage des données du formulaire. De plus, ce profil (CRD) assure la capacité de convertir les données dans un standard de cahier d'observation clinique (ODM) [24].

Les limites de cette solution RFD sont liées d'une part à la technologie standard adoptée pour ce profil (XForms) qui n'est pas prise en compte par toutes les plates-formes, ce qui pose un grand problème d'interopérabilité et de mise en œuvre. Par ailleurs, la solution de pré-remplissage du cahier d'observation électronique par les données cliniques pré-existant dans les dossiers repose sur une définition préalable et générique (indépendante d'un essai clinique en particulier) d'un ensemble de données communes au soin et à la recherche clinique. Cet ensemble de données ne constituera qu'une part très limitée des données collectées dans le cadre de l'essai clinique et ne résoudra donc pas, selon nous, la problématique de double saisie.

3 Matériel et méthodes

Nous adoptons une méthodologie s'appuyant sur les principes du Processus Unifié (*Unified Process*) de l'OMG afin d'analyser et de concevoir la solution d'intégration du Dossier Patient Informatisée (DPI) et du Système de Gestion des Données Cliniques (SGDC) de l'HEGP. Nous décrivons ici les étapes de modélisation de cette solution appelée « REUSE » réalisées au sein d'une itération de développement ainsi que le prototype développé à l'issue de cette itération.

3.1 Analyse, conception et implémentation du module REUSE

Les tâches d'analyse et de conception ont été réalisées en tenant compte à la fois des contraintes fonctionnelles et techniques de l'HEGP mais en intégrant également des

artefacts de connaissance du domaine issus des travaux de standardisation réalisés par HL7 ou CDISC. Nous avons utilisé l'outil de modélisation UML Entreprise Architect pour réaliser les modèles [25].

L'analyse de processus de recherche clinique permet de définir le périmètre fonctionnel du module « REUSE » en précisant les étapes du processus de mise en œuvre d'essais cliniques qui peuvent être simplifiées et automatisées grâce à une utilisation du DPI. Nous avons conduit une série d'entretiens avec les acteurs de l'Unité de Recherche Clinique (URC) afin de réaliser les modèles de processus décrivant l'activité de mise en œuvre d'un essai clinique dans un contexte de promotion institutionnelle. Nous avons intégré aux modèles de processus réalisés des informations issues de l'analyse des modèles dynamiques du projet BRIDG.

Nous avons défini l'architecture du module REUSE comme étant une extension de celle décrite dans le cadre du profil d'intégration RFD permettant d'importer un cahier d'observation électronique et les informations nécessaires à la mise en correspondance des items cliniques de ce cahier d'observation électronique avec ceux du DPI.

Dans le cadre de l'analyse du module REUSE, les interactions entre les acteurs de la recherche clinique et REUSE ont été représentées graphiquement par des diagrammes de séquence UML pour exprimer la chronologie des transactions.

Tout comme pour l'étape de modélisation des processus, la réalisation des modèles statiques de REUSE (modèles d'information du domaine) a tenu compte des contraintes fonctionnelles et techniques du projet ainsi que de l'analyse d'artefacts de connaissance de domaine fournis par HL7 ou CDISC. Nous nous sommes notamment appuyés sur l'analyse du format ODM de CDISC.

Une implémentation partielle du module REUSE a été réalisée sous forme d'un médiateur CDISC ODM mettant en œuvre l'export au format ODM de données cliniques du DPI pertinentes dans le cadre d'un essai clinique et leur intégration dans un SGDC. Nous avons utilisé l'API (Application Programme Interface) JDOM [26] de JAVA pour sa simplicité par rapport au parseur SAX [27] et DOM [28] et pour la richesse de ses fonctionnalités. Le médiateur CDISC ODM interroge une copie de la base (Oracle) de production du DPI et reformate les résultats des requêtes au format XML tout en respectant le standard CDISC ODM. Le paramétrage du médiateur est stocké dans un fichier XML. Le message ODM de sortie et importé par la suite dans le SGDC (MARVIN) afin de générer le cahier d'observation électronique et de stocker les données instanciées dans la base de données de l'essai clinique.

3.2 Expérimentation de REUSE dans le cadre de l'étude « Dyspnée »

Notre expérimentation a été menée à l'HEGP qui dispose de l'application DxCare de la société MedaSys© [29] pour gérer le DPI et de l'application MARVIN de la société XClinical [30] en tant que SGDC à l'URC. MARVIN comporte un outil d'élaboration des cahiers d'observation électronique appelé ODM Study Composer.

Le domaine d'application de cette expérimentation concernait les épreuves fonctionnelles respiratoires (EFR). Dans ce service, les cliniciens collectent les données cliniques relatives à ces examens d'EFR dans le DPI (DxCare). Une partie de ces données sont pertinentes dans le contexte de l'étude de recherche clinique « Dyspnées » gérés par l'URC en utilisant MARVIN comme SGDC.

Après une première étape d'analyse comparative des items nécessaires pour le soin et pour l'étude « Dyspnée » ainsi que de leurs formats de représentation, un formulaire mixte soin/recherche clinique du service EFR a été paramétré au sein de l'application du DPI

(DxCare).

Le médiateur CDISC ODM de la solution REUSE a été utilisé pour extraire du DPI (DxCare) les items choisis pour l'étude Dyspnée et les aligner au format CDISC ODM afin de pouvoir les partager avec un SGDC tel que MARVIN.

4 Résultats

4.1 Module REUSE

4.1.1 Modèles de processus

La modélisation du déroulement des activités actuelles de la recherche clinique de promotion institutionnelle dans un environnement hospitalo-universitaire nous a permis d'analyser les différentes étapes d'essai clinique et de préciser l'avantage et les limites d'intégration du DPI dans chaque étape.

Un essai clinique peut être décomposé en trois grandes phases : la phase soumission où l'investigateur principal et l'URC travaillent à la rédaction du protocole de recherche et à sa soumission au promoteur ; la phase instruction où l'URC prépare tous les outils pour la gestion des cahiers d'observations et achève les tâches réglementaires et financières ; et la phase mise en œuvre où sont gérés les consentements des patients, la collecte des données et les analyses statistiques.

Les diagrammes d'activité réalisés ont permis de mieux identifier le rôle des acteurs principaux lors de chaque activité du processus de mise en œuvre d'essai clinique (médecin investigateur, méthodologiste, coordonateur de la structure, Coordonateur des Etudes Cliniques (CEC), statisticien, Data-Manager, Technicien d'Étude Clinique (TEC), Assistant de Recherche Clinique (ARC), Responsable de la Gestion Administrative et Financière (RGAF) et promoteur). Le diagramme de processus a une structure hiérarchique : chacune des phases modélisée est découpée en plusieurs activités composites, elles-mêmes constituées de sous-activités. Il comporte au total 15 diagrammes d'activité. Cette structure permet de naviguer selon l'ordre des activités et d'avoir à la fois une vision globale et détaillée des différentes phases.

Nous avons identifié au sein du diagramme d'activité les activités et les tâches susceptibles d'être simplifiées et automatisées grâce à une utilisation du DPI. La majorité de ces tâches concerne la gestion des données. Par exemple, dans la phase d'instruction de l'étude et en particulier pendant l'étape de gestion des données, les tâches d'élaboration de la base de données de l'essai et du plan de data-management pourraient être automatisées à l'aide du DPI (figure 1).

La solution d'intégration REUSE permet d'effectuer certaines tâches des processus analysés en utilisant le DPI. Les diagrammes de séquence ont été réalisés afin de préciser les interactions des acteurs utilisant la solution REUSE lors du déroulement des cinq étapes suivantes : « Elaborer les eCRF », « Reporter des données existantes dans le DPI », « Modifier des données administratives des patients », « Modifier des données de recherche clinique », « Monitorer les données ».

4.1.2 Architecture de REUSE

L'architecture de la solution REUSE est une extension de celle décrite dans le cadre du profil d'intégration RFD. Nous intégrons à la solution REUSE les composants « *Form Manager* », « *Form Filler* » et « *Form Receiver* » définis dans le profil RFD mais définissons un nouveau composant « *Form Mapper* » permettant d'importer un cahier d'observation électronique ainsi que les informations nécessaires à la mise en

correspondance des items cliniques de ce cahier d'observation électronique avec ceux du DPI. Cette fonctionnalité repose sur la définition d'une nouvelle transaction utilisant le message ODM (MetaData) pour transférer les métadonnées du cahier d'observation électronique du « *Form Manager* » vers le « *Form Mapper* ». Nous avons appelé « *Retrieve & Integrate Form for Data-capture* » (RIFD), ce profil d'intégration qui permet de générer le cahier d'observation électronique de l'essai clinique au sein du DPI (figure 2). Ce dernier étant intégré entièrement dans le DPI, le pré-remplissage des données cliniques préexistant dans le DPI est réalisé grâce à des mécanismes de report des données propres au DPI. Ces mécanismes de report de données reposent sur le résultat d'un travail de mise en correspondance au sein du DPI entre les items de différents formulaires.

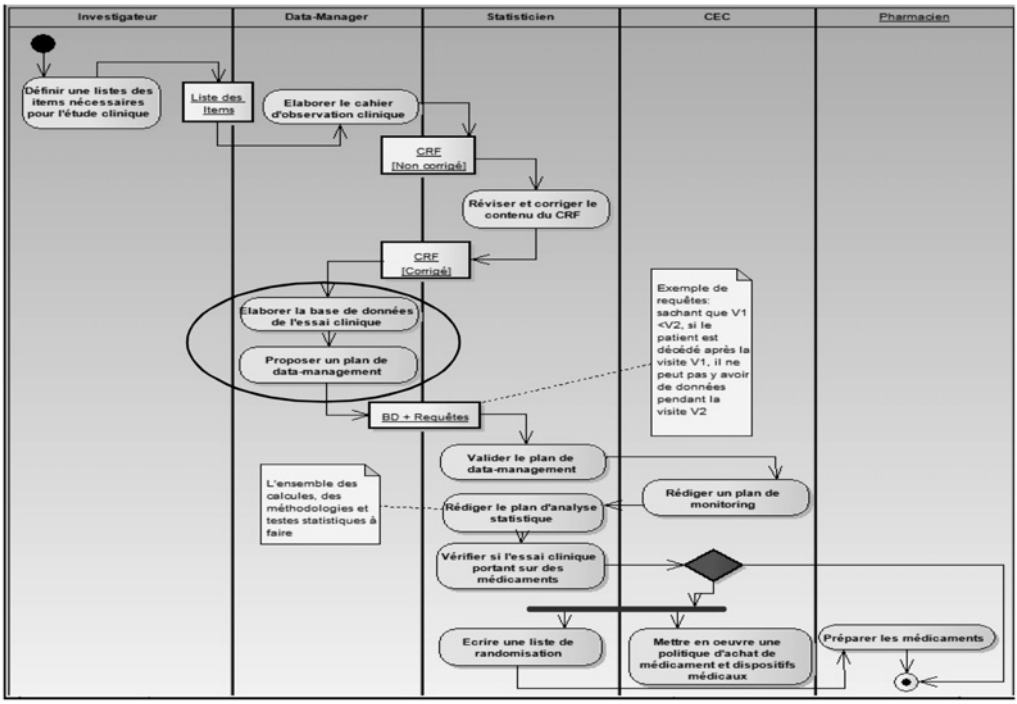


Figure 1 - Exemple d'activités du processus de mise en œuvre d'un essai clinique pouvant bénéficier de l'utilisation du DPI : « Élaborer la base de données de l'essai clinique » et « Proposer un plan de data-management ».

Dans l'étape de collecte de données, le profil RIFD se base sur un message ODM pour transférer les données et les métadonnées correspondantes via une transaction « *Submit Form* » du « *Form Filler/Mapper* » au « *Form Receiver* ». La validation des données se fait dans le SGDC, c'est-à-dire dans une application dont la gestion ne dépend ni de l'investigateur, ni du promoteur (« *Trusted Third Party* »).

4.1.3 Médiateur CDISC ODM de REUSE

Une implémentation partielle du module REUSE a été réalisée sous forme d'un médiateur CDISC ODM. Le médiateur CDISC ODM interroge une copie de la base (Oracle) de production du DPI (DxCare®, MEDASYS©) et reformate les résultats des requêtes au

format XML tout en respectant le standard CDISC ODM. Le paramétrage du médiateur est stocké dans un fichier XML. Le message ODM de sortie est importé par la suite dans le SGDC (MARVIN, XClinical) afin de générer le cahier d'observation électronique et de stocker les données instanciées dans la base de données de l'essai clinique. Le message ODM généré par le médiateur CDISC ODM est produit à l'issue de l'étude clinique, il peut donc contenir plusieurs visites pour plusieurs patients concernés par le même essai clinique. Le patient est référencé par un identifiant unique.

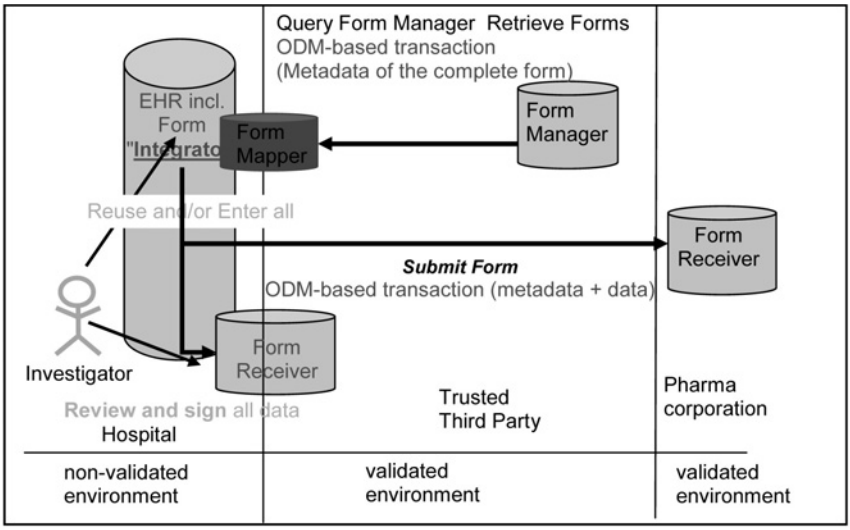


Figure 2 - Acteurs et transactions impliqués dans le profil d'intégration proposé « Retrieve & Integrate Form for Data Capture » (RIFD), basé sur le profil d'intégration IHE RFD.

Les médecins identifiés sont des investigateurs et co-investigateurs (et non plus les médecins responsables comme c'est le cas dans le DPI). Le schéma XML du message ODM se compose de trois rubriques principales :

- Study :
  - GlobalVariables : référence de l'essai clinique.
  - BasicDefinitions : unités de mesure nécessaires à l'essai clinique.
  - MetaDataVersion et ReferenceData : références et définitions des données (exemple : libellés, types, contraintes, codage, etc).
- AdminData : références des investigateurs et des sites.
- ClinicalData : patient (anonyme), visites, formulaires, groupes d'items, items.

4.2 Utilisation de REUSE dans l'étude « Dyspnée »

Dans le cadre de l'étude « Dyspnée », l'expérimentation a consisté à exporter les données d'observation d'Épreuves Fonctionnelles Respiratoires du DPI (DxCare, MEDASYS©) vers le SGDC (MARVIN, XClinical).

La figure 3 met en évidence comment le médiateur CDISC ODM permet de transférer une partie des données cliniques saisies dans le DPI dans le cadre d'une activité de soins vers le SGDC.

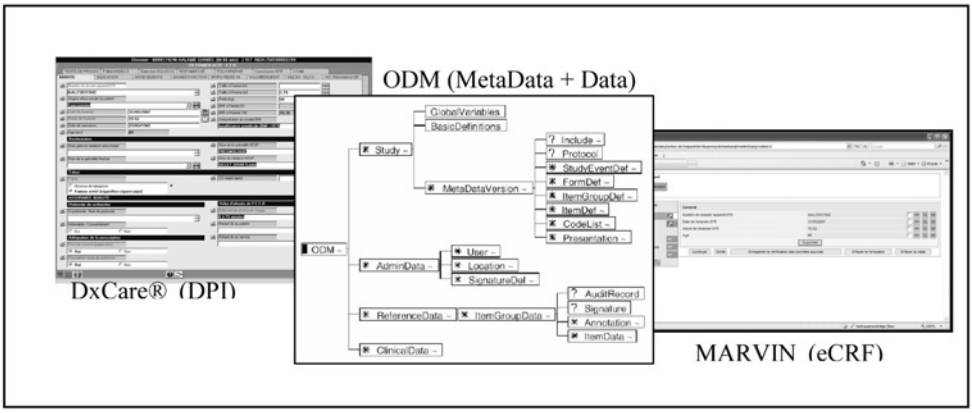


Figure 3 - Export des méta données et des données cliniques d'un essai clinique depuis le DPI (DxCare) vers le SGDC (MARVIN) en utilisant le médiateur CDISC ODM de la solution REUSE.

5 Discussion-Conclusion

Nous avons modélisé le processus de mise en œuvre d'essais cliniques de promotion institutionnelle dans un Centre Hospitalo-Universitaire afin de concevoir, développer et expérimenter une solution d'intégration entre Dossier Patient Informatisé (DPI) et Système de Gestion d'Essais Cliniques (SGDC). Nous avons montré qu'il est possible d'exploiter les données cliniques du DPI dans le cadre de la recherche clinique grâce au transfert automatique de données administratives et médicales du DPI pertinentes pour la recherche clinique vers un SGDC. Les bénéfices pour le professionnel de santé consistent en l'économie d'une double saisie consommatrice de temps et génératrice d'erreur. Les bénéfices pour le patient sont liés à la conservation dans le DPI des données cliniques recueillies dans un contexte de recherche clinique susceptibles d'être nécessaires à la prise en charge diagnostique et thérapeutique du patient tels que les événements indésirables des traitements délivrés dans le cadre de l'essai clinique. Les bénéfices pour l'URC consistent à faciliter le contrôle de la qualité des données dans les cahiers d'observations cliniques. Le transfert automatique des données diminuant le risque d'erreur, le monitorage concerne essentiellement les données non issues du DPI.

L'architecture de la solution REUSE repose sur une extension du profil d'intégration IHE « RFD » que nous proposons et qui permet d'étendre les possibilités de réutilisation des données du DPI dans un contexte de recherche clinique.

L'obligation réglementaire concernant les SGDC impose que la validation des données cliniques du DPI saisies dans le cahier d'observation électronique s'effectue toujours au niveau du SGDC (indépendant de l'investigateur et du promoteur) et non du DPI. Ainsi, la solution REUSE doit fournir un appel contextuel de l'application du SGDC à partir du DPI au moment de la validation des données.

Une des limites actuelles de notre travail est d'assurer uniquement l'interopérabilité syntaxique entre deux applications (DxCare et MARVIN). En effet, dans le contexte de notre expérimentation, le formulaire de soin (des explorations fonctionnelles respiratoires) et le cahier d'observation électronique ont été paramétrés de manière simultanée en veillant

à ce que les items de l'étude « Dyspnée » soient définis de manière rigoureusement identique à ceux de DxCare (libellé(s), types de données, lien vers le dictionnaire de concepts médicaux). Ce matériel expérimental particulier a donc facilité la mise en œuvre de la solution d'alignement réalisée par le médicateur CDISC-ODM. En réalité, dans deux contextes différents (soin et recherche clinique), des items similaires sont souvent exprimés en utilisant des vocabulaires différents spécifiques du contexte. Dans une itération ultérieure de développement de REUSE, nous implémenterons dans le médiateur CDISC ODM une solution d'alignement sémantique entre les items des formulaires de soin et de recherche clinique. Par ailleurs, nous poursuivons l'implémentation de la solution REUSE en complétant l'implémentation des transactions du profil RIFD que nous proposons. Nous allons notamment développer le module d'intégration de cahier d'observation électronique ODM dans DPI.

Remerciements

Nous souhaitons tout d'abord remercier le personnel de l'HEGP en particulier de l'URC et du département d'informatique hospitalière de l'HEGP pour la convivialité de leur accueil et sans qui notre travail n'aurait pu être réalisé dans de si bonnes conditions. Par ailleurs, nous remercions tout particulièrement M. Nicolas De Saint Jorre et le Pr Joël Menard.

Références

[1] Kahn MG, Kaplan D. Configuration Challenges: Implementing Translational Research Policies in Electronic Medical Records. *Academic Medicine* 2007; 82(7): 661-9.

[2] Health Level Seven [URL: http://www.hl7.org] [Consulté le 17/11/2008].

[3] Clinical Data Interchange Standards Consortium [URL: http://www.cdisc.org].

[4] Object Management Group [URL: http://www.omg.org].

[5] *The Future Vision of Electronic Health Records as eSource for Clinical Research (Version 1.0).* Anonymous eds. *eClinical Forum / PhRMA EDC/eSource Taskforce Joint Team* September 14, 2006.

[6] Dusserre L, Ducrot H, Allaert FA. L'information médicale : l'ordinateur et la loi. *sl : Éditions médicales internationales* 1996; 24-40.

[7] Pansier FJ, Garay A. Le médecin, le patient et le droit. *Rennes : Éditions ENSP*, 1999; 75-6, 196.

[8] Société française de santé publique [URL : http://www.sfsp.fr] [Consulté le 17/11/2008].

[9] Conseil national de l'ordre des médecins [URL : http://www.conseil-national.medecin.fr] [Consulté le 17/11/2008].

[10] Picard S, Pellet J, Brulet JF, et Trombert B. Les aspects juridiques et éthiques de la protection des données issues du dossier médical informatisé et utilisées en épidémiologie : un point de la situation. *Santé Publique 2006*; 18 : 107-17.

[11] Commission nationale de l'informatique et des libertés : [URL : http://www.cnil.fr] [Consulté le 17/11/2008].

[12] Méthodologie de référence pour les traitements de données personnelles opérés dans le cadre des recherches biomédicales : [URL : http://www.cnil.fr/fileadmin/documents/declarer/mode_d-emploi/sante/MR-001.pdf] [Consulté le 22/11/2008]

[13] Food and Drug Administration [URL :http://www.fda.gov] [Consulté le 17/11/2008].

[14] Operational Data Model [URL: http://www.cdisc.org/models/odm/v1.3/index.html] [Consulté le 17/11/2008].

[15] Comité européen de normalisation, Technical Committee 251 [URL: http://cen.iso.org] [Consulté le 17/11/2008].

[16] Klein A, Prokosch HU, Müller M, Ganslandt T. Experiences with an Interoperable Data Acquisition Platform for Multi-centric Research Networks Based on HL7 CDA. *Methods of Information in Medicine* 2007; 46 5: 580-5.

[17] National Cancer Institute [URL: http://www.nci.nih.gov] [Consulté le 17/11/2008].

[18] Biomedical Research Integrated Domain Group (BRIDG) model [URL : http://www.bridgmodel.org] [Consulté le 17/11/2008].

[19] Souza T, Kush R, Evans J. Global Clinical Data Interchange standards are here.

[20] Fridsma DB, Evans J, Haastak S, Mead CN. The BRIDG Project: A Technical Report. *J Am Med Inform Assoc* 2008; 15: 130-7.

[21] Integrating the Healthcare Enterprise, IHE [URL: http://www.ihe.net] [Consulté le 17/11/2008].

[22] Healthcare Information and Management Systems Society, HIMSS. [URL: http://www.himss.org] [Consulté le 17/11/2008].

[23] RFD Supplement. Anonymous eds. IHE IT Infrastructure Technical Framework 2006.

[24] *CRD Supplement.* Anonymous eds. IHE Quality, Research and Public Health (QRPH) Technical Framework 2008.

[25] Enterprise Architect [URL : http://www.sparxsystems.com.au] [Consulté le 17/11/2008].

[26] JDOM [URL: http://www.jdom.org] [Consulté le 17/11/2008].

[27] Java API for XML Processing, SAX [URL: http://java.sun.com/j2se/1.4.2/docs/api/org/xml/sax/package-summary.html] [Consulté le 17/11/2008].

[28] Java API for XML Processing, DOM [URL: http://java.sun.com/j2se/1.4.2/docs/api/org/w3c/dom/package-summary.html] [Consulté le 17/11/2008].

[29] MedaSys [URL: http://www.medasys.com] [Consulté le 17/11/2008].

[30] XClinical [URL: http://www.XClinical.com] [Consulté le 17/11/2008].

Adresse de correspondance

Naji El Fadly
SPIM-INSERM UMR_S 872, Eq 20
Faculté de Médecine Paris V
15, rue de l'École de Médecine
75006 Paris – France

E-mail: nelfadly@yahoo.fr

Informatique et Santé, 2009 (17) : 239-250
© Springer – Verlag France

Prédire l'accès à la liste d'attente de transplantation rénale : comparaison de deux méthodes de fouille de données

Sahar Bayat[1], Marc Cuggia[1], Delphine Rossille[1], Luc Frimat[2]

[1]*INSERM U936, Université Rennes 1, IFR 140, Rennes, France*
[2]*EA 4003, Nancy université, France*

Abstract

The study compares the effectiveness of Bayesian networks versus Decision Trees for predicting access to the renal transplant waiting list in a French healthcare network. The data set consisted in 809 patients starting renal replacement therapy. The data were randomly devised in a training set (90 %) and a validation set (10 %). Bayesian networks and CART decision tree were built on the training set. Their predictive performances were compared on the validation set. The Age variable was found to be the most important factor for predicting registration on the waiting list in both models. Both models were highly sensitive and specific: sensitivity 90.0 % (95 % CI: 76.8-100), specificity 96.7% (95 % CI: 92.2-100). Moreover, the models were complementary since the Bayesian network provided a global view of the variables' associations while the decision tree was more easily interpretable by physicians. These approaches provide insights on the current care process. This knowledge could be used for optimizing the healthcare process.

Keywords

Bayesian network; Decision tree; access to the renal transplant waiting list.

1 Introduction

L'insuffisance rénale chronique (IRC), définie par une clairance de la créatinine (Clcr) inférieure à 60 mL/min, est due à une destruction irréversible du rein par une maladie chronique. Le degré de l'IRC est corrélé à l'intensité des lésions rénales. Sa progression s'étale sur plusieurs mois ou années. L'expression clinique de la maladie est tardive, puisqu'elle apparaît au stade préterminal (Clcr < 30 mL/min). Le stade terminal (Clcr < 15 mL/min) correspond à la destruction totale des reins. Afin de prolonger la vie du patient, il faut alors avoir recours aux techniques de suppléance de la fonction rénale : hémodialyse, dialyse péritonéale et transplantation rénale.

Par rapport à la dialyse, la transplantation rénale permet au patient de récupérer un état de santé proche de la normale avec une réinsertion sociale optimale, le tout à moindre coût. Elle représente la stratégie thérapeutique la plus performante et la plus efficiente[1, 2]. En particulier, le transplanté bénéficie d'une diminution rapide et significative du risque de morbi-mortalité [3]. Cependant, la réussite de la transplantation rénale nécessite une logistique complexe et l'enjeu est de pouvoir l'offrir au plus grand nombre de patients dans le contexte de pénurie des greffons. De plus, en France, comme dans l'ensemble des pays industrialisés, l'incidence et la prévalence de l'IRCT sont en augmentation [4].

Avant de pouvoir bénéficier d'une transplantation rénale, il faut avoir été inscrit sur une liste d'attente. L'inscription sur la liste a des déterminants spécifiques. Une fois inscrit sur la liste, le patient attend plus ou moins longtemps la transplantation à laquelle il n'accédera pas forcément.

Les études sur l'accès à la liste ont montré qu'il existe des variations importantes d'un pays à l'autre, voire d'une région à l'autre au sein d'un même pays [5-9]. Une récente comparaison européenne a démontré que le risque relatif d'accès à la transplantation rénale dans un délai de 4 ans après le début de l'IRCT (après ajustement sur l'âge, le sexe, l'année du début de traitement de suppléance et les pathologies associées) varie de 0,23 en Italie à 3,86 en Norvège [8]. Cette variation peut provenir du manque de référentiel consensuel pour le choix des candidats potentiels à la greffe rénale. Des recommandations européennes, canadiennes et américaines sur les critères d'inscription sur liste ont été publiées [10-12] ; mais elles ne sont pas appliquées [8-13]. Ceci s'explique à la fois par leur imprécision, par la faiblesse des preuves scientifiques ayant conduit à leur élaboration, mais surtout par les différences importantes dans l'organisation des soins des différents pays.

Plusieurs études ont été menées, pour la majorité d'entre elles, dans les pays anglo-saxons afin d'identifier les facteurs limitants de l'accès à liste d'attente et à la transplantation rénale. Elles ont identifié deux types de facteurs : ceux d'ordre strictement médical (l'âge élevé, le diabète, les maladies associées graves) et ceux non médicaux (le sexe féminin, la distance entre le domicile du patient et le centre de transplantation, la prise en charge par les cliniques privées et l'orientation tardive vers un néphrologue) [14-19].

À l'échelle de la France métropolitaine, il n'y a pas de recommandations sur le choix des candidats potentiels à la greffe rénale. Une étude restreinte à 3 centres hospitaliers universitaires en Rhône-Alpes a mis en évidence que les diabétiques et les personnes âgées de plus de 60 ans ont moins de chance d'être inscrits sur liste [18]. En 2005, nous avons mené une étude sur une cohorte de 1 495 adultes ayant démarré un traitement de suppléance au sein de l'ensemble des structures néphrologiques du réseau de soins NEPHROLOR (région Lorraine). Cette étude a démontré le rôle défavorable pour l'inscription sur liste de l'âge élevé et de certaines caractéristiques médicales : diabète, maladies cardio-vasculaires, dénutrition... mais aussi du suivi néphrologique initial dans une structure ne pratiquant pas l'activité de transplantation rénale [15]. Ces déterminants ont été identifiés en appliquant les méthodes biostatistiques classiques ainsi que les méthodes de fouille de données [20].

Après avoir identifié les déterminants de l'accès à la liste, l'objectif de cette étude est de prédire l'accès à la liste d'attente de transplantation rénale au sein de NEPHROLOR avec deux méthodes de fouille de données : les réseaux bayésiens et les arbres de décision et de comparer leur performance.

2 Matériel et méthodes

2.1 Réseau de soins NEPHROLOR

Le réseau de soins NEPHROLOR a été créé autour des besoins du patient insuffisant rénal chronique dans la région Lorraine en 2002. Il réunit les néphrologues exerçant dans 13 structures distinctes, réparties dans 5 hôpitaux publics, 2 cliniques privées, 3 établissements du secteur participant au service public hospitalier et 2 associations. Le seul service de transplantation de la région est situé au CHU de Nancy [21].

2.2 Sources de données

Les données étudiées sont issues des systèmes d'information du réseau NEPHROLOR et de l'Agence de la Biomédecine (CRISTAL) [22]. Le réseau NEPHROLOR dispose de deux systèmes d'information complémentaires fonctionnant en permanence sur la région Lorraine : le Registre REIN (Réseau Épidémiologie et Information en Néphrologie) [4] et DIVAT (Données Informatisées, Validées en Transplantation). Ces bases de données sont nominatives suite aux autorisations CNIL obtenues.

2.2.1 Registre REIN

Le système d'information principal du réseau NEPHROLOR est le registre REIN qui a pris la suite de l'étude EPIREL (EPidémiologie de l'Insuffisance REnale en Lorraine). EPIREL était une étude prospective dont l'objectif était d'évaluer le traitement de l'insuffisance rénale chronique terminale en Lorraine. Tous les patients consécutifs présentant une IRCT et ayant commencé un traitement de suppléance de la fonction rénale, du 15 juin 1997 au 1er janvier 2001, ont été inclus dans l'étude EPIREL [23].

Le registre REIN est un registre national de l'IRC terminale traitée par dialyse ou par transplantation rénale préemptive. Initié en 2001, il s'étend progressivement sur les différentes régions françaises pour couvrir l'ensemble du territoire. REIN a pour finalité de contribuer à l'élaboration et à l'évaluation de stratégies sanitaires visant à améliorer la prévention et la prise en charge de l'insuffisance rénale chronique.

Le dossier d'inclusion des patients dans REIN comporte des informations sur [4-23] :

- les caractéristiques socio-démographiques des patients ;
- la période relative à l'initiation du traitement : cause de l'insuffisance rénale (maladie initiale), données biologiques et activité du patient au début du traitement ;
- l'état du patient : données bio-cliniques, co-morbidités, handicaps, modalités de traitement, inscription sur la liste d'attente de transplantation.

Les patients sont également suivis, au moyen de l'enregistrement continu d'un certain nombre d'événements (changements de méthode ou de modalité de prise en charge en dialyse, transferts entre structures, sevrage, greffe rénale, décès) et d'un point annuel.

La jonction des bases de données EPIREL et REIN permet à la région Lorraine de disposer d'un système d'information comportant un enregistrement exhaustif des patients atteints d'IRCT débutant un traitement de suppléance, ainsi que de leur évolution depuis juillet 1997 [21].

2.2.2 Base de données DIVAT

La collaboration du centre de transplantation de Lorraine à DIVAT permet au réseau NEPHROLOR de disposer d'informations très détaillées sur les patients transplantés. La base de données DIVAT a été créée en 1990 à l'initiative du service de néphrologie du CHU de Nantes afin de regrouper les informations médicales des patients bénéficiaires de greffe rénale et/ou pancréatique. Actuellement, les CHU de Nantes, Nancy, Montpellier, Limoges, Toulouse et l'hôpital Necker sont reliés au système d'information DIVAT. DIVAT contient les informations sur le receveur, le donneur, les paramètres immunologiques, les traitements d'induction, les infections/complications post-greffe ainsi

que la survie du patient et du greffon. De plus, les paramètres de suivi sont recueillis à 3 mois, 6 mois et 1 an après la greffe puis tous les ans.

2.2.3 Système d'information CRISTAL

CRISTAL est le système d'information de l'Agence de la Biomédecine portant sur l'activité de transplantation réalisée sur le territoire français. Le patient est enregistré dans CRISTAL dès son inscription sur la liste d'attente de transplantation rénale. CRISTAL contient les informations suivantes : bilan médical du receveur à l'inscription sur la liste, suivi annuel avant la greffe, rapport de greffe (compte rendu des évènements péri-opératoires), suivi annuel post greffe ainsi que les informations sur le donneur et sur l'allocation des greffons [22]. Ainsi, CRISTAL nous permet de disposer des données sur le décès, la mise en contre-indication temporaire et la sortie de liste des patients en attente de greffe.

2.3 Population de l'étude

Il s'agit de patients :

- ayant débuté le traitement de suppléance de la fonction rénale entre le 1er juillet 1997 et le 30 juin 2003 (cas incidents) ;
- âgés de plus de 17 ans ;
- résidant dans la région Lorraine ;
- et ayant été pris en charge au sein de NEPHROLOR.

Ces patients ont été extraits du registre REIN. Pour identifier ceux d'entre eux inscrits sur la liste d'attente, avant ou après le début du traitement de suppléance, nous avons extrait de la base de données DIVAT et du système d'information CRISTAL la totalité des résidants lorrains inscrits sur la liste entre le 1er janvier 1996 et le 31 décembre 2004. La jonction des 3 fichiers extraits de ces 3 bases de données nous a permis d'identifier les patients inscrits sur la liste.

2.4 Données recueillies

Nous avons étudié 3 catégories de variables :

- les caractéristiques socio-démographiques : sexe, âge, distance entre le lieu de résidence et le centre de transplantation ;
- les données bio-cliniques au 1er traitement de suppléance : pathologies associées (maladies cardiovasculaires, diabète, insuffisance respiratoire, cirrhose, antécédent de cancer, trouble sévère du comportement, handicap physique), Indice de Masse Corporelle (IMC) (en 3 classes : < 20 ; 20-24,99 et $\geq$ 25kg/m²), albuminémie (en 3 classes : < 3 ; 3-3,49 et $\geq$ 3,5g/L), hémoglobine (en 2 classes : <11 et $\geq$ 11g/dL) ;
- les caractéristiques de prise en charge au sein de NEPHROLOR : prise en charge par le centre pratiquant la transplantation *versus* les autres centres et l'organisation de la prise en charge (publique *versus* privée).

2.5 Méthodes d'analyses

Les patients ayant des données manquantes n'ont pas été inclus dans cette analyse.

Nous avons donc réalisé cette étude sur 809 des 1495 patients inclus initialement dans l'étude. Cependant, les caractéristiques des 686 patients ayant des données manquantes et les autres patients ont été comparées. Il n'y avait pas de différence significative entre les caractéristiques sociodémographiques (sexe, âge, distance entre le lieu de résidence et le centre de transplantation) et les modalités de prise en charge au sein de NEPHROLOR (prise en charge par le centre pratiquant la transplantation, le statut public/privé de structure de prise en charge) des deux groupes. Toutefois, excepté le diabète, toutes les pathologies associées étaient plus fréquentes chez les patients n'ayant pas de données manquantes. Enfin les taux d'inscription sur la liste étaient similaires dans les deux groupes.

Il faut signaler que la majorité des données manquantes sont dues à un défaut d'exhaustivité des enregistrements (taille, poids...) durant l'étude EPIREL et avant la mise en place du registre REIN en 2001. Depuis l'exhaustivité des enregistrements s'est améliorée d'année en année.

Les patients ont été repartis aléatoirement dans 2 ensembles : 90 % dans l'ensemble d'apprentissage (les données servant à bâtir les modèles de prédiction) et 10 % dans l'ensemble de validation afin d'évaluer les performances des modèles de prédiction sur un échantillon indépendant [24]. Nous avons effectué une analyse avec le test de χ^2 pour vérifier la comparabilité de ces deux groupes.

2.5.1 Réseaux Bayésiens

Un réseau bayésien est un graphe orienté acyclique dans lequel les nœuds représentent les variables aléatoires et les flèches entre les nœuds représentent les relations entre les variables [25]. Ces flèches n'impliquent pas une relation causale. La signification intuitive d'une flèche entre le nœud parent et le nœud enfant est que le nœud parent influence directement le nœud enfant, ces influences sont quantifiés par des probabilités conditionnelles. Ainsi le graphe est associé à la distribution des probabilités conditionnelles des variables qu'il contient [26-28]. Une fois qu'un réseau bayésien a été construit, il peut être utilisé pour calculer la probabilité de survenue d'un évènement connaissant d'autres événements (inférence).

L'analyse du réseau bayésien a été réalisée en 3 étapes avec le logiciel BAYESIALAB Software (version 3.3 ; Bayesia SA) :

1. Découverte des associations entre les variables.
2. Analyse des associations afin de caractériser la variable d'intérêt : *l'inscription sur la liste d'attente* (le nœud cible).
3. Inférence pour étudier l'importance des nœuds en termes d'apport d'information sur la connaissance du nœud cible.

Le réseau obtenu a été ensuite utilisé pour prédire l'inscription sur la liste d'attente dans l'ensemble de validation. Le seuil de probabilité retenu pour classer les patients dans les groupes inscrits/non inscrits était de 0,50.

2.5.2 Arbres de décision

Un arbre de décision est à la fois un outil d'aide à la décision et d'exploration des données. C'est une méthode de classification et de prédiction représentée sous la forme d'une arborescence. L'arbre de décision est construit en partitionnant récursivement et le plus

efficacement possible l'ensemble d'apprentissage en sous-groupes les plus homogènes possibles, par des tests définis sur les variables descriptives. À chaque nœud ou test, une variable descriptive est sélectionnée parce qu'elle classifie au mieux les individus selon un critère de séparation donné. La construction de l'arbre est stoppée quand les conditions d'arrêt sont rencontrées. De par l'algorithme de construction, l'arbre permet d'identifier les variables descriptives les plus pertinentes pour la classification des individus, correspondant aux variables effectivement retenues à chaque nœud de l'arbre.

Les arbres de décision sont populaires, en particulier dans le domaine médical, car ils sont simples à construire et à interpréter du fait que tout arbre est équivalent à un ensemble de règles décisionnelles. Toutefois les arbres sont sensibles à toute variation de l'ensemble d'apprentissage et ils se construisent de manière séquentielle sans remettre en cause les tests précédemment choisis.

Plusieurs algorithmes de construction d'arbres de décision ont été proposés, les plus connus étant CHAID, CART, C5.0. Dans cette étude l'analyse est réalisée avec l'algorithme CART (*Classification And Regression Tree*) [29, 30] qui construit un arbre binaire basé sur l'indice de Gini. De plus, la construction de l'arbre s'arrête quand la profondeur maximale de 5 est atteinte (nombre maximal de niveaux entre le nœud racine et les feuilles). L'analyse par méthode CART a été réalisée avec le logiciel SPSS Clementine (version 10.1).

Les performances des modèles de prédiction (réseau bayésien et arbre de décision CART) sur le groupe de validation ont été mesurées en termes de sensibilité, de spécificité et de valeur prédictive positive. Nous avons effectué le test de McNemar pour étudier la concordance de ces 2 modèles de prédiction.

3 Résultats

3.1 Caractéristiques des patients

L'âge moyen des 809 patients inclus dans l'étude était de 62,1 ± 14,2 ans. Six patients sur dix étaient des hommes (59,6 %). 279 (34,5 %) patients étaient atteints de diabète, 90 (11,1 %) d'insuffisance respiratoire et 358 (44,2 %) de maladies cardiovasculaires. 114 (14,1 %) patients avaient un antécédent de cancer, 19,5 % souffraient d'un handicap physique et 5,9 % d'un trouble sévère du comportement. 212 d'entre eux (26,2 %) ont été inscrits sur la liste d'attente.

Les 809 patients inclus ont été repartis aléatoirement dans 2 groupes : 90 % (729 patients) dans l'ensemble d'apprentissage et 10 % (80 patients) dans l'ensemble de validation. Il n'y avait pas de différence significative entre les caractéristiques de base des patients des 2 groupes.

3.2 Réseaux bayésiens

La figure 1 montre le réseau bayésien construit à partir du groupe apprentissage. L'inscription sur la liste d'attente est associée directement avec le diabète, les maladies cardiovasculaires, l'insuffisance respiratoire, l'âge, l'albuminémie et la prise en charge par le centre pratiquant la transplantation. L'importance relative des variables en termes d'apport d'information sur la connaissance de « Inscription sur la liste d'attente » est la suivante : âge : 1,000 ; maladies cardiovasculaires : 0,3582 ; diabète : 0,2287 ; albuminémie : 0,1221 ; insuffisance respiratoire : 0,0815 ; prise en charge par le centre pratiquant la transplantation : 0,0749 ; handicap physique : 0,0413 ; IMC : 0,0138 ;

distance entre le lieu de résidence et le centre de transplantation : 0,0108 ; structure de prise en charge (publique/privée) : 0,0037.

Le réseau a été évalué sur l'ensemble de validation. Parmi les 80 patients du groupe validation, 20 avaient été inscrits sur la liste. Le réseau bayésien a prédit correctement 18 inscriptions sur 20 et 58 non-inscriptions sur 60. La sensibilité, la spécificité et la valeur prédictive positive du modèle pour la prédiction de l'inscription sur la liste étaient respectivement de : 90,0 % (IC à 95 % : 76,8-100), 96,7 % (IC à 95 % : 92,2-100) et 90,0 % (IC à 95 % : 76,8-100).

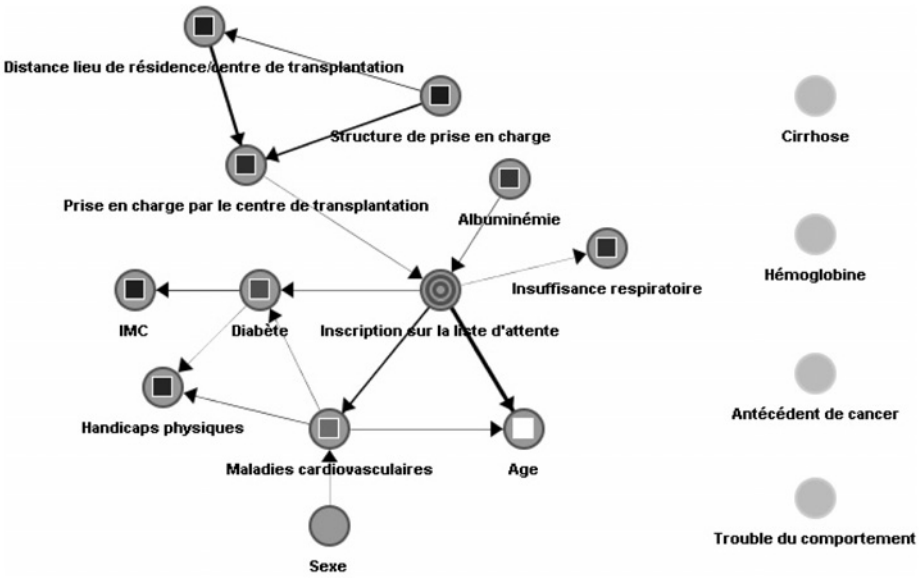


Figure 1 - Réseau bayésien construit à partir du groupe apprentissage.

3.3 Arbre de décision

La figure 2 montre l'arbre de décision CART généré à partir du groupe apprentissage. L'âge était la variable la plus discriminante pour la prédiction de l'inscription sur la liste d'attente. Le modèle CART a prédit correctement 18 inscriptions sur 20 et 58 non-inscriptions sur 60. Les performances de l'arbre de décision ont été évaluées sur l'ensemble de validation. La sensibilité, la spécificité et la valeur prédictive positive du modèle pour la prédiction de l'inscription sur la liste d'attente étaient respectivement de : 90,0 % (IC à 95 % : 76,8-100), 96,7 % (IC à 95 % : 92,2-100) et 90,0 % (IC à 95 % : 76,8-100).

Le test de McNemar n'a pas mis en évidence de différence significative entre les prédictions du réseau bayésien et de l'algorithme CART. Les prédictions des deux modèles étaient discordantes pour 2 patients. L'indice Kappa de concordance entre les prédictions du réseau bayésien et de l'arbre de décision CART était de 0,93.

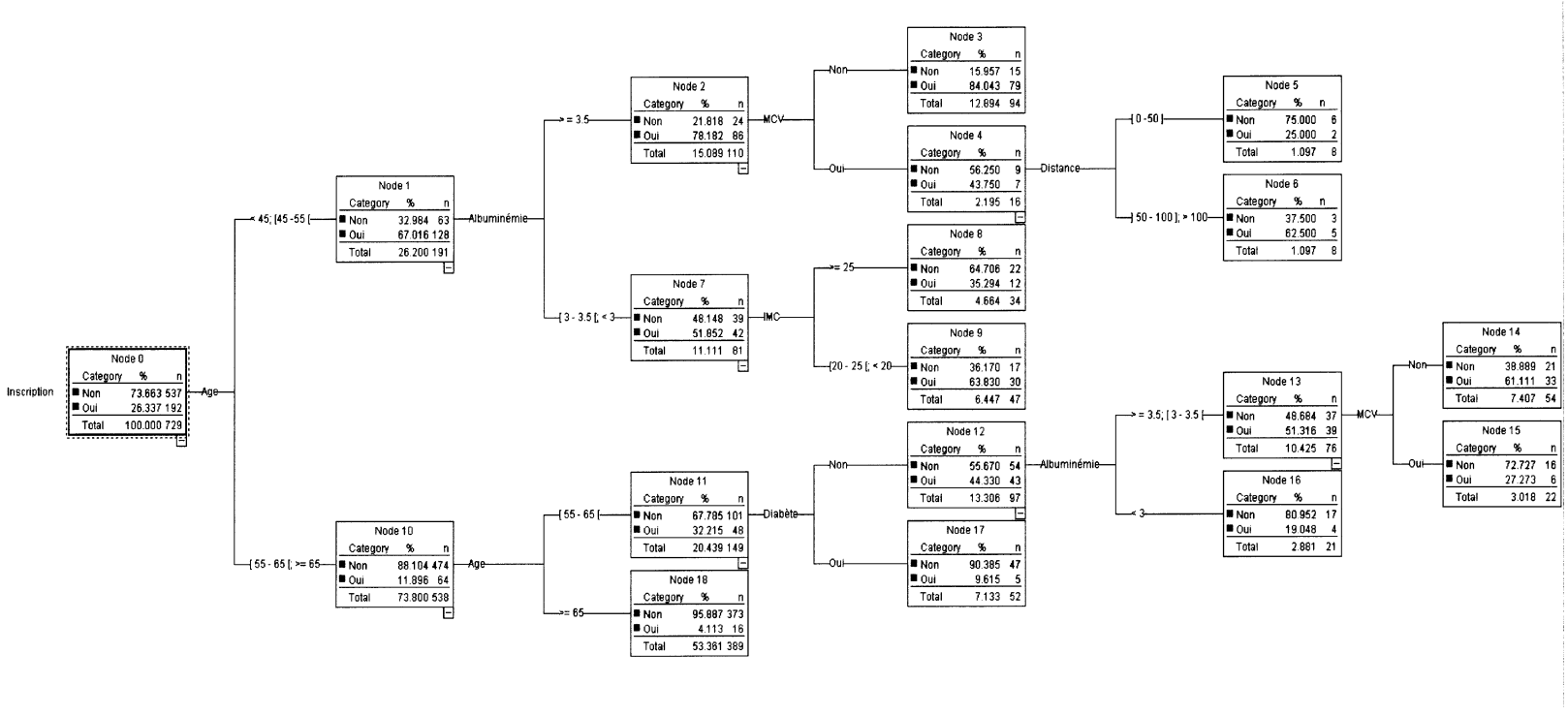


Figure 2 - Arbre de décision CART construit à partir du groupe apprentissage (MCV : maladies cardiovasculaires).

4 Discussion

Cette analyse montre que le réseau bayésien et l'arbre de décision CART sont deux approches intéressantes et performantes pour prédire l'accès à la liste d'attente de transplantation rénale au sein de NEPHROLOR. Pang *et al.* ont comparé les performances de différentes méthodes de fouille de données incluant les réseaux bayésiens et les arbres de décision pour prédire le pronostic des patients atteints d'un trauma crânien [29- 31]. Cette étude menée sur 513 patients a démontré la meilleure performance des arbres de décision. Chapman et al. ont comparé la performance des réseaux bayésiens et des arbres de décision avec les règles définies par les experts dans la cadre du diagnostic de pneumonie sur 292 patients. Par rapport au réseau bayésien, l'arbre de décision était moins sensible et plus spécifique [32]. Contrairement à ces études, les prédictions faites par les deux modèles étaient très concordantes dans notre étude. L'*âge* était la variable la plus importante pour la prédiction de l'inscription sur la liste d'attente dans les deux modèles. Toutefois, les modèles diffèrent sur les facteurs discriminants sélectionnés pour la prédiction de l'inscription sur la liste d'attente. Les variables communes aux deux modèles sont : les *maladies cardiovasculaires*, le *diabète* et l'*albuminémie*. Le réseau bayésien intègre en plus *l'insuffisance respiratoire*, et *la prise en charge par le centre pratiquant la transplantation*, tandis que l'arbre de décision CART utilise l'*IMC* et *la distance entre le lieu de résidence et le centre de transplantation*.

Janssens *et al.* ont montré que le réseau bayésien est plus performant que l'arbre de décision. Dans notre étude, le réseau bayésien et l'arbre de décision CART ont des performances élevées et similaires. Janssens *et al.* ont également montré que le réseau bayésien est plus adapté aux situations complexes d'aide à la décision [33]. Il peut notamment prendre en compte toutes les relations d'interdépendance entre les variables. Dans notre étude, le réseau bayésien a permis de visualiser de nouvelles associations entre les variables, telles que l'association entre *la distance entre le lieu de résidence et le centre de transplantation*, *la structure publique/privée de prise en charge* et *la prise en charge par le centre pratiquant la transplantation*. Cette association peut être expliquée par le fait que la majorité des patients habitant près du centre de transplantation sont pris en charge par ce centre qui est par ailleurs une structure publique. Cette association montre que bien que deux variables diffèrent entre les modèles du réseau bayésien et de l'arbre de décision : *la prise en charge par le centre pratiquant la transplantation* et *la distance entre le lieu de résidence et le centre de transplantation*, ces variables sont en fait fortement liées.

Néanmoins, le réseau bayésien peut associer plusieurs variables, directement et indirectement, d'une façon complexe, ce qui rend l'interprétation difficile. Par exemple, dans notre étude, le *diabète* et les *maladies cardiovasculaires* sont associés entre eux ainsi qu'avec l'inscription sur la liste d'attente. De plus ils sont également associés au *handicap physique* et à l'*IMC*, influençant indirectement l'inscription sur la liste d'attente. En comparaison, les règles décisionnelles dérivées de l'arbre de décision sont simples et facilement interprétables par les cliniciens. Pour cette raison, les arbres de décision sont populaires dans le milieu médical.

Les résultats des deux modèles peuvent être utilisés pour optimiser le processus d'inscription sur la liste d'attente au sein de NEPHROLOR. L'intégration de ces algorithmes dans le système d'information REIN permettrait de calculer la probabilité d'inscription sur la liste d'attente de chaque nouveau patient au moment de son enregistrement. Ceci pourrait inciter les néphrologues à informer les patients sur la transplantation et à commencer les évaluations pré-greffe plus tôt. Le processus de sélection des candidats potentiels à la greffe rénale pourrait être ainsi homogénéisé.

5 Conclusion

Nous avons utilisé deux méthodes de fouille de données afin de prédire l'accès à la liste d'attente de transplantation rénale au sein du réseau de soins NEPHROLOR. Les deux modèles étaient très performants et leurs prédictions proches. Il est important de noter que ces deux approches sont très complémentaires. Le réseau bayésien produit une vue globale des associations entre les variables étudiées tandis que les règles de décision extraites de l'arbre de décision sont simples et plus faciles à interpréter.

Ces approches permettent de calculer la probabilité d'inscription sur la liste d'attente, en fonction des caractéristiques de base du patient. Leur utilisation au moment de l'enregistrement initial du patient dans le système d'information de l'insuffisance rénale chronique terminale traitée, *i.e.* le registre REIN, pourrait dynamiser et optimiser le processus d'inscription sur la liste d'attente et améliorer l'accès à la greffe rénale.

Références

[1] Rabbat CG, Thorpe KE, Russell JD, Churchill DN. Comparison of mortality risk for dialysis patients and cadaveric first renal transplant recipients in Ontario, Canada. *J Am Soc Nephrol* 2000; 11(5): 917-922.

[2] Wolfe RA, Ashby VB, Milford EL, Ojo AO, Ettenger RE, Agodoa LY *et al.* Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *N Engl J Med* 1999; 341(23): 1725-1730.

[3] Meier-Kriesche HU, Schold JD, Srinivas TR, Reed A, Kaplan B. Kidney transplantation halts cardiovascular disease progression in patients with end-stage renal disease. *Am J Transplant* 2004; 4(10): 1662-1668.

[4] Couchoud C, Stengel B, Landais P, Aldigier JC, de Cornelissen F, Dabot C *et al.* The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrol Dial Transplant* 2006; 21(2): 411-418.

[5] Fritsche L, Vanrenterghem Y, Nordal KP, Grinyo JM, Moreso F, Budde K *et al.* Practice variations in the evaluation of adult candidates for cadaveric kidney transplantation: a survey of the European Transplant Centers. *Transplantation* 2000; 70(10): 1492-1497.

[6] Ramos EL, Kasiske BL, Alexander SR, Danovitch GM, Harmon WE, Kahana L et al. The evaluation of candidates for renal transplantation. The current practice of U.S. transplant centers. *Transplantation* 1994; 57(4): 490-497.

[7] Satayathum S, Pisoni RL, McCullough KP, Merion RM, Wikstrom B, Levin N et al. Kidney transplantation and wait-listing rates from the international Dialysis Outcomes and Practice Patterns Study (DOPPS). *Kidney Int* 2005; 68(1): 330-337.

[8] Stel VS, van Dijk PC, van Manen JG, Dekker FW, Ansell D, Conte F et al. Prevalence of co-morbidity in different European RRT populations and its effect on access to renal transplantation. *Nephrol Dial Transplant* 2005; 20(12): 2803-2811.

[9] Tonelli M, Klarenbach S, Manns B, Culleton B, Hemmelgarn B, Bertazzon S et al. Residence location and likelihood of kidney transplantation. *Cmaj* 2006; 175(5): 478-482.

[10] The EBPG expert group on renal transplantation. European best practice guidelines for renal transplantation (part 1). *Nephrol Dial Transplant* 2000; 15 (Supplement 7).

[11] Kasiske BL, Cangro CB, Hariharan S, Hricik DE, Kerman RH, Roth D *et al.* The evaluation of renal transplantation candidates: clinical practice guidelines. *Am J Transplant* 2001; 1 Suppl 2: 3-95.

[12] Knoll G, Cockfield S, Blydt-Hansen T, Baran D, Kiberd B, Landsberg D *et al.* Canadian Society of Transplantation: consensus guidelines on eligibility for kidney transplantation. *Cmaj* 2005; 173(10): S1-25.

[13] Kiberd B, Boudreault J, Bhan V, Panek R. Access to the kidney transplant wait list. *Am J Transplant* 2006; 6(11): 2714-2720.

[14] Alexander GC, Sehgal AR. Why hemodialysis patients fail to complete the transplantation process. *Am J Kidney Dis* 2001; 37(2): 321-328.

[15] Bayat S, Frimat L, Thilly N, Loos C, Briancon S, Kessler M. Medical and non-medical determinants of access to renal transplant waiting list in a French community-based network of care. *Nephrol Dial Transplant* 2006; 21(10): 2900-2907.

[16] Garg PP, Frick KD, Diener-West M, Powe NR. Effect of the ownership of dialysis facilities on patients' survival and referral for transplantation. *N Engl J Med* 1999; 341(22): 1653-1660.

[17] Oniscu GC, Schalkwijk AA, Johnson RJ, Brown H, Forsythe JL. Equity of access to renal transplant waiting list and renal transplantation in Scotland: cohort study. *Bmj* 2003; 327(7426): 1261.

[18] Villar E, Rabilloud M, Berthoux F, Vialtel P, Labeeuw M, Pouteil-Noble C. A multicentre study of registration on renal transplantation waiting list of the elderly and patients with type 2 diabetes. *Nephrol Dial Transplant* 2004; 19(1): 207-214.

[19] Winkelmayer WC, Glynn RJ, Levin R, Mittleman MA, Pliskin JS, Avorn J. Late nephrologist referral and access to renal transplantation. *Transplantation* 2002; 73(12): 1918-1923.

[20] Bayat S, Cuggia M, Kessler M, Briancon S, Le Beux P, Frimat L. Modelling access to renal transplantation waiting list in a French healthcare network using a Bayesian method. *Stud Health Technol Inform* 2008; 136: 605-610.

[21] Frimat L, Kessler M. Prise en charge des transplantés rénaux au sein d'un réseau de soins - l'exemple de NEPHROLOR (Lorraine). *Le courrier de la transplantation* 2004; 4.

[22] Strang WN, Tuppin P, Atinault A, Jacquelinet C. The French organ transplant data system. *Stud Health Technol Inform* 2005; 116: 77-82.

[23] Kessler M, Frimat L, Panescu V, Briancon S. Impact of nephrology referral on early and midterm outcomes in ESRD: EPidemiologie de l'Insuffisance REnale chronique terminale en Lorraine (EPIREL): results of a 2-year, prospective, community-based study. *Am J Kidney Dis* 2003; 42(3): 474-485.

[24] Guyon I. *A scaling law for the validation-set training-set size ratio.* AT&T Bell Laboratories, Berkely, California 1997.

[25] Goldszmidt M, Friedman N. *Learning Bayesian Network from Data.* SRI International 1998.

[26] Jouffe L. Nouvelle classe de méthodes d'apprentissage de réseaux bayésiens. *Journées francophones d'Extraction et de Gestion des Connaissances (EGC)*, 2002.

[27] Jouffe L, Munteanu P. New Search Strategies for Learning Bayesian Networks. *Proceedings of Tenth International Symposium on Applied Stochastic Models, Data Analysis*, 2001.

[28] Munteanu P, Bendou M. The EQ Framework for Learning Equivalence Classes of Bayesian Networks. *First IEEE International Conference on Data Mining (IEEE ICDM)*, 2001.

[29] Breiman L, Friedman JH, Ohlsen RA, Stone CJ. *Classification and regression trees*. Belmont: Wadsworth International Group 1984.

[30] Lewis RJ. An introduction to classification and regression tree (CART) analysis. *Annual Meeting of the Society for Academic Emergency Medicine*, San Francisco, CA 2000.

[31] Pang BC, Kuralmani V, Joshi R, Hongli Y, Lee KK, Ang BT et al. Hybrid outcome prediction model for severe traumatic brain injury. *Journal of neurotrauma* 2007; 24(1): 136-146.

[32] Chapman WW, Fizman M, Chapman BE, Haug PJ. A comparison of classification algorithms to automatically identify chest X-ray reports that support pneumonia. *Journal of biomedical informatics* 2001; 34(1): 4-14.

[33] Janssens D, Wets G, Brijs T, Vanhoof K, Timmermans HJP, Arentze TA. Improving the performance of a multi-agent rule-based model for activity pattern decisions using Bayesian networks. In: *Electronic Conference Proceedings of the 83rd Annual Meeting of the Transportation Research Board Journal of the Transportation Research*, 2004.

Adresse de correspondance

Sahar Bayat

Département d'Information Médicale

CHU de Rennes, Pontchaillou

2, rue Henri Le Guilloux

35033 Rennes cedex, France

Sahar.bayat@chu-rennes.fr

Informatique et Santé, 2009 (17) : 251-262
© Springer – Verlag France

Approche sémantique pour l'identification homogène d'effets indésirables et de médicaments dans huit bases de données de patients : une contribution au projet européen eu-ADR

Paul Avillach[1,2,3], Fleur Mougin[1], Michel Joubert[3,4], Frantz Thiessard[1,2], Antoine Pariente[5,6], Jean-Charles Dufour[3,4], Marius Fieschi[3,4]

[1]*LESIM - ISPED, Université Victor Segalen Bordeaux 2, France*
[2]*Service d'Informatique Médicale, CHU de Bordeaux, France*
[3]*LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille, France*
[4]*Service de Santé Publique et d'Information Médicale, Hôpital Timone, AP-HM, Marseille, France*
[5]*Département de pharmacologie, Université Victor Segalen Bordeaux 2, France*
[6]*Service de Pharmacologie, CHU de Bordeaux, France*

Abstract

The overall objective of the eu-ADR project is the design, development, and validation of a computerised system that exploits data from electronic health records and biomedical databases for the early detection of adverse drug reactions. Eight different databases, containing health records of more than 30 million European citizens, are involved in the project. Unique queries cannot be performed across different databases because of their heterogeneity: Medical record and Claims databases, four different terminologies for coding diagnoses, and two languages for the information described in free text. The aim of our study was to provide database owners with a common basis for the construction of their queries. Using the UMLS, we provided a list of medical concepts, with their corresponding terms and codes in the four terminologies, which should be considered to retrieve the relevant information for the events of interest from the databases.

Keywords

Terminology; UMLS; Semantic Interoperability ; Database extraction.

1 Introduction

Les informations médicales recueillies dans le cadre du suivi clinique peuvent être réutilisées pour une grande variété de finalités annexes : médico-économique, épidémiologique [1], générations d'alertes [2, 3]. Ces informations, colligées à toutes les étapes du processus de soin, souvent exprimées en texte libre, sont souvent codées dans une ou plusieurs terminologies médicales spécifiques [4]. La sélection d'un code approprié pour représenter une information médicale donnée est chronophage et a l'avantage de préciser sans ambiguïté le sens de cette information. Le codage de l'information permet son traitement automatisé et facilite l'interopérabilité sémantique entre systèmes d'information. Les informations médicales dont la sémantique a été précisée par un codage approprié peuvent être transmises, interprétées et traitées par différents systèmes, permettre ainsi un partage et une réutilisation des données entre systèmes d'information.

Prenons l'exemple de la déclaration d'effets indésirables médicamenteux en France. Elle est spontanée et basée sur le volontariat des cliniciens. Elle est transmise aux centres de pharmacovigilance pour être codée dans des terminologies conçues pour décrire les effets indésirables de médicament : MedDRA [5] ou WHO-ART [6]. Ce processus de recueil est loin d'être optimal puisqu'on estime que seulement 4 % des effets indésirables sont ainsi déclarés [7-10]. La recherche de tels événements, en scrutant les bases de données médicales disponibles au sein d'une institution (locale, régionale, nationale, voire internationale), est une solution envisageable pour améliorer le taux de recueil des effets indésirables. Des travaux antérieurs, dans un cadre différent, ont exploré le regroupement des effets indésirables similaires grâce à des définitions formelles issues de SNOMED CT [11]. Les expérimentations existantes n'explorent qu'une seule source à la fois [12, 13]. Or les informations pertinentes pour la détection d'effets indésirables nouveaux ou déjà connus peuvent provenir de sources de nature différente pour un même patient et peuvent être recherchées chez de multiples patients au travers de bases situées dans des pays différents qui, assemblées, permettent d'augmenter la puissance de la détection.

C'est dans ce cadre que s'inscrit le projet européen eu-ADR. Son objectif est la conception, le développement, et la validation d'un système informatisé qui exploite les données de dossiers de santé électroniques et de bases de données biomédicales pour la détection précoce de vingt-trois effets indésirables de médicaments [14]. Chacune des bases de données utilisées comporte des informations codées dans des terminologies spécifiques et possède des caractéristiques propres dépendant de l'objectif initial et de la fonction locale qu'elle assure (médico-économique, soins, audit, etc.). Du fait de cette hétérogénéité (structurelle et sémantique) des bases en présence, la construction d'une requête unique et réutilisable *in extenso* sur les différentes bases pour effectuer la même recherche d'effets indésirables est impossible.

L'objectif de ce travail est de fournir une méthode pour extraire les informations pertinentes concernant les effets indésirables étudiés et les médicaments associés au sein de différentes bases de données. Le projet eu-ADR explore 8 bases de données provenant de différents pays. Notre travail inclut la recherche d'une cohérence du processus d'amélioration de notre méthode d'extraction entre les différents acteurs du projet.

2 Matériel et méthodes

Des terminologies différentes sont utilisées pour coder les événements dans les différentes bases de données, un socle commun a été nécessaire afin d'harmoniser les requêtes. L'objectif était de fournir aux administrateurs de bases de données la liste des concepts et des termes associés qu'ils devraient impérativement utiliser afin de récupérer les informations pertinentes concernant les événements d'intérêts.

2.1 Les bases de données

Les huit bases de données utilisées dans le projet eu-ADR contiennent des informations issues de dossiers médicaux de plus de 30 millions de citoyens européens. Les caractéristiques de ces huit bases sont différentes à plusieurs niveaux (Tableau 1). En particulier, divers types de données sont stockés : données cliniques, données de prescriptions, etc. Selon les bases, les données médicales peuvent concerner des adultes ou des enfants ou encore rassembler indifféremment les deux populations. Par ailleurs, chaque base de données utilise ses propres terminologies pour coder les événements rencontrés et les médicaments prescrits lors de la prise en charge du patient dans leur structure médicale. Enfin, certaines bases stockent ce type d'information en texte libre.

2.2 Les terminologies utilisées dans le projet eu-ADR

Il existe différents types de systèmes terminologiques (notamment les classifications, les nomenclatures) [15]. Pour simplifier, nous parlerons par la suite de « terminologie ».

2.2.1 Terminologies des évènements

Les terminologies médicales sont structurées sous forme de listes de concepts[1], généralement organisées en hiérarchie. Un concept peut être désigné de multiples façons, puisque les termes[2] le désignant peuvent provenir de différentes langues. Chaque langue peut utiliser des termes distincts synonymes ou quasi synonymes pour le même concept.

Dans le projet eu-ADR, quatre terminologies sont utilisées (Tableau 1).

La « classification statistique internationale des maladies et des problèmes de santé connexes » est développée et maintenue par l'Organisation mondiale de la santé (OMS)[3].

- La 9e révision, ou ICD9 [16], à l'origine peu appropriée pour le codage en général et plus particulièrement pour certaines spécialités [17, 18], a été complétée par le Centre américain des statistiques médicales (USNCHS). La version finale, dénommée ICD9-CM [19], est composée de 19 chapitres et contient plus de 21 000 concepts hiérarchisés.

- La 10e révision ou ICD10 [20] est la dernière révision de la classification internationale des maladies, composée de 21 chapitres. Elle contient plus de 12 000 concepts également organisés de manière hiérarchique. Les codes associés à chaque concept sont différents dans les révisions ICD9-CM et ICD10.

Tableau 1 - Description des huit bases de données.

Database	Terminology		Text libre	Type de données*	Patients†
	Event	Drug			
Pedianet – Italia (ITA)	ICD9-CM	ATC	oui (ITA)	DPI	E
Health Search (ITA)	ICD9-CM	ATC	oui (ITA)	DPI	A/E
Lombardy Regional DB (ITA)	ICD9-CM	ATC	non	RSS, D	A/E
Tuscany Regional - ARS (ITA)	ICD9-CM	ATC	non	RSS, D, L, M	A/E
IPCI – Netherlands (NL)	ICPC	ATC	oui (NL)	EHR	A/E
PHARMO (NL)	ICD9-CM	ATC	non	RSS, P,L, M	A/E
QRESEARCH United Kingdom (UK)	RCD	BNF/ ATC	non	DPI	A/E
Aarhus University Hospital DB (DK)	ICD10	ATC	non	RSS, D, L, M	A/E

*DPI : Dossier Patient Informatisé, RSS : Résumés Standardisés de Sortie, D : dispensation, L : Laboratoire, M : Mortalité, P : Prescription ; †E : Enfant, A : Adulte

- La classification internationale des soins primaires ICPC [21] a été publiée en 1987 par l'organisation mondiale des médecins de famille (WONCA). Elle permet le classement des données concernant les patients et l'activité clinique dans les domaines de la médecine

[1] Un concept est une unité de pensée [ISO 5963].

[2] Un terme est la désignation d'un concept défini dans une langue par son expression linguistique [ISO 1087].

[3] http://www.who.int/en/

générale et des soins de santé primaires. L'ICPC contient près de 7 400 concepts et a une structure bi-axiale avec 17 chapitres sur un axe, et 7 composantes sur l'autre : 1) symptômes et plaintes ; 2) dépistage et prévention ; 3) traitements ; 4) procédures et médicaments ; 5) résultats de tests de laboratoire ; 6) administration ; 7) diagnostic et 8) maladies et autres.

- La classification READ CODE ou RCD [22] est un standard de classification des données de médecine générale qui contient plus de 82 000 concepts. Cette terminologie décrit non seulement les maladies mais aussi l'histoire clinique et les symptômes, les signes et examens cliniques, les procédures, les médicaments et les appareils, ainsi que les activités et les informations sociales. Elle est exclusivement utilisée au Royaume-Uni.

2.2.2 Terminologies utilisées pour les médicaments dans le projet eu-ADR

Sept bases de données utilisent la classification Anatomical Therapeutic Chemical ou ATC[4] [23]. Le principe de l'ATC est de classer les médicaments en différents groupes selon l'organe ou le système sur lequel ils agissent et selon leurs propriétés chimiques, pharmacologiques et thérapeutiques. Cette classification est un standard international pour le codage des médicaments.

Dans la dernière base, QRESEARCH, les médicaments sont codés en British National Formulary (BNF) [24]. Il faut noter qu'il existe des correspondances, réalisées par cette équipe, entre les codes ATC et BNF.

Ainsi, pour interroger les huit bases de données, il est possible d'utiliser la classification ATC sans qu'il soit nécessaire d'effectuer un travail de mise en correspondance particulier comme cela sera le cas avec les classifications codant les effets indésirables.

2.3 Un système d'unification des terminologies médicales : l'UMLS

L'Unified Medical Language System® (UMLS)® [25, 26] est un système d'intégration de terminologies biomédicales intégrant plus de 150 terminologies. Les quatre terminologies utilisées dans le projet eu-ADR sont incluses dans l'UMLS. Le Métathesaurus®, composant principal de l'UMLS, est un vocabulaire central, une base de donnée [26, 27], constitué d'environ 1,5 million de concepts reliés par plus de 3,75 millions de relations hiérarchiques. Un concept UMLS, identifié par un identifiant unique (*Concept Unique Identifier*, CUI), décrit une notion médicale unique qui peut être exprimée avec différents synonymes (termes) dans des langues différentes. Par exemple, le concept UMLS *Haematemesis* (CUI : C0018926) a pour code « 578.0 » dans la CM-ICD9, « K92.0 » dans la ICD10, « D14 » dans la CISP et « J680 » dans la RCD. Certains termes en texte libre sont les suivants : « Ematemesi/vomito sanguinolento » en italien, « Bloed; braken » en néerlandais, « Haematemesis/vomiting blood » en anglais, etc. La figure 1 illustre la hiérarchie entre les concepts de l'UMLS, et présente la représentation du concept *Kidney Failure*[5], *Acute* (*insuffisance rénale aiguë)*, et l'un de ses « fils » dans la hiérarchie *Kidney Tubular Necrosis, Acute* (*nécrose tubulaire aiguë)*.

[4] http://www.whocc.no/atcddd/

[5] Cette police de caractères sera utilisée pour l'écriture des concepts UMLS

2.4 Méthode de projection des concepts définissant un évènement dans les terminologies

2.4.1 Création de la liste des évènements

Une liste de vingt-trois événements, identifiés comme ceux principalement impliqués dans les réactions indésirables aux médicaments, a été créée. Les événements inclus dans cette liste fermée seront ceux utilisés pour la suite du projet eu-ADR. Parmi ces 23 événements, cinq ont été identifiés comme devant être étudiés en priorité [14] : éruption cutanée bulleuse, choc anaphylactique, infarctus aigu du myocarde, rhabdomyolyse et insuffisance rénale aiguë. L'élaboration de la méthode et des procédures employées pour la projection des concepts dans les terminologies dans le cadre du projet eu-ADR a été réalisée par l'étude de ces 5 événements. Pour tester et améliorer la méthode, nous avons également étudié l'événement « hémorragie digestive haute » (UGIB) dont la définition médicale complexe amène des difficultés pour sa recherche dans les bases de données.

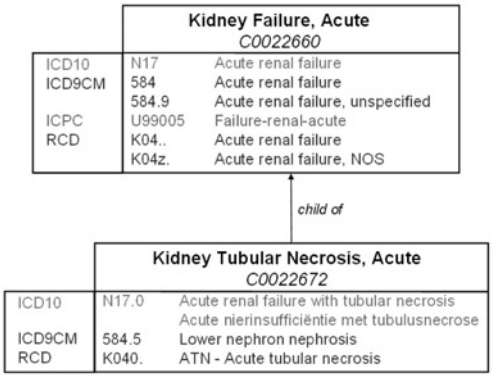


Figure 1 - Représentation du concept Kidney Failure Acute *et un de ses* fils Kidney Tubular Necrosis, Acute *dans l'UMLS.*

Après avoir décrit la méthode proposée pour effectuer la projection d'un événement, nous exposons en détail les différentes étapes pour effectuer la projection des concepts définissant l'UGIB dans les quatre terminologies.

2.4.2 Principe de création d'un socle sémantique commun pour réaliser les requêtes

Une requête unique n'est pas réalisable pour extraire des informations des bases de données utilisées car, nativement, 4 terminologies différentes sont présentes. Nous avons élaboré un socle sémantique commun identique pour les huit bases. Les éléments de ce socle commun sont des concepts de l'UMLS (regroupent les termes de différentes terminologies ayant le même sens médical) et non des termes.

La définition précise de chaque événement doit tenir compte de son expression dans les différentes sources et doit être validée par des experts. Pour chacun des événements, nous avons recherché tous les concepts UMLS sous-tendus par les différents termes de la définition. Le socle commun des requêtes est ainsi constitué par les concepts de l'UMLS et leur projection dans les quatre terminologies, et ce dans les différentes langues employées dans les bases de données étudiées.

2.4.3 Processus pour réaliser la projection des concepts dans les terminologies

Un groupe de travail a été constitué. Il comprenait des experts en informatique médicale, des experts en pharmacologie et pharmacovigilance, et des administrateurs des bases de données. Six phases distinctes ont été identifiées pour réaliser cette tache :

1. Élaboration d'une fiche définissant l'événement d'intérêt (construite à partir des critères cliniques et biologiques de diagnostic de l'événement).
2. Création d'une première projection des concepts dans les terminologies par recherche des concepts UMLS représentant l'événement d'intérêt à partir de la fiche.
3. Présentation des concepts UMLS (ainsi que les codes dans les terminologies associées) pour examen par tous les partenaires impliqués dans eu-ADR.
4. Harmonisation des avis au sein du groupe de travail sur les mises en correspondance des concepts dans les différentes terminologies.
5. Création d'une seconde version de la projection des concepts dans les terminologies, soumis pour validation à tous les partenaires.
6. Diffusion de la version finale de la projection des concepts dans les terminologies à tous les administrateurs de bases de données.

3 Résultats

3.1 Élaboration d'une fiche définissant l'évènement par les experts en pharmacologie

Pour la recherche d'information et afin d'améliorer la sensibilité des extractions, nous avons retenu une définition « large » pour l'élaboration de la fiche de définition de l'événement UGIB. Cette définition a été établie en fonction de manuels de référence clinique et validée par des spécialistes en gastro-entérologie. La définition était composée de critères d'inclusion, liés à des symptômes ou des diagnostics associés à l'événement et de critères d'exclusion.

Suite aux commentaires des administrateurs des bases de données, la structure de la fiche de définition de l'événement (EDF) a été simplifiée et les critères d'exclusion ont été supprimés car pouvant conduire à l'élimination d'événements potentiellement intéressants. Ainsi, la définition finale de l'EDF pour UGIB est uniquement fondée sur l'inclusion de critères : *Upper gastrointestinal haemorrage*, *Oesophagal haemorrage*, *Gastrointestinal haemorrhage*, *Bleeding from peptic ulcer*, *Haematemesis/blood vomiting et Melaena.*

Ces critères ont été explorés au cours de la projection des concepts dans les terminologies.

3.2 Création d'une première projection des concepts dans les terminologies

Pour chaque expression littérale (ou libellé) correspondant aux critères d'inclusions de la définition de l'événement, nous avons effectué une recherche automatisée, à l'aide du Knowledge Source Server (UMLSKS)[6] et les installations locales de l'UMLS (version 2008AA), pour identifier le concept UMLS (figure 2) et tous les codes reliés au concept dans les terminologies du projet : ICD9-CM, ICD10, ICPC et RCD.

Lorsque cette recherche automatisée ne permettait pas d'identifier un terme correspondant à un concept dans l'une des terminologies étudiées, nous avons effectué une recherche manuelle dans la terminologie concernée afin d'identifier les termes qui pourraient être correctement liés à ce concept. Les termes identifiés par le biais de cette procédure

[6] http://kswebp2.nlm.nih.gov/UMLSKS

manuelle ont été rajoutés à ceux qui avaient été produits par la recherche automatisée afin de compléter la liste des termes pertinents.

Cette liste de termes a permis aux administrateurs des différentes bases de données d'exécuter des requêtes syntaxiquement différentes mais ayant un sens identique, en utilisant les termes appropriés pour coder/exprimer les concepts UMLS communs recherchés dans chacune des différentes bases. Toutes les observations faites après la constitution et l'exécution de leur première requête (sur UGIB) ont été prises en compte pour affiner la liste des concepts et réitérer les requêtes au sein des différentes bases.

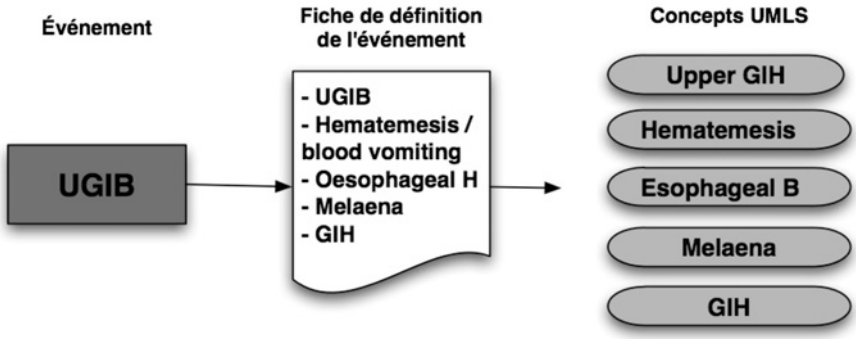


Figure 2 - Identification des concepts UMLS pour l'événement « hémorragie digestive haute » (UGIB).

3.3 Processus de critique et d'amélioration

Le forum Internet mis en place dans le cadre du projet (accès restreint aux membres du consortium) a été un outil très utile pour discuter de la pertinence des concepts proposés à l'étape précédente. Chaque événement avait sa propre section sur le forum. Le forum permet facilement l'affichage de toutes les discussions sur une même page Internet pour chaque événement, garantissant ainsi une visualisation plus globale et une meilleure interactivité (supérieure à celle qui est obtenue par l'usage du courrier électronique) entre les membres. Plus de 300 messages ont été postés sur le forum en moins de 6 semaines.

Ainsi, par exemple, de nouveaux concepts, recherchés *in fine* dans l'UMLS, ont été proposés et discutés : *Acute gastric ulcer with hemorrhage*, *Acute gastric ulcer with hemorrhage AND perforation, Acute duodenal ulcer with hemorrhage, Acute duodenal ulcer with hemorrhage AND perforation, Acute peptic ulcer with hemorrhage, Acute peptic ulcer with hemorrhage and perforation, Acute gastric mucosal erosion, Atrophic gastritis, with haemorrhage, Other specified gastritis, with hemorrhage, Unspecified gastritis and gastroduodenitis, with hemorrhage.* L'ajout de concepts correspondant aux codes ICD9-CM 531-4.1* (*i.e. Gastric ulcer*, *Duodenal ulcer*, *Peptic ulcer*, *site unspecified*, *Gastrojejunal ulcer*: *Acute *with perforation** (*with or without obstruction*) : soit les concepts UMLS : *Acute gastric ulcer with perforation, Acute duodenal ulcer with perforation, Acute peptic ulcer with perforation)* a également été proposé.

L'examen final de la pertinence de chacun de ces concepts potentiels a été réalisé par le groupe de travail sur la base des discussions postées sur le forum. L'harmonisation a été réalisée par le biais de conférences téléphoniques et de réunions plénières. Concernant l'événement UGIB, tous les nouveaux concepts suggérés ont été conservés.

3.4 Diffusion de la version finale de la projection des concepts dans les terminologies à tous les administrateurs de bases de données

Afin d'harmoniser les requêtes, nous avons établi un tableau de tous les concepts UMLS liés aux critères d'inclusion proposés dans l'EDF. Les codes et libellés des termes correspondants dans les quatre terminologies utilisées par les huit bases de données ont également été fournis (un extrait est présenté dans le tableau 2).

Tableau 2 - Extrait de la projection des concepts dans les quatre terminologies pour l'événement hémorragie digestive haute.

C0041909		**Upper gastrointestinal**
RCD	X30Be	Upper gastrointest haemorrhage
C0018926		**Hematemesis**
ICD10	K92.0	Haematemesis
ICD9CM	578.0	Hematemesis
ICPC	D14	Hematemesis/vomiting blood
ICPC2EDUT	D14	Haematemesis
ICPC2EENG	D14	Haematemesis/vomiting blood
ICPC2ICD10DUT	MTHU011334	bloed; braken
ICPC2ICD10DUT	MTHU012496	braken; bloed
ICPC2ICD10ENG	MTHU011334	blood; vomiting
ICPC2ICD10ENG	MTHU012496	vomiting; blood
ICPC2ICD10ENG	MTHU033134	hematemesis
ICPC2P	D14001	vomiting out blood
ICPC2P	D14001	Vomiting; blood
ICPC2P	D14003	Haematemesis
ICPC2P	D14003	Hematemesis
ICPCDUT	D14	Haematemesis/bloed braken
ICPCITA	D14	Ematemesi/vomito
RCD	J680.	Haematemesis
RCD	J680.	Vomiting blood
C0025222		**Melena**
ICD9CM	578.1	Melena
ICD10	K92.1	Melaena
ICPC	D15	Melena/black tarry stools
ICPC2EDUT	D15	Melaena
ICPC2EENG	D15	Melaena
ICPC2ICD10ENG	MTHU048082	melena
ICPC2P	D15003	Melaena
ICPC2P	D15003	Melena
ICPC2P	D15004	Stool(s); tarry
ICPC2P	D15004	tarry stool
ICPC2P	D15005	black stool
ICPC2P	D15005	Stool(s); black
ICPCDUT	D15	Melaena/zwarte ontlasting
ICPCITA	D15	Melena/feci nere picee
RCD	X76fM	Black faeces symptom
RCD	XE0bI	Alt blood passed per rectum
RCD	XE0bI	Altered blood in stool
RCD	XE0bI	Altered blood passed per
RCD	XE0bI	Black, tarry stool
RCD	XE0bI	Melaena
RCD	XE0rH	Black faeces

La création des requêtes a fait l'objet de recommandations aux administrateurs des bases de données : chaque concept listé devait obligatoirement être présent dans leur requête par l'intermédiaire des codes et/ou termes (libellé associé au code) issus des différentes

terminologies. Les codes et/ou termes choisis devaient être séparés par l'opérateur booléen « OR » afin de récupérer tous les dossiers ayant trait à au moins un des critères d'inclusion. Les termes en anglais, italien et néerlandais sont fournis en plus des codes car certaines bases de données stockent une partie de l'information en texte libre.

La liste des codes et des termes que nous avons fournie n'était pas restrictive : les administrateurs des bases de données étaient libres d'y ajouter tous les termes qu'ils jugeaient pertinents pour retrouver l'événement "UGIB" dans leur base de données, à l'unique condition que ces termes soient une nouvelle façon de décrire les concepts sélectionnés. En particulier, même si nous n'avons pas inclus dans la liste tous les concepts affiliés à un concept d'intérêt, les administrateurs de bases devaient le faire si cela leur paraissait nécessaire. Ainsi, lorsqu'un code donné a des fils (c'est-à-dire une description plus précise), la requête devait également inclure les descendants de ce code jugés pertinents pour récupérer les données.

Un exemple de requête effectuée (avant harmonisation et donc rajout d'autres concepts) pour rechercher les hémorragies gastro-intestinales supérieures est :

(*Oesophageal haemorrhage* OR *Upper gastrointestinal haemorrhage* OR *Haematemesis/blood vomiting* OR *Melaena* OR *Gastrointestinal haemorrhage*)

4 Discussion

Le processus mis en place a permis d'aboutir à un consensus concernant les concepts choisis pour la projection dans les terminologies des cinq événements d'intérêts jugés prioritaires. Ceci a permis de proposer une liste de termes incluant leurs codes et leurs libellés pour uniformiser les requêtes et donc les extractions à partir des huit bases de données participant au projet eu-ADR. Dans le cadre de l'événement hémorragie digestive haute (UGIB), une première version a conduit à proposer 5 concepts identifiés initialement. Le processus de discussion et d'harmonisation a permis de rajouter de nouveaux concepts à cette liste pour aboutir à un total de 21 concepts potentiellement utilisés pour le codage de cet événement dans les bases de données.

Ces bases de données étaient hétérogènes de par la terminologie utilisée, de par la présence ou non de données en texte libre (utilisé dans deux langues : italien et néerlandais), et de par le type de données qu'elles contiennent (médico-administratives et dossier patient informatisé). L'UMLS permet la mise en relation des bases de données hétérogènes [20] et participe à la réalisation d'une interopérabilité sémantique entre ces différentes bases.

Plusieurs limites du processus ont cependant été mises en évidence au cours de l'étude des cinq autres événements sélectionnés. La première concerne l'utilisation des résultats des tests de laboratoires, disponibles dans certaines bases de données, pour l'identification de certains événements (par exemple, une insuffisance rénale aiguë). Il s'agit là de l'identification d'un concept non par sa dénomination ou sa place dans une description nosologique mais par sa signature dans des résultats biologiques. Notre approche ne répond pas à cette question. La seconde concerne les différences de règles de codage entre les différentes classifications. Chaque base peut définir une granularité différente de codage. L'ATC peut, par exemple, être codé par l'utilisateur avec 5 ou 7 caractères en fonction de degrés de précision souhaitée. Le niveau d'information saisi n'est donc pas toujours identique. L'analyse des résultats de l'extraction des bases de données devra tenir compte de ces éléments.

5 Conclusion

La projection des concepts UMLS dans les terminologies et les ajustements manuels ont été validés dans les quatre terminologies utilisées dans notre étude. Ces terminologies sont toutes les quatre intégrées à l'UMLS. Ceci a permis de fournir un socle sémantique commun pour l'élaboration des requêtes adaptées aux huit bases de données hétérogènes que nous avons exploitées. La liste des concepts accompagnée de la liste de codes associés et leurs libellés en texte libre a été utilisée par les administrateurs des bases de données pour construire les requêtes afin d'interroger leur base en utilisant la terminologie adaptée. Cette méthode a été utilisée avec le même succès pour les 5 autres événements et sera utilisée pour les 17 événements restants du projet. Les extractions des mêmes concepts médicaux, dans les huit bases, ont permis aux biostatisticiens du projet, de disposer de bases de données comparables quant à la définition des événements recherchés malgré la grande hétérogénéité entre les bases.

Remerciements

Ces recherches ont reçu un financement de la communauté de l'Union européenne dans le cadre du FP7/2007-2013 convention de subvention n ° 215847 - the eu-ADR project.

Nous remercions les membres du consortium eu-ADR : http://193.144.6.8/partners , en particulier ceux du service de pharmacologie de l'Université de Bordeaux 2 et ceux de l'institut NEUROLESI de l'Université de Médecine de Messine.

Les auteurs remercient la NLM pour la mise à disposition gratuite de l'UMLS.

Références

[1] Goldberg M. Les bases de données d'origine administrative peuvent-elles être utiles pour l'épidémiologie ? *Rev Epidemiol Sante Publique* 2006; 54(4): 297-303.

[2] Cimino JJ. Collect once, use many. Enabling the reuse of clinical data through controlled terminologies. *J Ahima* 2007;78(2): 24-9.

[3] Giannangelo K. Making the connection between standard terminologies, use cases and mapping. *Him J* 2006; 35(3): 8-12.

[4] Hayrinen K, Saranto K, Nykanen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. *Int J Med Inform* 2008; 77(5): 291-304.

[5] Brown EG, Wood L, Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf* 1999; 20(2): 109-17.

[6] WHO. The WHO Adverse Reaction Terminology – WHO-ART. 2005 [cited 2008 Nov 19th]; Available from: http://www.umc-products.com/graphics/3149.pdf

[7] Pouyanne P, Haramburu F, Imbs JL, Begaud B. Admissions to hospital caused by adverse drug reactions: cross sectional incidence study. French Pharmacovigilance Centres. *BMJ* 2000; 320(7241): 1036.

[8] Lacoste-Roussillon C, Pouyanne P, Haramburu F, Miremont G, Begaud B. Incidence of serious adverse drug reactions in general practice: a prospective study. *Clin Pharmacol Ther* 2001; 69(6): 458-62.

[9] Begaud B, Martin K, Haramburu F, Moore N. Rates of spontaneous reporting of adverse drug reactions in France *Jama* 2002; 288(13): 1588.

[10] Moride Y, Haramburu F, Requejo AA, Begaud B. Under-reporting of adverse drug reactions in general practice. *Br J Clin Pharmacol* 1997; 43(2) :177-81.

[11] Alecu I, Bousquet C, Mougin F, Jaulent MC. Mapping of the WHO-ART terminology on Snomed CT to improve grouping of related adverse drug reactions. *Studies in health technology and informatics* 2006; 124: 833-8.

[12] Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J Am Med Inform Assoc* 2003; 10(2): 115-28.
[13] Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. *J Am Med Inform Assoc* 2005; 12(4): 448-57.
[14] Trifiro' G, Pariente A, Polimeni G, Miremont-Salame G, Catania MA, Salvo F, et al. Data Mining on Large Health Record Databases for Detecting Adverse Reactions: Which Events to Monitor? *Drug Saf* 2008; 31(10): 885.
[15] de Keizer NF, Abu-Hanna A, Zwetsloot-Schonk JH. Understanding terminological systems. I: Terminology and typology. *Methods Inf Med* 2000; 39(1): 16-21.
[16] WHO. International Classification of Diseases. 9th Revision. 1975 [cited 2008 nov 19]; Available from: http://www.cdc.gov/nchs/about/otheract/ICD9-CM/abtICD9-CM.htm
[17] Kurtzke JF. ICD 9: a regression. *Am J Epidemiol* 1979; 109(4): 383-93.
[18] White KL. Restructuring the International Classification of Diseases: need for a new paradigm. *J Fam Pract* 1985; 21(1): 17-8, 20.
[19] Commission of Professional and Hospital Activities. International Classification of Diseases. Ninth Revision with Clinical Modifications (ICD9-CM). Ann Arbor; 1978.
[20] WHO. International Classification of Diseases, manual of the International Statistical Classification of diseases, injuries and causes of death: 10th revision. 1993 [cited 2008 nov 19]; Available from: http://www.who.int/classifications/apps/icd/icd10online/
[21] Lamberts H, Wood M, eds. ICPC: *International Classification of Primary Care. Oxford:* Oxford University Press 1987.
[22] O'Neil M, Payne C, Read J. Read Codes Version 3: a user led terminology. *Methods Inf Med* 1995; 34(1-2): 187-92.
[23] Miller GC, Britt H. A new drug classification for computer systems: the ATC extension code. *Int J Biomed Comput* 1995; 40(2): 121-4.
[24] New additions in the BNF 1976-78. *Drug Ther Bull* 1976; 14(24): 93-6.
[25] Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004; 32(Database issue): D267-70.
[26] Lindberg DA, Humphreys BL, McCray AT. The Unified Medical Language System. *Methods Inf Med* 1993; 32(4): 281-91.
[27] Humphreys BL. The 1994 Unified Medical Language System knowledge sources. *Health Libr Rev* 1994; 11(3): 200-3.

Adresse de correspondance

Paul Avillach

Laboratoire d'Epidémiologie, Statistique et Informatique Médicales (LESIM), ISPED

Université Victor Segalen Bordeaux 2, 146 rue Léo-Saignat

33076 Bordeaux cedex, France

E-mail : Paul.Avillach@isped.u-bordeaux2.fr

Informatique et Santé, 2009 (17) : 263-274
© Springer – Verlag France

Constitution d'un référentiel probabiliste d'aide au codage

Grégoire Thillay[1], Laurent Lecornu[1,3], Clara Le Guillou[2,3], Pierre-Jean Garreau[2], Vanessa Hery[2] et Jean Michel Cauvin[2,3]

[1]*INSTITUT TELECOM ; TELECOM Bretagne UEB ; Dpt ITI, Brest, France*
[2]*CHU Brest, Département d'Information Médicale, Brest, France*
[3] *INSERM, U650, IFR 148 ScInBioS, Brest, France*

Abstract

Because giving a code to something is a non-intuitive operation, it leads us towards committing mistakes, which may have some bad consequences. In medicine, physicians and such have to assign a code for everything they do and they are not spared with these kinds of errors. This report is about a way to help the medical coding task in order to minimize error without losing time by giving some codes that may occur in consequences of what the practionner is doing and of his patient characteristics. It describes different medical information sources in order to have these predicted codes and tries to determine which one is better. In order to achieve this objective, a method based on the probability to have a diagnostic given a specific medical information (which is already presents in the pmsi) has been developed.

Keywords

PMSI; coding prediction; Medical Informatics.

1 Introduction

Durant ces dernières décennies, avec le développement d'outils informatiques de plus en plus performants, les prédictions ont pris une part de plus en plus importante dans la vie de l'Homme, car pouvant être réalisées à l'aide de méthodes rigoureuses et scientifiques donc reproductibles et relativement fiables. La prédiction comportementale peut aider le domaine médical notamment dans le codage des maladies. En effet, depuis la réforme des hôpitaux (ordonnance du 24/04/1994 sur la réforme de l'hospitalisation), chaque acte et chaque diagnostic se doivent d'être codés et associés à un patient. Cependant, à la vue des ressemblances existantes dans ces codes et des innombrables autres origines possibles de mauvais codage si aucune aide n'est proposée au médecin codeur, chaque hôpital se voit exposé à de nombreuses erreurs pouvant conduire à des statistiques complètement faussées et, plus grave, à des défauts de crédits le menant à un bilan déficitaire.

Le projet présenté dans cet article propose d'aider les médecins codeurs à coder les différents diagnostics en prédisant des codes probables à partir de connaissances issues d'une base de codage régionale anonymisée et d'informations contenues dans le résumé standardisé de sortie de leurs patients (âge, sexe, actes déjà codés…). Ce travail recouvre un aspect du projet TECSAN/ANR intitulé Med1dex qui se propose de développer une plate-forme d'aide au codage combinant plusieurs modules spécifiques utilisant chacun une source d'information médicale et une méthode d'aide adaptée.

Après avoir présenté le contexte de cette étude, la méthode développée qui a permis de prédire les codes diagnostics sera introduite. La conclusion sera précédée des résultats obtenus.

2 Contexte

2.1 Présentation

Cet article s'inscrit dans le cadre de la création d'outils d'aide au codage des actes et des pathologies. Le but des travaux présentés ici est d'indiquer au médecin une liste ordonnée de codes (avec leurs libellés) de diagnostics classés selon leurs pertinences respectives en fonction des diagnostics, actes, unité médicale, sexe, âge et autres caractéristiques du patient contenues dans le RUM. Le médecin choisit parmi ces codes ceux qui correspondent aux diagnostics qu'il désire coder.

2.1.1 Rappel sur le PMSI

Le Programme de Médicalisation des Systèmes d'Information ou PMSI est un dispositif faisant partie de la réforme du système de santé français qui permet de mesurer l'activité et les ressources des établissements grâce à des informations quantifiées et standardisées [1]. Dans l'optique de ce travail, une seule partie du fonctionnement du PMSI est considérée: la partie Médecine Chirurgie Obstétrique (MCO). Cette partie se base sur deux constituantes : Le Résumé de Sortie Standardisé (RSS) et le Groupe Homogène de Malade (GHM).

Le RSS correspond au séjour d'un patient à l'hôpital. Il est constitué d'un (ou plusieurs) Résumé d'Unité Médicale (RUM) qui contient obligatoirement les informations du patient (âge, sexe...) ainsi qu'un diagnostic principal qui a fait l'objet de la plus importante consommation de ressources. Il peut aussi contenir un diagnostic relié ainsi qu'un ou plusieurs diagnostics associés. Lorsque l'on synthétise le RSS à partir de ce ou de ces RUM, un diagnostic sera établi comme principal, sur la même règle explicitée précédemment, avec le diagnostic relié associé s'il existe puis des diagnostics associés qui seront désignés sur la base des autres RUM avec un ajout de qualificatif : « significatifs » s'ils ont consommé des ressources ou « documentaires » dans le cas contraire. Le tableau 1 décrit le format du RUM. Enfin, les RSS sont anonymisés et deviennent alors des Résumés de Sortie Anonymisés (RSA) pour transmission à l'Agence Régionale de l'Hospitalisation (ARH) et ce en conformité avec la Commission Nationale de l'Informatique et des Libertés (CNIL).

La CCAM (Classification Commune des Actes Médicaux) est une nomenclature française destinée à coder les gestes pratiqués par les médecins, gestes techniques dans un premier temps puis, par la suite, les actes intellectuels cliniques. Cette classification sert à établir :

- en médecine libérale, les honoraires des actes techniques réalisés lors des consultations ;
- dans les cliniques privées, les honoraires pour les interventions réalisées ;
- dans les hôpitaux publics, le PMSI et sa tarification des séjours hospitaliers transmis à l'assurance maladie dans le cadre de la tarification à l'activité (T2A).

Le codage CCAM se fait selon un code composé de quatre lettres suivies de trois chiffres (figure 1). Chaque lettre possède une signification permettant ainsi de connaître un minimum le contexte de l'acte technique effectué par le médecin.

Chaque série de 3 chiffres, quant à elle, permet de singulariser chacun des actes ne pouvant être différenciés par les quatre lettres du codage [2, 3].

Tableau 1 - Description du format du RUM.

Intitulé		position	taille
Numéro FINESS		1-9	9
Version du format du RUM (011)		10-12	3
N° de RSS	(Equivalent de HOSP-PMSI sauf si IPP)	13-19	7
N° Administratif local de séjour		20-39	20
Date de naissance		40-47	8
Sexe		48-48	1
Numéro de l'unité médicale		49-52	4
Type d'autorisation de l'unité médicale		53-54	2
Type d'autorisation du lit dédié		55-56	2
Zone réservée		57-57	1
Date d'entrée dans l'unité médicale		58-65	8
Mode d'entrée dans l'unité médicale		66-66	1
Provenance (si mode d'entrée est mutation, transfert ou domicile)		67-67	1
Date de sortie de l'unité médicale		68-75	8
Mode de sortie de l'unité médicale		76-76	1
Destination (si mode de sortie est mutation, transfert ou domicile)		77-77	1
Code postal de résidence (ou 99 suivi du code Insee du pays pour les patients résidant hors de France)		78-82	5
Poids du nouveau-né à l'entrée de l'unité médicale (en grammes)		83-86	4
Nombre de séances		87-88	2
Nombre de diagnostics associés significatifs (nDAS) dans ce RUM		89-90	2
Nombre de données à visée documentaire (nDAD) dans ce RUM		91-92	2
Nombre de zone d'actes (nZA) dans ce RUM		93-94	2
Diagnostic principal (DP)		95-102	8
Diagnostic relié (DR)		103-110	8
IGS 2		111-113	3
Zone réservée		114-128	15
DAS n° 1			8
.....			* 8
DAS n° nDAS			8
DAD n° 1			8
.....			* 8
DAD n° nDAD			8
Zone d'acte n° 1	Date de réalisation		8
	Code CCAM		7
	Phase		1
	Activité		1
	Extension documentaire		1
	Modificateurs		4
	Remboursement exceptionnel		1
	Association non prévue		1
	Nombre d'exécutions de l'acte n° 1 pendant le séjour		2
.....			*26
Zone d'acte n° nZa	Date de réalisation		8
	Code CCAM		7
	Phase		1
	.....		8
	Nombre d'exécutions de l'acte n° nZA pendant le séjour		2

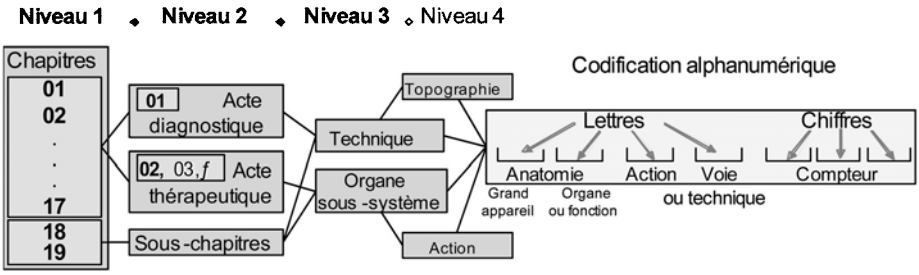


Figure 1 - Structure hiérarchique et codification de la CCAM.

L'appellation complète de la CIM10 est « Classification statistique internationale des maladies et des problèmes de santé connexes » et il s'agit de la 10e révision [4]. La CIM permet le codage des maladies, des traumatismes et de l'ensemble des motifs de recours aux services de santé. Elle est publiée par l'Organisation mondiale de la santé (OMS) et est utilisée à travers le monde principalement pour enregistrer les causes de morbidité et de mortalité ainsi qu'organiser les services de santé. La CIM10 rassemble près de 15 000 entrées réparties dans 22 chapitres. Les affections (symptômes, maladies, lésions traumatiques, empoisonnements) et les autres motifs de recours aux soins y sont répertoriés avec une précision qui dépend de leur importance. Le codage d'un diagnostic dépend directement du chapitre auquel il appartient mais chaque caractère du code ne possède pas une signification donnée comme pour la CCAM. Les codes diagnostiques sont composés d'une lettre suivie de 2 à 4 chiffres. Ce sont particulièrement les codes CIM10 que l'on ambitionne de prédire pour aider le médecin codeur.

2.1.2 Problématique

L'utilisation de plusieurs milliers de codes différents via la CCAM ou la CIM10 codant chaque fait et geste d'un médecin et de son patient conduit inévitablement à des erreurs de codage non négligeables. Ajouté à cela que de telles erreurs sont courantes et peuvent représenter plus de 20 % (voire jusqu'à 50 %) des codages effectués, il apparaît donc urgent de trouver un dispositif permettant de limiter ces erreurs afin d'améliorer la qualité du codage. Il faut aussi préciser que ces erreurs de codages sont plus souvent du côté du diagnostic que du côté de l'acte lui-même. C'est donc dans cette optique de correction d'erreur qu'est née l'idée de créer un outil aidant le médecin à coder les diagnostics qu'il constate.

La méthode d'aide au codage présentée ici s'appuie uniquement sur les informations déjà codées dans le RUM et sur une base de Résumés de Séjour Anonymisés régionale regroupant un peu plus d'un million de RSA (tableau 2) à partir de laquelle il a été extrait une base de connaissance qui nous permet de prédire les codes diagnostics les plus pertinents.

Une telle base de cas peut sembler plus que conséquente mais il faut savoir qu'il existe 7 619 actes différents dans la CCAM, 18 794 diagnostics différents dans la CIM10. En ajoutant quelques éléments tels que l'âge, le sexe, l'IGS2 ou encore la durée de séjour, cela fait plus de 1 000 milliards de combinaisons différentes !

Afin de réduire le nombre de combinaisons, l'âge a été regroupé en une vingtaine de tranches d'âges, le code CCAM a été réduit aux sous-paragraphes, et seules les trois premières lettres de diagnostics issus de la CIM10 (famille diagnostique) seront prédites.

La réduction à la famille diagnostique a été demandée par notre expert codeur qui voulait que le codeur fasse son choix sur l'ensemble des diagnostics d'une famille diagnostique. Même avec cette réduction, le nombre de combinaisons différentes est supérieur à 500 millions.

Afin de prédire les différents diagnostics possibles, il a été choisi, après discussion avec l' expert codeur, d'utiliser l'âge, le sexe, la durée du séjour, le service ou l'unité fonctionnelle où se trouve le patient, les actes codés et les diagnostics déjà codés.

Tableau 2 - Étude du nombre de RSA, du nombre de diagnostics et d'actes par RSA dans la base régionale utilisée.

Nombre de RSA	1 068 449
Diagnostics codés	3 052 079
Nombre de diagnostics par RSA	1 à 50
Moyenne et écart type par RSA	2,86 ± 2,24
Actes codés	2 769 008
Nombre d'actes différents par RSA	0 à 132
Moyenne et écart type par RSA	2,63 ±3,90

3 État de l'art

La problématique de ce projet s'inscrit donc dans le domaine de la prédiction où le nombre d'attributs prédictifs et à prédire varie suivant le cas étudié. Ce paragraphe dresse un bref état de l'art des méthodes de prédiction existantes.

Abraham B. et Ledolter J [5] divisent les méthodes de prédiction en deux catégories : les méthodes quantitatives et les méthodes qualitatives.

Les méthodes qualitatives sont intuitives et reposent pour l'essentiel sur une supposition subjective qui peut dépendre de données passées ou non. Habituellement, ce type de méthode ne peut être reproduit par quelqu'un d'autre que celui qui effectue la prédiction car ce dernier ne spécifie pas explicitement comment sa prédiction se fait. Bien que ce genre de prédiction s'avère relativement peu rigoureux, elle peut se révéler assez appropriée dans certains cas, mais elle ne l'est clairement pas dans le notre car c'est précisément ces méthodes subjectives qui entraînent des erreurs de codage de diagnostic chez les médecins. Le deuxième type de méthode, les méthodes quantitatives, est basé sur des modèles statistiques ou/et mathématiques [6] . Une fois le modèle sous-jacent choisi, les prédictions correspondantes sont automatiquement déterminées et sont donc complètement reproductibles par n'importe qui.

D.A. Zighed *et al.* présentent dans [7] différentes méthodes d'explication et de prédiction. Ils indiquent que ces méthodes ont pour objectif de trouver un modèle qui permet de faire le lien entre des attributs prédictifs et des attributs à prédire. Ce modèle est produit pas le biais d'une base de données (dit base d'apprentissage), dans notre cas, une base de RSA régionale. On distingue alors une phase « inductive » ou d'apprentissage qui établit le lien entre attribut suivie d'une phase « prédictive » qui indique la (ou les) valeur(s) prédite(s) [7] présente de manière synthétique les principales familles de méthodes d'explication et de prédiction comme par exemple les graphes d'induction, les réseaux de neurones, les méthodes de régression, l'analyse discriminante, les réseaux bayésiens, les règles d'association ainsi que d'autres approches prédictives (CN2 ou SVM).

Soit une population d'individus ou d'objets, notée Ω concernés par le problème d'apprentissage. À cette population est associé un attribut particulier appelé « attribut endogène » noté C. À chaque individu $?$ peut être associée sa classe $C(?)$. C peut prendre ses valeurs dans un ensemble quelconque pouvant être discret ou continu.

$$C : \Omega \rightarrow \Gamma$$

$$\omega \ ? \ C(\omega)$$

Dans notre cas, nous associons à ω non pas une seule classe mais un ensemble de classes qui correspond à l'ensemble des diagnostics. À chacune des classes, il est associé un degré de pertinence qui permet de la classer par ordre d'importance.

Le choix d'une méthode s'avère ici difficile car d'une part, le nombre d'actes codés dans le RSA et utilisé dans le processus d'aide au codage varie et d'autre part, le nombre de codes diagnostiques à coder dépend de la complexité du cas et du médecin codeur.

4 Méthode d'aide au codage proposée

4.1 Théorie

Pour rappel, il faut trouver un moyen d'établir une liste de résultats, du plus plausible au moins plausible, afin de donner un choix facile à l'utilisateur pour diminuer des éventuelles erreurs de typographie ou autre. Avec cette approche probabiliste, les résultats sont classés en fonction de leur probabilité d'apparition.

4.1.1 Estimation d'une probabilité conditionnelle

La méthode utilisée est basée sur des probabilités conditionnelles avec la probabilité d'apparition d'un événement sachant qu'un autre, lié, s'est déjà produit. Soit E un événement et R un autre événement, la probabilité d'apparition de E sachant que $R_1, \ldots, R_n$ se sont déjà produits est donnée par la formule suivante :

$$P\left(E_{|R_1,\ldots,R_n}\right) = \frac{P\left(R_1 \cap \ldots \cap R_n \cap E\right)}{P\left(R_1 \cap \ldots \cap R_n\right)} \qquad (2)$$

Comme précisé dans l'état de l'art, avec une base de données conséquente, il est possible d'utiliser cette méthode, les probabilités étant estimées à partir d'une base de données de référence.

Comme la base de données fournie est conséquente, il est donc possible d'approcher cette probabilité conditionnelle grâce à la formule (3). En effet, en comptant les occurrences de chaque événement E et de chaque couple d'évènements $(E, R_1, \ldots, R_n)$, on peut trouver une valeur s'approchant de la probabilité $\hat{P}\left(E_{|R}\right) \hat{P}(E|R)$ comme suit :

$$\hat{P}\left(E_{|R_1,\ldots,R_n}\right) = \frac{occ\left(E, |R_1,\ldots,R_n\right)}{occ\left(|R_1,\ldots,R_n\right)} \qquad (3)$$

$occ(E,R_1,\ldots,R_n)$: définit l'occurrence du triplet $(E,R_1,\ldots,R_n)$, c'est-à-dire le nombre de fois où le couple $(E, R_1, \ldots, R_n)$ est présent dans la base de référence.

Bien entendu, cette méthode n'est pas parfaite et comporte plusieurs limites :

- Premièrement, il s'agit d'un estimateur et, comme tout estimateur, il est soumis à un degré d'erreur. Ainsi, si la base de données d'apprentissage s'avère trop peu peuplée ou même possède des cas isolés avec un nombre d'occurrences trop faible, l'erreur s'en retrouvera augmentée. Dans un cas extrême, si cette base ne possède pas certains cas, il sera impossible de prédire ces cas.

- Deuxièmement, si la base de données possède beaucoup de bruit, l'erreur d'estimation sera également augmentée. On peut toutefois espérer qu'avec une base de données d'apprentissage assez importante, ce bruit se compensera assez pour faire ressortir les bons résultats.

4.1.2 Fusion de probabilité

Imaginons deux sources d'information (ou capteurs) (resp. C_1 et C_2) donnant chacun une information probabiliste de l'avènement d'un diagnostic D_1 (noté resp. $P_{C_1}(D_1)$ et $P_{C_2}(D_1)$), il est nécessaire de fusionner ces deux informations afin de prendre une décision.

[6] nous indique plusieurs méthodes permettant de réaliser cette fusion. La première méthode est de considérer le résultat du capteur qui donne la probabilité la plus importante. $P(D_1) = \max(P_{C_1}(D_1), P_{C_2}(D_1))$. La deuxième approche est de considérer que chaque capteur a la même importance d'où $P(D_1) = \frac{1}{2}[P_{C_1}(D_1) + P_{C_2}(D_1)]$. La dernière approche consiste à combiner la probabilité de chaque capteur en fonction de sa performance intrinsèque et $P(D_1) = \alpha_1 \cdot P_{C_1}(D_1) + \alpha_2 \cdot P_{C_2}(D_1)$.

4.2 Méthode proposée

Il n'est pas possible d'estimer directement la probabilité du diagnostic D_i sachant l'âge, le sexe, la durée de séjour, IGS2, l'unité médicale/unité fonctionnelle (UMUF), les actes et les diagnostics déjà codés. C'est pourquoi il a été décidé de distinguer quatre sources d'informations médicales :

(1) âge, sexe, durée de séjour, IGS2 ;
(2) unité médicale / unité fonctionnelle (UMUF) ;
(3) actes ;
(4) diagnostics déjà codés.

D'autre part des contraintes technologiques sur le projet pour lequel est effectué cette étude imposent que les estimations de probabilité soient stockées dans une base de type Microsoft Access et que le temps nécessaire pour retrouver les estimées (pour un cas donné) soit très court.

Pour chacune de ces sources d'informations, nous estimerons la probabilité d'avoir chaque diagnostic. Ensuite, ces probabilités seront fusionnées en tenant compte de la performance relative de chaque source.

- **Probabilité du diagnostic selon la source (1) avec classement personnalisable**

Suivant l'avis de notre consultant, 4 éléments à prendre en compte ont été déterminés pour pouvoir mieux prédire l'occurrence d'un diagnostic : l'âge du patient, le sexe du patient, l'Indice de Gravité Simplifié II (IGS2) du patient et la durée de séjour du patient. Afin de réduire l'explosion combinatoire, l'âge a été découpé en 22 intervalles, nous distinguons les cas où la valeur de l'IGS2 est supérieure à 5 ou non, de même pour la durée du séjour, nous considérons les séjours de plus de 24 heures et les autres.

Ceci a donc conduit à la formule ci-dessous pour le calcul de probabilité :

$$\check{P}(D_j | age, sexe, dur?edes?jour, IGS2) = \frac{occ(Diag_j, age, sexe, dur?edes?jour, IGS2)}{occ(age, sexe, dur?edes?jour, IGS2)} \quad (4)$$

L'estimation de cette probabilité est effectuée pour l'ensemble des diagnostics, des tranches d'âge, sexe, durée de séjour et IGS2. Cet estimateur a été calculé à partir de la base régionale (Bretagne) de RSA 2006.

- **Probabilité du diagnostic selon la source (2)**

Dans un second temps, c'est l'estimée de la probabilité d'un diagnostic en fonction de l'endroit où se trouve le patient dans l'hôpital, c-à-d l'unité médicale/ unité fonctionnelle (UMUF) qui a été calculée par :

$$\check{P}(D_j|UMUF) = \frac{occ(Diag_j, UMUF)}{occ(UMUF)} \qquad (5)$$

La composition des unités médicales/unités fonctionnelles est propre à chaque hôpital. C'est la raison pour laquelle nous avons fait nos calculs à partir des probabilités calculées au sein du CHU de Brest.

- **Probabilité du diagnostic selon la source (3)**

Dans un troisième temps, l'ensemble des actes codés est utilisé pour estimer les probabilités des diagnostics. Le nombre d'actes codé varie de 0 à N. Il n'est pas possible de vérifier l'indépendance entre les actes (d'un point de vue statistique, il est fort vraisemblable que certains actes sont dépendants entre eux et d'autres non).

C'est la probabilité d'un diagnostic selon un unique acte $\hat{P}(D_j|acte_i)$ qui a été estimée comme précédemment à partir de la base régionale de RSA. Afin de fusionner l'information issue de chaque acte, on supposera dans un premier temps que chaque acte a la même importance, dans un second temps on distinguera les actes classants et non classants.

$$\check{P}(D_j|acte_1,...acte_N) = \alpha_1 \cdot \frac{1}{L} \sum_{k=1, acte_k \text{ est classant}}^{L} \check{P}(D_j|acte_k) + \alpha_2 \cdot \frac{1}{M} \sum_{k=1, acte_k \text{ est non-classant}}^{M} \check{P}(D_j|acte_k) \qquad (6)$$

avec $L + M = N$ et α_1 et α_2 dépendant de la performance des actes classants par rapport aux actes non classants. Cette performance a été définie de manière expérimentale à partir d'une base de 1 000 RUM ne contenant que des actes non classants et une base de 1 000 RUM contenant uniquement des actes classants. Avec les actes classants, on obtient une meilleure performance. Dans notre expérimentation, la performance est définie en fonction du pourcentage de diagnostics à trouver (ou de famille diagnostique) qui sont listés dans les 10 premiers.

- **Probabilité du diagnostic selon la source (4)**

Dans un quatrième temps, l'ensemble des diagnostics déjà codés est utilisé pour estimer les probabilités des diagnostics. Le nombre de diagnostics codés varie de 0 à M. Tout comme pour les actes, il n'est pas possible de vérifier l'indépendance entre les diagnostics (d'un point de vue statistique, il est fort vraisemblable que certains diagnostics sont dépendants entre eux et d'autres non).

C'est la probabilité d'un diagnostic selon un unique diagnostic $\hat{P}(D_j|D_i)$ qui a été estimée comme précédemment. Afin de fusionner l'information issue de chaque diagnostic, on supposera que chaque diagnostic a la même importance.

$$\check{P}(D_j|D_1,...D_M) = \frac{1}{M} \sum_{k=1}^{M} \check{P}(D_j|D_k) \qquad (7)$$

- **Probabilité du diagnostic en fonction des 4 sources d'informations**

$$\hat{P}(D_j|D_1,...D_M) = \beta_1 \cdot \hat{P}(D_j|age, sexe, durée\ de\ séjour, IGS2) + \beta_2 \cdot \hat{P}(D_j|UMUF)$$
$$\beta_3 \cdot \hat{P}(D_j|acte_1,...acte_N) + \beta_4 \cdot \hat{P}(D_j|D_1,...D_M)$$

β_1 , β_2 , β_3 et β_4 dépendent de la performance individuelle de chacun des capteurs. Afin de les déterminer, nous avons évalué expérimentalement leur performance.

4.3 Algorithme

On rappelle que l'on effectue la prédiction sur la famille diagnostique (les 3 premiers caractères du code CIM10). Le codeur devant alors choisir le (ou les) diagnostic(s) parmi ceux qui appartiennent à chaque famille proposée.

L'algorithme permettant de proposer des codes diagnostics s'effectue en plusieurs étapes :

1. Affichage des 10 familles diagnostiques ayant les plus hautes probabilités (ces probabilités sont déterminées comme indiqué précédemment).
2. Choix des diagnostics parmi les familles diagnostiques présentées par le codeur.
3. Affichage des 10 familles diagnostiques suivantes (en tenant compte des diagnostics et familles diagnostiques choisis et non choisis) (si nécessaire).
4. Choix des diagnostics présents.
5. Affichage des 10 familles diagnostiques suivantes (en tenant compte des diagnostics choisis et non choisis).
6. Choix des diagnostics présents.

5 Résultats - Discussion

Afin de valider la faisabilité de l'application de cette méthode, nous avons appliqué cette méthode sur plusieurs bases de test comportant 1 000 RUM. Ces bases sont créées par tirage aléatoire de cas réels déjà codés par un médecin codeur. Lors de cette expérience, nous supposons que le médecin codeur aurait choisi dans l'ensemble des codes proposés l'ensemble des codes qui lui sont nécessaires. Ces bases n'ont pas fait l'objet de contrôle interne sur la qualité de codage.

L'objectif de cette expérience est de montrer qu'il est possible d'utiliser les informations déjà codées dans le RUM et que l'on peut proposer au médecin codeur une liste de codes/libellés contextualisée qui contienne une partie ou la totalité des codes dont il a besoin.

Chaque base a des propriétés propres. Nous avons créé une base ne comportant que des cas (RUM) avec des actes non classants et une base possédant des cas ayant à la fois des actes classants et non classants. C'est cette dernière qui illustrera les résultats obtenus. Chaque RUM comporte une liste d'actes, l'âge du patient, la durée du séjour, le sexe, l'IGS2 et une liste de diagnostics (que l'on recherche).

La figure 2 présente les résultats obtenus sous la forme d'histogramme. Chaque barre verticale est décomposée en 3 parties qui représente successivement en partant du haut, le pourcentage de RUM dont on n'a pas trouvé un seul diagnostic, le pourcentage de RUM

dont on n'a trouvé qu'une partie des diagnostics et enfin le pourcentage de RUM dont tous les diagnostics ont été détectés. Cette figure présente les taux de classification après la première et après la troisième itération.

Les résultats obtenus par le capteur - source (1) - (sexe, âge, durée de séjour, IGS2) correspondent aux colonnes intitulées « quadruplet » , pour le capteur - source (3) - utilisant les actes, les colonnes sont intitulés « actes », pour la source (2) - unité médicale/unité fonctionnelle -, il s'agit de « umuf » et enfin dans le cas de la fusion, la colonne est intitulée « tout ».

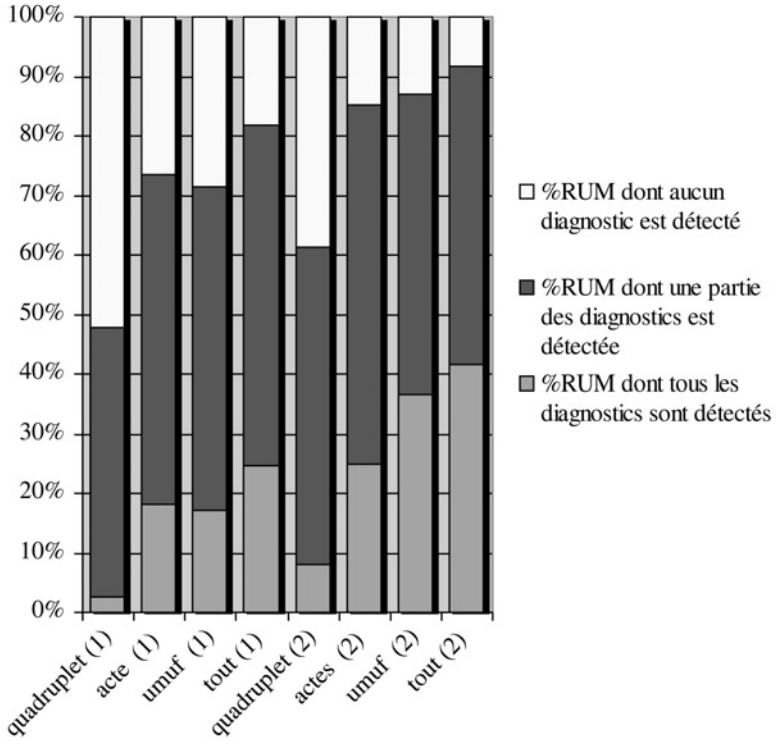


Figure 2 - Résultats obtenus après les deux 1res étapes de l'algorithme (1) et à l'issue de l'algorithme (2).

Ces résultats nous permettent de classer la qualité de classification de chaque source d'information et d'en déduire une partie des coefficients β_i (proportionnels au taux de RUM dont les diagnostics ont été totalement trouvés). Une modification légère de ces coefficients n'a qu'une très faible influence sur le résultat.

Le critère retenu est d'abord le pourcentage de RUM dont l'ensemble des diagnostics a été trouvé puis le pourcentage de RUM dont seule une partie des diagnostics est détectée.

- **Résultat obtenu à partir de** $\hat{P}\left(D_j \middle| age, sexe, durée\ de\ séjour, IGS2\right)$, $\hat{P}\left(D_j \middle| acte_1, ...acte_N\right)$ et $\hat{P}\left(D_j \middle| UMUF\right)$

La première étape a consisté à évaluer successivement les 3 sources d'informations (sexe, âge, durée de séjour, IGS2), les actes et l'unité médicale/unité fonctionnelle où se trouve le patient).

Les résultats obtenus nous indiquent que l'information la plus pertinente est l'unité médicale/unité fonctionnelle où se trouve le patient, puis celle issue des actes qui lui sont prodigués et enfin celle issue des données administratives.

On remarque également qu'en augmentant le nombre d'itération, un plus grand nombre de résultat est retrouvé (gain de 10 %).

- **Résultat obtenu en ajoutant** $\hat{P}\left(D_j \middle| D_1,...,D_M\right)$

Afin de valider l'apport de la connaissance des diagnostics, celle-ci a été insérée lors de l'estimation des probabilités au cours de la deuxième et de la troisième itération. Les diagnostics sont fournis lors de la première itération et de la deuxième. On a pu montrer que cette connaissance permettait d'améliorer le pourcentage de RUM où l'ensemble des diagnostics est trouvé. Le gain dans le nombre de RUM dont les diagnostics sont totalement détectés est de 3 à 5 %. On a remarqué que plus ce taux était élevé, plus faible était le gain.

- **Résultat obtenu par combinaison de toutes les sources d'information**

Le principal problème a été de définir la valeur des coefficients β_i. Concernant les trois premiers coefficients, leurs valeurs sont proportionnelles aux pourcentages de RUM dont l'ensemble des diagnostics a été trouvé. Le quatrième coefficient a été réglé de manière expérimentale. On a choisi la valeur qui optimise le critère.

Lors de la première itération, le résultat « tout » n'inclut que les trois premières sources d'information tandis que pour la troisième itération, il utilise l'ensemble des sources d'information.

Les résultats montrent qu'il n'est pas possible de retrouver la totalité des codes diagnostics. Seuls les codes les plus fréquents, pour une situation donnée, seront proposés.

6 Conclusion

Cette étude nous a permis de développer une méthode d'aide au codage relativement satisfaisante et simple d'approche. Cette méthode s'appuie sur le développement d'une méthode de prédiction probabiliste fusionnant le résultat obtenu par plusieurs sources d'information dont chacune donne une information sur le diagnostic à prédire. Sans donner la totalité des codes désirés par le médecin codeur, elle a permis de lui proposer des résultats crédibles. S'il existe encore de nombreux points améliorables et à améliorer, nous pouvons malgré tout dire que cette méthode se révèle prometteuse.

Cependant, ces résultats, aussi satisfaisants soient-ils, ne doivent pas nous faire oublier qu'ils sont loin d'être parfaits. Ainsi, utiliser une base de données d'apprentissage comportant déjà des erreurs ne pouvait nous conduire qu'à un résultat imparfait et, quelque soit la méthode d'apprentissage choisie, si la base de données se révèle trop bruitée, il est impossible d'obtenir des résultats convenables, et ce sont bien là les limites de cette approche.

D'autre part, cette méthode ne propose que les termes qui ont la plus forte probabilité d'être codés. Les diagnostics rares seront difficilement détectés dans les premiers termes.

La prochaine étape à ce travail visera à valider cette approche en condition réelle. Cette étude vient de démarrer par l'utilisation d'une interface utilisateur, en connexion avec le

logiciel de codage, qui permet de proposer à un codeur des familles diagnostiques et des codes diagnostiques. Cette interface permet de comparer l'utilisation de ce travail en condition réelle et de la comparer à une méthode d'aide au codage qui utilise une recherche par mots clefs. Cette étude nous permettra de connaître l'apport (ou non) de ce travail in situ, le gain de temps dans le codage, le gain en termes de diagnostics codés et si cela peut entraîner un apport dans le codage.

À terme, ce travail intègrera la plate-forme Med1dex et les résultats de l'approche probabiliste seront fusionnés à ceux d'autres modules pour optimiser l'aide au codage

Remerciements

Ce projet est un projet TECSAN soutenu par l'ANR sous le nom Med1dex associant le CHU de Brest, TELECOM Bretagne, le LaTIM Unité INSERM U650 et la société PRISMEDICA.

Références

[1] Noury JF. *La gestion médicalisée des établissements de santé, Le PMSI et l'information médicale*. CNRS Editions, 2000.

[2] Agence Technique de l'Information sur l'Hospitalisation (ATIH), Édition consultable et imprimable de la CCAM Version 10, http://www.atih.sante.fr/index.php?id=0003100027FF, consulté le 24/01/2008.

[3] Maravic M, Le Bihan C, Landais P. La classification commune des actes médicaux (CCAM) : de la description à la tarification, *Revue du Rhumatisme* 2003 ; 70 9: 785-9.

[4] *Classification statistique internationale des maladies et des problèmes de santé connexes: CIM-10*. volumes 1, 2 et 3. In: Organisation mondiale de la santé (OMS) (ed). 1998.

[5] Abraham B et Ledolter J. *Statistical methods for forecasting*. In: Wiley-Interscience (ed), 1983.

[6] Grandin JF. *Fusion de données : Théorie et méthode*. In : Technique de l'ingénieur (ed) 2006.

[7] Zighed DA, Rakotomolala R. *Extraction de connaissances à partir de données (ECD)*. In. Techniques de l'ingénieur, traité informatique (ed) 2002; 14-21.

Adresse de correspondance

L. Lecornu

Institut TELECOM ; TELECOM Bretagne, Dpt ITI
Technopole de Brest Iroise
CS 83818
29238 Brest Cedex 3, France

E-mail : Laurent.Lecornu@telecom-bretagne.eu

Chapitre 7

Web et eSanté

Informatique et Santé, 2009 (17) : 277-286
© Springer – Verlag France

CEMARA : une plate-forme Web pour les maladies rares

Claude Messiaen[1], Loïc Le Mignot[1], Ana Rath[3], Jean-Baptiste Richard[1], Éric Dufour[2], Mohamed Ben Said[1], Jean-Philippe Jais[1], Alain Verloes[4], Martine Le Merrer[5], Christine Bodemer[6], Geneviève Baujat[5], Marion Gerard-Blanluet[4], Eva Bourdon-Lanoy[6], Rémi Salomon[7], Ségolène Aymé[3], et Paul Landais[1]

[1] *Université Paris Descartes, UPRES EA 4067 et Service de Biostatistique et d'Informatique Médicale, APHP – Hôpital Necker-Enfants Malades ;* [2]*SELIC;* [3]*Orphanet et INSERM SC 11 ;* [4]*Département de Génétique, APHP- Hôpital Robert Debré ;* [5]*Département de Génétique,* [6] *Service de Dermatologie et* [7]*Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Paris.*

Abstract

A rare disease is a condition the prevalence of which is lower than 5 cases per 10,000 inhabitants. At least 7000 rare disease have been discovered. In France, rare diseases affect nearly 30,000 patients and often only few cases for a given disease. Three to 4% of children and 6% of the population in Europe present with a rare disease. Most of these diseases do not have any curative treatment yet. It is thus a true public health stake. In France, a National Rare Diseases Plan was initiated in 2004. Thirty three out of 132 labelled Reference Centres (RC) share a common Information System called CEMARA. It is dedicated to collect continuous and complete records of patients presenting with a rare disease, and their follow-up. The main objective of CEMARA is to analyze the epidemiological patterns of rare diseases. It is also to contribute to the missions of the RC regarding the registration and description of their activities, coordination of the network of their correspondents, organization of the follow-up of rare diseases. A description of CEMARA is provided as well as its cooperation with Orphanet and GenAtlas, and a presentation of the 30,119 current records collected by more than 400 health care professionals belonging to more than 100 specialized sites.

Keywords

Dynamic Web interface; Information System; Thesaurus; Rare Diseases.

1 Introduction

Les maladies rares englobent un groupe d'affections caractérisées par une faible prévalence. En Europe, une maladie est « rare » quand elle atteint moins d'une personne sur 2 000. En France, les maladies rares pourraient affecter 30 000 patients. Pour la Communauté européenne, cela signifie jusqu'à 230 000 patients. De 5 000 à 7 000 maladies rares sont actuellement identifiées. Trois à 4 % des enfants et 6 % de la population en Europe est affectée. Il s'agit d'un véritable problème de santé publique, puisque la plupart des maladies n'ont pas de traitement curatif. En France, le ministère de la Santé a lancé un Plan national pour les maladies rares en 2004. Cent trente-deux centres

de référence (CR) [1] ont été labellisés pour un groupe de maladies comme les « anomalies du développement » ou les « maladies rénales rares », ou pour une maladie telle que « le syndrome de Prader-Willi » ou la « maladie de Willebrand ». Les CR composés d'équipes pluridisciplinaires ont plusieurs missions :

- Faciliter le diagnostic et définir une stratégie de prise en charge thérapeutique, psychologique et d'accompagnement social ;
- Définir et diffuser des protocoles de prise en charge, en lien avec la Haute Autorité de Santé et l'Union nationale des caisses d'assurance maladie ;
- Coordonner les travaux de recherche et participer à la surveillance épidémiologique, en lien avec l'Institut national de veille sanitaire ;
- Participer à des actions de formation et d'information pour les professionnels de santé, les malades et leurs familles, en lien avec l'Institut national de prévention et d'éducation pour la santé (INPES) ;
- Animer et coordonner les réseaux de correspondants sanitaires et médico-sociaux ;

Un système d'information nommé CEMARA (Centres Maladies Rares) a été mis en place afin de recueillir de manière continue et exhaustive les enregistrements des patients présentant une maladie rare et, le cas échéant, leur suivi dans les CR adhérents. L'objectif principal de CEMARA est l'analyse de l'épidémiologie des maladies rares. Cette plate-forme contribue également aux missions des CR en ce qui concerne l'enregistrement et la description de leurs activités, la coordination du réseau de leurs correspondants, ou du suivi des maladies rares.

2 Matériel et méthodes

2.1 Réseau des professionnels CEMARA

Actuellement, 33 centres sont membres du réseau CEMARA en France. Chaque centre est généralement composé de plusieurs unités cliniques distribuées sur plusieurs sites couvrant l'ensemble du territoire français[1]. Un centre est dirigé par un coordonnateur et un médecin correspondant est désigné par ses pairs pour chaque site. Près de 400 membres sont inscrits pour plus de 100 sites. Une charte de confidentialité du réseau assure la préservation du respect des données nominatives propres à chaque site. Tous les utilisateurs ont donné leur accord et ont signé la charte qui a obtenu l'accord de la CNIL et du Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS). Par une décision récente, le ministère de la Santé a décidé d'étendre le maillage territorial des centres maladies rares par la labellisation de Centres de Compétence (CC). Chaque centre de compétence est affilié à un centre de référence. Pour prendre en compte cette nouvelle organisation, CEMARA ouvrira progressivement 400 nouveaux sites.

2.2 Thésaurus : conception et constitution

Pour les intitulés diagnostiques des dossiers patients, une classification a été mise en place avec les professionnels de santé correspondant à leurs besoins. Pour les maladies rares, la Classification Internationale des Maladies 10e version (CIM10) ne fournit pas le support informationnel adapté car sa granularité n'est pas appropriée pour une telle cible, et de nombreuses maladies génétiques rares n'y sont pas actuellement listées. De plus, la structure de CIM10 n'est pas adaptée à un domaine où la nosologie évolue très rapidement.

[1] https://cemara.org/Présentation/show.jsp?sm=membres (19 jan 2009).

Ce sont en effet près de 300 nouvelles entités qui sont découvertes chaque année. Or il a fallu 17 années pour passer de la CIM9 (1976) à la CIM10 (1993). Nous nous sommes donc associés pour ce travail ontologique à Orphanet [2] qui a conçu un thésaurus spécifique dédié aux maladies rares. Orphanet a un savoir-faire dans le domaine et est membre de la *Rare Disease Task Force* européenne[2]. Il fournit un soutien ontologique à des experts pour chaque domaine de CEMARA. La spécificité de cette collaboration prévoit une ontologie pour chaque thésaurus de CEMARA avec un caractère lexical unique pour la labellisation de chaque maladie, y compris la gestion des synonymes et des éponymes associés à certaines maladies ou de syndromes complexes (par exemple, syndrome de De Toni-Debré-Fanconi et tubulopathie proximale). En outre, les nosologies et nosographies associées sont fournies aux professionnels de différentes spécialités afin de leur permettre de partager des termes communs. Orphanet peut aussi bénéficier d'un retour d'information via CEMARA par l'étude, à rythme régulier, des associations de diagnostics avec des mots clés et/ou notes émises par les médecins utilisateurs, qui peuvent conduire à des évolutions du thésaurus.

En ce qui concerne les mutations génétiques, nous utilisons GENATLAS[3], qui contient des informations pertinentes à l'égard de la cartographie génique et des maladies génétiques. Un lien spécifique a été élaboré vers cette base de données. Les codes OMIM [2] (*Online Mendelian Inheritance in Man*) de la base de données génotypiques et phénotypiques seront également mis à disposition. Un lien sera également assuré avec la base « Ensembl »[4] [3]. Il s'agit d'un projet collaboratif débuté en 2000, conçu comme un navigateur pour le génome humain.

Un aspect important concerne la granularité du thésaurus. En effet, un choix était à faire entre une granularité fine, source de variabilité de codage (difficultés de codifier des règles de codage à l'échelon du symptôme par exemple), et une granularité plus importante (par exemple, les ensembles syndromiques, retard mental avec ou sans dysmorphie), source d'hétérogénéité de sens. Nous avons choisi de nous placer au niveau des syndromes et des maladies et non au niveau des signes en raison de la présence en nombre notable de syndromes, en particulier pour les dysmorphies. Ceci permet de recueillir une information pertinente pour le niveau de description épidémiologique.

Autre aspect important, l'arborisation du thésaurus. Le thésaurus a une structure à profondeur variable d'une spécialité à l'autre ou dans une même spécialité. Par exemple, le syndrome du triple A (profondeur 6), le syndrome de Aase-Smith (profondeur 3) et maladie de Willebrand, forme non acquise (profondeur 2). En dépit de cette variabilité de profondeur, le niveau de pertinence sémantique a été maintenu congruent avec la granularité définie précédemment.

2.3 Système d'information (SI) : présentation et mise en œuvre

Le SI de CEMARA rassemble dans une représentation standard un enregistrement minimum patient élaboré par les professionnels des spécialités médicales des enfants et des adultes. CEMARA vise à remplir plusieurs exigences : l'évolutivité, la modularité, la fiabilité, l'accessibilité et l'efficacité sur la base de logiciels non propriétaires.

L'architecture de CEMARA (figure 1) est fondée sur une architecture n-tier. Via un navigateur Web, le client se connecte à un « middle-tier », qui est relié à plusieurs bases de données : la base de données de production, le dictionnaire géographique, et la base de

[2] Rare Disease Task Force : http://www.rdtf.org/ (19 nov 2008)

[3] GENATLAS : http://www.dsi.univ-paris5.fr/genatlas/ (19 nov 2008).

[4] Ensembl : http://www.ensembl.org/ (19 jan 2009).

données thesaurus. Un entrepôt de données et un système d'information géographique permettant des requêtes et leurs représentations sont en cours de réalisation. Cette logique de conception est à rapprocher d'une application déjà développée et utilisée pour l'insuffisance rénale terminale depuis 2002 [4].

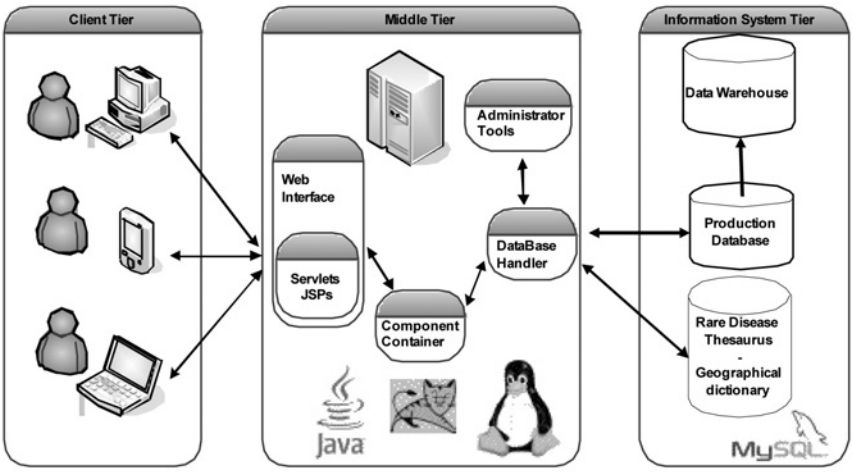


Figure 1 - CEMARA : une architecture n-tier.

Le « Middle-tier » appuie les services aux clients par le biais de conteneurs Web et de la logique métier. Cette dernière est interfacée dans le « middleware » avec une bibliothèque d'interrogation de la base de données gestionnaire-dépendante qui supporte les opérations vers la base de données. Du côté client, CEMARA s'appuie sur les installations locales de réseaux Internet et sur des configurations ordinateurs très diverses dans les services médicaux. Un accès par l'intermédiaire d'un assistant numérique personnel est également à l'étude.

La maintenance et les évolutions sont centralisées, ce qui réduit les coûts et les délais. La structure de la base de données de production correspond à la structure du réseau CEMARA et respecte la confidentialité et la sécurité des données [5]. Une fonction de dédoublonnage a été spécifiquement développée afin d'éviter tout doublon d'identité [6].

Le déploiement de CEMARA a commencé le 23 mai 2007, initié par 15 centres. Les coordonnateurs et les utilisateurs ont été formés, par l'intermédiaire de sessions centralisées et sur site. Un assistant de recherche clinique a organisé la formation et est en charge de l'assistance technique.

CEMARA s'appuie également sur des collaborations avec :

- des institutions telles que le ministère de la Santé, la mission maladie rare, l'Institut national de Veille Sanitaire, la Haute Autorité de Santé et la Caisse nationale d'assurance maladie, concernant la conception des protocoles nationaux de diagnostic, de traitement et de suivi, ainsi que l'analyse des coûts ;
- les hôpitaux universitaires pour la coordination et la connexion des professionnels par l'intermédiaire du réseau CEMARA ;
- les associations de familles de patients pour améliorer les soins aux patients.

2.4 Tronc commun

Tous les centres partagent un tronc commun d'informations qui comprend des données d'identification (pour le cas-index et les éventuels membres de sa famille), le diagnostic, le contexte et l'activité médicale. L'identification des patients est fondée sur : nom, prénom, sexe, date de naissance, date de décès, lieu de résidence. Les modalités des activités médicales et des modes de recrutement sont enregistrées. Des informations facultatives sont recueillies sur les antécédents médicaux du patient, par exemple des données prénatales et/ou néonatales.

Un ou plusieurs diagnostics sont étiquetés à l'aide du thésaurus correspondant, ainsi que le statut de ce diagnostic : probable, confirmé, en cours, infirmé, non déterminé, non classable. Une description chromosomique de l'anomalie par bras est également disponible, ainsi que des mots clés supplémentaires pour les signes atypiques des symptômes, ou pour les patients présentant un diagnostic encore inconnu ou non étiquetable. Des mots-clés ont été proposés par le comité de pilotage. Ils permettent d'annoter des syndromes ou des maladies dont l'expression phénotypique diffère de l'expression courante. Cette annotation a pour objectif d'identifier les patients qui présentent des phénotypes atypiques. La sélection d'un diagnostic utilise des requêtes de type Asynchronous JavaScript et XML (AJAX).

2.5 Pétales

Des informations complémentaires pour des maladies spécifiques sont décrites et recueillies dans ce que nous appelons les « pétales », rattachés à la base de données, en mettant l'accent sur la collecte de données spécifiques. Actuellement, six «pétales» ont été déployés et plusieurs autres sont en cours de développement. Ces pétales sont développés de manière modulaire (figure 2) à l'image de « plug-in ». De cette façon, nous pouvons les attacher à (voire les détacher de) la base de données du tronc commun avec un minimum d'investissement et de développement. Ils sont fondés sur l'enregistrement d'informations spécifiques et de données biologiques, mais aussi sur certains aspects sociaux comme la scolarité ou le handicap, ou psychologiques. Des pétales peuvent être développés spécifiquement dans le cadre de suivi d'études de type Protocole Hospitalier de Recherche Clinique (PHRC).

Ces pétales donnent une capacité d'enrichissement du tronc commun de CEMARA permettant une personnalisation des données par centre et/ou site et ceci à 3 niveaux d'intégration : sur le diagnostic, sur l'activité, ou plus globalement sur le dossier.

Deux formes d'implémentation pour l'Interface Homme-Machine (IHM) sont alors possibles soit par l'ajout de champs dans le bordereau CEMARA (cas d'informations complémentaires sur diagnostic), soit par l'ajout d'un bordereau complémentaire.

2.6 Interface utilisateur

L'interface permet à l'utilisateur d'accéder à plusieurs fonctions : les fonctions dossiers (création et consultation), saisie d'activité, outils (recherches, exports), documents, messagerie sécurisée intégrée et un annuaire (structures et membres de CEMARA). Des paramètres personnalisables permettent également de modifier l'affichage.

L'ergonomie générale de l'application et plus précisément celle des bordereaux de saisie a été pensée et développée pour limiter le nombre de clics. Ainsi toutes les actions et les options sont visibles en permanence et suivent le défilement vertical.

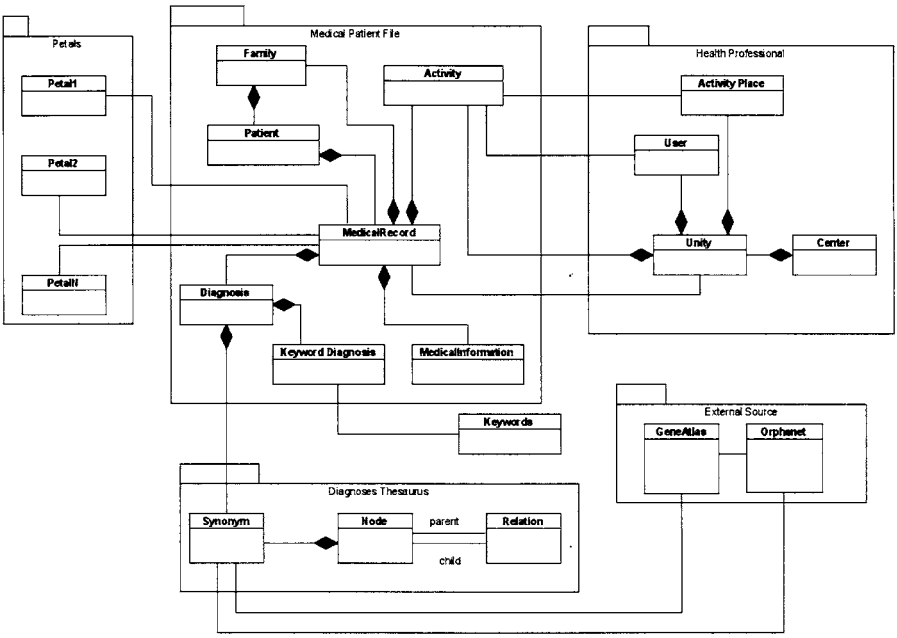


Figure 2 - Schéma UML de CEMARA.

D'autre part, une aide à la saisie par recherche AJAX est appliquée en cours de frappe. Enfin, l'historisation est représentée par une visualisation via des onglets intégrés aux bordereaux (figure 3).

2.7 Interopérabilité

CEMARA a été conçu afin de communiquer avec d'autres sources d'information : l'usage de XML, en tant que format d'échange, permet une plus grande flexibilité et une meilleure capacité d'échanger des données avec d'autres systèmes d'information français tels que le système de l'Assurance Maladie ou des systèmes d'information des hôpitaux. Il permet également l'importation de bases de données préexistantes (figure 4).

Ainsi deux schémas XML CEMARA distincts ont été conçus pour les flux entrant et sortant permettant un contrôle d'intégrité des données et un contrôle sémantique.

Sur le flux entrant, le schéma XML donne :

- la capacité de gérer la diversité des données sources par gestion de briques type « Intégration d'Application d'Entreprise » (EAI) avec pour objectif de transformer les données émettrices en flux XML CEMARA.
- la capacité de gestion de jeton de synchronisation avec contrôle à la source.

Il permet également une amélioration de la sécurité par la possibilité d'envoi différé des données d'identification et des données médicales.

Figure 3 – Interface CEMARA : exemple de la partie « activité » du bordereau.

3 Résultats

Plus de 400 professionnels de santé appartenant à plus de 100 sites sur les 33 CR contribuent à la collecte des données. Actuellement, l'ensemble des données rassemble 30 119 dossiers patients et comprend 15.456 patients de sexe masculin (963 fœtus) et 13 784 de sexe féminin (906 fœtus). Pour 877 fœtus le sexe n'a pas été déterminable. L'âge médian est de 8,36 ans.

Les patients atteints de maladies rares ont été principalement adressés aux CR par les pédiatres (42 %) et les médecins spécialistes (26 %), tandis que pour les fœtus ce sont principalement par des gynécologues (65 %) et par des centres de diagnostic prénatal (24 %). Ces patients ont été majoritairement recrutés par le biais des consultations (77 %) ou d'avis en salle (7 %). Comme attendu, le nombre de patients par diagnostic est limité, même pour les diagnostics les plus fréquents. En outre, environ 1/3 des dossiers sont classés comme « diagnostic inconnu », ce qui signifie que la conclusion diagnostique n'a pas été atteinte, c'est un phénomène courant chez les patients atteints de malformations multiples, dont le syndrome ne correspond pas aux entités actuellement étiquetées. Pour les autres pathologies, le tableau ci-après présente les 10 diagnostics les plus fréquents.

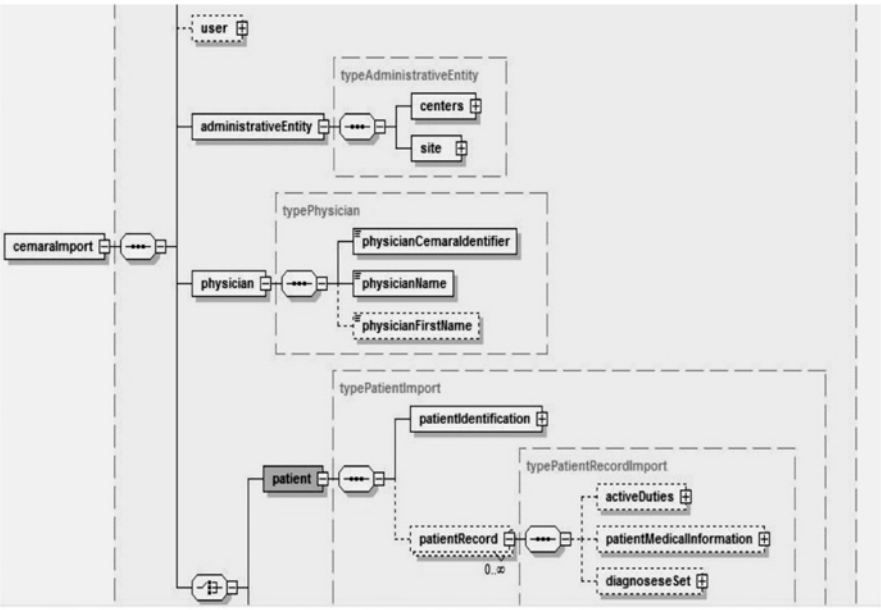


Figure 4 – CEMARA – partie du schéma XML d'import de deonnées

4 Discussion

Le programme de la Communauté Européenne pour les maladies rares, initié en 2003, visait à améliorer l'information et la connaissance pour le développement de la santé publique [7]. En France, le plan maladies rares a joué un rôle clé dans la promotion de la prise en charge de ces maladies [8]. Les CR doivent normaliser leurs informations et structurer leur manière de les colliger. Ceci a déclenché le développement de CEMARA. L'expérience partagée avec Orphanet et GenAtlas a permis d'étendre l'expertise engrangée avec le programme national français REIN (Réseau Épidémiologie et Information en Néphrologie) [4] que nous avons développé pour l'insuffisance rénale terminale. Elle renforce la pérennité de notre application. En outre, l'application est fondée sur les logiciels libres, qui permettent des coûts de développement plus limités. La conception d'une structure évolutive transmise par sa modularité assure son extensibilité.

D'autre part, nous avions constaté une limitation concernant le thésaurus. En effet, le thésaurus ne comporte pas de définitions des entités, ce qui aurait pu nuire à son homogénéité sémantique. La solution mise en place pour y pallier a été de mettre à disposition des professionnels un lien direct vers les descriptions Orphanet[5].

La question de l'adhésion des professionnels a trouvé sa réponse dans le modèle organisationnel mis en place et hérité de REIN. En effet, la constitution, du réseau « humain » a été un préalable à la mise en place de la plate-forme informatique. Ce réseau humain respecte la structure des centres de référence (centre/sites), il est fondé sur l'étanchéité des données nominatives, matérialisé par la structure alvéolaire de type « ruche » du modèle. Chaque site gère ses données nominatives de façon étanche, dans le

[5] Orphanet : http://www.orpha.net (19 jan 2009).

respect des directives de la loi Informatique et Liberté, et matérialisé par une charte, signée par chaque utilisateur. Ces dispositions ont permis un consensus fort qui est l'un des piliers de la réussite de CEMARA. Sur une période de 9 mois, près à 12 000 cas de maladies rares ont été enregistrées et plus de 30 000 en 18 mois. C'est un reflet de l'adhésion et de l'implication des professionnels et de l'utilité de cette application.

Tableau 1 - Les dix pathologies les plus fréquentes dans CEMARA.

Pathologie	Effectifs
Neurofibromatose de type 1	574
Ostéogenèse imparfaite	420
Autisme	400
Trisomie 21	387
Mucoviscidose	375
Syndrome de Di George	363
Syndrome néphrotique cortico-sensible	298
Oligoarthrite juvénile	293
Syndrome de Noonan	284
Syndrome de Turner	281

Parallèlement, d'autres registres en ligne existent ou sont émergents. Le plus important d'entre eux est celui de l'ESID (*European Society for Immunodeficiencies*) [9, 10]. Mais il se différencie sur plusieurs points. Tout d'abord, il s'agit d'un registre européen dans un champ médical unique, les déficits immunitaires. Par ailleurs, il diffère aussi sur un point essentiel : l'objectif. En effet, le registre de l'ESID se veut une aide à la décision diagnostique par une procédure de validation fondée sur la description des symptômes. Alors que CEMARA se pose comme un outil épidémiologique et d'aide à la décision en Santé publique.

Les utilisateurs de bases de données affirment que souvent ils ne disposent pas d'un véritable retour sur les données qu'ils renseignent. Par conséquent, nous fournissons tous les trois mois des rapports sur l'activité et sur les diagnostics, et tous les ans un document reprenant les éléments du référentiel d'évaluation de la Haute Autorité de Santé (HAS) et de la Direction de l'Hospitalisation et de l'Organisation des soins (DHOS). Chaque participant peut aussi utiliser un outil de recherche multicritères pour étudier ses cas.

Enfin, les informations seront stockées dans un entrepôt de données, le CEMARACube. Cet entrepôt sera accessible aux utilisateurs par un système d'information géographique, le SIGMARA via le web.

5 Conclusion

CEMARA permet d'évaluer la répartition géographique des cas existants pour améliorer l'offre de soins. Il permet de mieux décrire les nouveaux cas de maladies rares, d'en organiser le suivi, et d'identifier les besoins médicaux et psycho-sociaux. L'usage d'une interface Web a pour but d'aider l'enregistrement médical, la prise en charge et la connaissance des maladies rares. Il vise à contribuer à la santé publique, à l'adaptation de l'offre de soins avec une meilleure connaissance de ces maladies, offrant ainsi un appui décisionnel de santé publique. L'analyse épidémiologique sera interfacée avec le logiciel

de statistiques R [11] sur des données agrégées et préalablement anonymisées en utilisant notre expérience découlant de SIGNE, un Web-système d'information géographique [12].

Remerciements

Les coordonnateurs et les professionnels des 33 centres CEMARA sont chaleureusement remerciés pour leur participation au programme CEMARA.

Références

[1] Alliance Maladies Rares, Liste des centres labellisés, http://www.alliance-maladies-rares.org/IMG/pdf/Liste_Centres_de_Reference.pdf (19 nov 2008).

[2] CEMARA et Orphanet : une collaboration au service des centres de référence, http://www.orpha.net/actor/Orphanews/2007/070321.html (19 nov 2008).

[3] Hubbard T, Barker D, Birney E *et al.* The Ensembl genome database project. *Nucleic Acids Res* 2002; 30(1): 38-41.

[4] Landais P, Simonet A, Guillon D, Jacquelinet C, Ben Saïd M, Mugnier C, Simonet M. SIMS REIN: a multi-source information system for end-stage renal disease. *C R Biol* 2002 ; 325 : 515-28.

[5] Rindfleish TC. Privacy, Information Technology and Health Care. *Communications of the ACM* 1997; 40: 93-100.

[6] Le Mignot L, Mugnier C, Ben Saïd M, Jais JP, Richard JB, Le Bihan-Benjamin C, Taupin P, Landais P. Avoiding doubles in distributed nominative medical databases: optimization of the Needleman and Wunsch algorithm. *Stud Health Technol Inform* 2005; 116 : 83-8.

[7] La Commission Européenne, Public Health - Programme 2003-2008 - Funded projects - 2004 http://ec.europa.eu/health/ph_projects/2004/action1/action1_2004_10_en.htm (19 nov 2008).

[8] Ministère de la Santé, de la Jeunesse, des Sports et de la Vie Associative, Les Maladies Rares, http://www.sante.gouv.fr/htm/actu/maladie_rare/plan.pdf (19 nov 2008).

[9] Eades-Perner AM, Gathmann B, Knerr V, Guzman D, Veit D, Kindle G, Grimbacher B; ESID Registry Working Party. The European internet-based patient and research database for primary immunodeficiencies: results 2004-06. *Clin Exp Immunol* 2007; 147(2): 306-12.

[10] Guzman D, Veit D, Knerr V, Kindle G, Gathmann B, Eades-Perner AM, Grimbacher B. The ESID Online Database network. *Bioinformatics* 2007; 23(5): 654-5.

[11] The R Project for Statistical Computing: http://www.r-project.org/ (19 nov 2008).

[12] Richard JB, Toubiana L, Le Mignot L, Ben Said M, Mugnier C, Le Bihan-Benjamin C, Jaïs JP, Landais P. A Web-based GIS for health care decision-support. *AMIA Annu Symp Proc* 2005 ; 365-9.

Adresse de correspondance

Pr Paul Landais

Service de Biostatistique et d'Informatique Médicale – Hôpital Necker-Enfants Malades

149, rue de Sèvres, 75743 Paris, France

E-mail : landais@necker.fr

Informatique et Santé, 2009 (17) : 287-298
© Springer – Verlag France

Les informations médicales en ligne dignes de confiance : étude des spécificités en Afrique francophone

[1,2]Cheick Oumar Bagayoko, [3] Celia Boyer, [1] Antoine Geissbühler

[1]Université de Genève et Hôpitaux Universitaires Genève, Suisse
[2]Laboratoire d'Enseignement et de Recherche sur le Traitement de l'Information Médicale, Faculté de Médecine Université de la Méditerranée Marseille, France
[2]Fondation Health On the Net, Genève, Suisse

Abstract
Access to quality medical information is an important part of a health system because it may help to improve people's health. With the development of information technology and communication, new tools are available to facilitate the production, publication and access to such information. However, these tools pose new challenges. It becomes increasingly difficult to assess the credibility of medical information on the Web. It is not clear if the techniques used to produce medical information online are reliable. In addition, it is unclear whether the confidence that we can give this information is based on reliable universal criteria. It seems important to take into account the diversity of social, economic and cultural contexts in which this information is used. This project offers recommendations for production and access to online medical information for health professionals. It studies the situation in French speaking Africa and particularly in Mali.

Keywords
Medical information; Health on the net; Specifities in Africa.

1 Introduction

Avec le développement de la société de l'information se pose le problème de la mondialisation des savoirs et des tensions provoquées par la méconnaissance des facteurs identitaires et culturels. Comme le dit Dominique Wolton [1], « Ce n'est pas parce que l'Autre est aujourd'hui plus accessible qu'il est plus compréhensible, c'est même précisément l'inverse. Plus nos différences sont visibles, plus elles créent des tensions. »
Afin que tous puissent bénéficier des outils d'Internet – notamment dans le domaine de la santé où ils permettent d'améliorer l'accès à l'information, le pilotage des systèmes de santé et le « *patient empowerment* » – il convient de développer les accès à Internet, des compétences et des règles d'usage. Chacun pourra ainsi participer au réseau mondial et y apporter ses connaissances et ses expériences [2].
C'est dans cette optique que la fondation Health On the Net (HON) s'est développée depuis dix ans [3]. Surtout active dans les pays développés, avec plus de 5 000 sites médicaux accrédités dans 72 pays, la fondation HON étend ses activités aux pays en

développement, en y établissant des représentations régionales, afin de mieux répondre aux besoins locaux d'accréditation et de valorisation des informations médicales en ligne.

L'Organisation mondiale de la santé (OMS) renforce son activité dans ce domaine, suite à la résolution sur la cybersanté votée par les états membres en mai 2005 [4]. Un des projets importants dans ce domaine est la Global Health Library (GHL) [5]. Son but est de fournir un point d'accès aux informations médicales fiables grâce à une plate-forme virtuelle.

Celle-ci rassemble des sources multiples et oriente les différents groupes d'utilisateurs (ministères de la Santé, acteurs de la santé, fournisseurs d'information, patients, public général) vers des contenus utiles à leurs besoins respectifs. Ce projet facilite également la relation entre les fournisseurs et les utilisateurs d'information et vise une couverture globale avec un effort particulier sur les pays en développement. Les outils basés sur Internet peuvent également être employés pour assurer la formation continue à distance des professionnels de la santé, pour autant que les contenus et l'organisation des cours soient adaptés aux réalités locales [6]. C'est un domaine de recherche important pour la chaire UNESCO de télémédecine de l'Université de Genève, établie en 2005.

L'Institut Universitaire d'Études du Développement étudie, depuis plus de quarante ans, les conséquences socio-économiques des développements technologiques sur les sociétés du Sud et les représentations sociales qui y sont associées. Plusieurs recherches ont analysé les transformations sociales issues du développement des nouvelles technologies de l'information [7]. En outre, l'expérience socio-anthropologique de l'Institut des Sciences Humaines du Mali a contribué à identifier les caractéristiques spécifiques des représentations sociales associées aux TIC.

2 Méthodes

Plusieurs activités furent menées de front : l'inventaire des principales sources de contenus médicaux en ligne en Afrique francophone ; l'application des principes du HONcode pour l'accréditation des sites Web identifiés et le recueil des problèmes rencontrés lors de leur évaluation ; l'identification des critères de confiance spécifiques aux contenus médicaux en ligne au moyen d'une approche socio-anthropologique (entretiens semi-directifs et questionnaires).

2.1. L'inventaire des principales sources de contenus médicaux en ligne en Afrique francophone

L'inventaire des sources d'information de santé en Afrique francophone a été effectué en mobilisant deux réseaux de professionnels actifs dans ce domaine. D'une part, l'OMS travaille sur le Global Index Medicus, qui rassemble et structure la liste des sources d'information de santé [8]. D'autre part, le Réseau en Afrique Francophone pour la Télémédecine (RAFT) qui mobilise chaque semaine des centaines de professionnels de la santé pour des cours de formation continue universitaire à distance. Les cadres et coordinateurs de ce réseau, actif dans dix pays d'Afrique francophone, ont participé à ce projet, de même que les correspondants académiques membres du réseau de l'UNESCO dans le domaine de la cybersanté.

2.2 L'application des principes du HONcode

L'évaluation des sites Web identifiés lors de l'inventaire a été effectuée par les collaborateurs du siège malien de la Fondation Health On the Net à Bamako, en utilisant la méthodologie développée et appliquée par la Fondation HON depuis dix ans. Il s'agit d'un questionnaire structuré permettant d'évaluer le respect de huit principes de bonne conduite constituant le HONcode [9]. La fondation HON a fourni, depuis Genève, un support méthodologique et une coordination technique. Les divergences par rapport aux critères du HONcode ont été systématiquement relevées et analysées pour identifier d'éventuelles régularités dont il faudrait déterminer si elles motivent une adaptation du HONcode, et/ou s'il s'agit de développer des compétences spécifiques pour la production de contenus dignes de confiance.

2.3 Approche socio-anthropologique

L'Institut Universitaire d'Études du Développement et l'Institut des Sciences Humaines du Mali ont identifié les critères qui déterminent la confiance que peuvent accorder les professionnels de la santé à l'information médicale en ligne. Cette enquête s'appuie sur quatre-vingt entretiens semi-directifs menés auprès d'un échantillon de professionnels de la santé maliens dans différents centres de santé (hôpitaux universitaires, Centres de santé de référence, Centre de santé communautaire) à Bamako et dans la région de Mopti et auprès des participants du réseau RAFT lors des Journées Francophones d'Informatique Médicale (JFIM) des 10 et 11 janvier 2007 à Bamako. Un questionnaire – distribué lors des JFIM – a permis de compléter les données qualitatives récoltées par le biais des entretiens semi-directifs

2.4 Enquête en ligne (Internet)

La fondation HON, l'IUED et l'ISH ont réalisé une enquête en ligne intitulée « Enquête sur l'utilisation de l'Internet médical et de santé en Afrique francophone » pour déterminer comment les utilisateurs et les éditeurs de sites Web accordent leur confiance à l'information en ligne. Ceux-ci ont ainsi appris l'existence d'initiatives pour améliorer la qualité de l'information de santé en ligne, tel que la certification HONcode.

3 Résultats

3.1 L'absence de marché local de l'information médicale en ligne

Les médecins spécialistes – généralement des hommes plus âgés et bien formés – disposent de multiples possibilités d'accès à de l'information médicale de qualité. Cette « élite » médicale (médecins-chefs, professeurs, chercheurs) ne produit pas d'informations médicales destinées au marché local. En général, ces professionnels s'intéressent aux informations spécialisées produites au Nord et n'œuvrent pas à la démocratisation de l'accès à l'information ou à la production de contenus à caractère plus local. À l'autre extrémité de la hiérarchie médicale, les infirmiers et d'autres professionnels de la santé n'ont pas accès à l'information et sont donc laissés de côté. Au plan intermédiaire, l'on trouve des médecins internes, des généralistes, des infirmiers-majors et des femmes médecins spécialistes qui ont un accès limité à l'information médicale. En outre, lorsqu'une connexion Internet existe, ils n'ont que rarement accès aux bases de données spécialisées. Les efforts pour améliorer l'accès à l'information médicale de qualité devraient donc se concentrer sur ce groupe de professionnels.

Ainsi, l'«élite» hospitalo-universitaire privilégie l'information produite au Nord, sans se préoccuper de la production locale. Il n'existe donc pas de production locale d'informations qui pourraient intéresser les autres professionnels. Cela n'incite pas ces professionnels à chercher des informations dans les sources existantes qui ne sont souvent pas applicables dans leur contexte clinique. Le sentiment de méfiance des professionnels de la santé vis-à-vis de l'information produite localement renforce cette carence d'informations locales. Il n'y a que peu de publicité sur les sites d'Afrique francophone et la présence de sites commerciaux est très faible : c'est un autre signe de l'absence d'un marché local de l'information médicale en ligne.

Figure 1 - Accès des professionnels de la santé aux différentes informations médicales. (Entretiens IUED, n= 80.)

L'Afrique sub-saharienne représente un faible marché économique pour les médicaments et les technologies de l'information et de la communication (figures 2, 3 et 4).

En outre, il n'existe que peu de moyens économiques pour diffuser des informations médicales éditées en Afrique (que ce soit sur papier ou en ligne), contrairement aux modèles occidentaux (grands éditeurs scientifiques qui possèdent des centaines de journaux, ouvrages de référence ou base de données en ligne).

Les contenus produits en Afrique ont donc tendance à être publiés dans les journaux et les sites du Nord.

Des moyens pour dynamiser l'accès et la production de telles informations existent pourtant, mais ils sont peu utilisés : OpenAccess (financement des frais d'édition par les auteurs) ou soutiens gouvernementaux à l'édition scientifique.

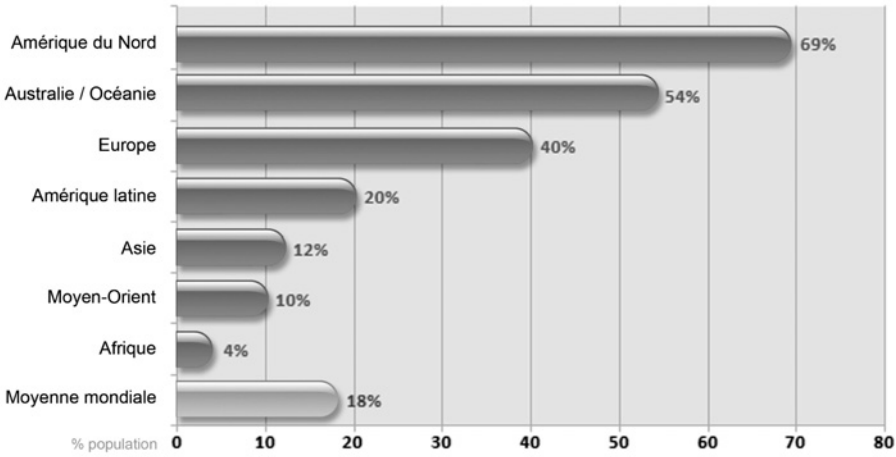


Figure 2 - Nombre d'utilisateurs de l'Internet, par région du monde. (www.internetworldstats.com, 2007).

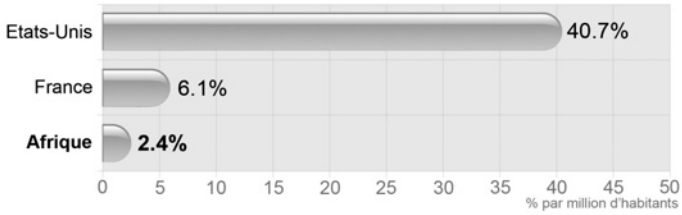


Figure 3 - Répartition par pays des sites de santé certifiés HONcode. (Résultats issus de la certification HONcode par la Fondation Health On the Net; n =6 258)

3.2 La qualité éditoriale du contenu des sites web produits en Afrique francophone

En termes de respect des critères HONcode, on se retrouve dans une situation similaire à celle qu'on mesurait il y a 5 ans dans les pays occidentaux. Internet est essentiellement utilisé pour un accès en lecture, généralement anonyme, ce qui explique en partie le peu de soucis liés à la confidentialité.

Remarque : le but de ce tableau est de mettre en évidence qu'en termes de respect des critères HONcode, on se retrouve en Afrique dans une situation similaire à celle qu'on mesurait entre cinq ans et deux ans auparavant dans les pays occidentaux. En effet les chiffres relatifs aux principes HONcode respectés entre l'Afrique en 2007 et l'Europe en 2001 et 2005 sont similaires pour les principes tels que l'autorité, la complémentarité, la date, la justification et le contact. À noter également que pour l'Europe de 2001 à 2006, nous constatons une augmentation du respect des principes (complémentarité, confidentialité, références, date, financement et publicité) en fonction du temps et liée à la prise de conscience des Webmasters et des éditeurs de sites Web aux standards du HONcode.

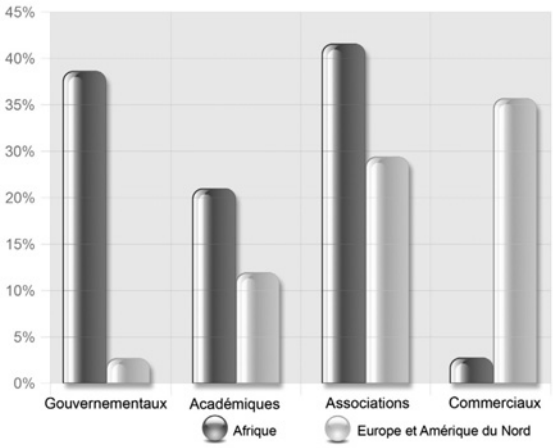


Figure 4 - Distribution des sites de santé certifiés HONcode en Afrique francophone, en Europe et Amérique du Nord.

(Étude du contenu des sites Web en Afrique francophone sub-saharienne par la Fondation Health On the Net, n=116; et n=5 281 sites certifiés HONcode en Europe et en Amérique du Nord.)

Nos travaux mettent en évidence un décalage entre les informations trouvées sur les sites médicaux et les attentes des professionnels de la santé. L'information scientifique avec 82 % semble la plus recherchée par les professionnels de la santé, alors que uniquement 25 % des sites en Afrique francophone offrent des articles scientifiques en ligne.

Les informations sur la pratique médicale, comme les expériences thérapeutiques et le diagnostic sont également recherchées (60 % et 56 %). Toutefois, moins de 8 % des sites en Afrique francophone répondent à cette attente. Alors que, près de la moitié des professionnels de la santé recherchent des formations universitaires, seuls 2,5 % des sites en Afrique francophone donnent accès à des cours en ligne. Ceci souligne la nécessité de fournir un contenu répondant aux besoins des utilisateurs en Afrique francophone.

Principes HONcode	AFRIQUE** 2007*	EUROPE 2001*	EUROPE 2005*	EUROPE 2006*
Autorité	93%	82.5%	87.3%	81.2%
Complémentarité	68.6%	64%	81%	86%
Confidentialité	4.8%	53.6%	70%	75.3%
Références	94.3%	83.8%	90.8%	91.9%
Date	58.1%	61%	76.3%	79.7%
Justification	99%	97.9%	99.7%	99.7%
Contact	96.2%	81.5%	98.2%	99%
Sources de financement	77.2%	46.4%	72.5%	77%
Honnêteté dans la publicité	82.8%	38.9%	61.1%	74.4%

* pourcentage de sites par principe respecté
** Afrique francophone

Figure 5 - Comparaison des principes HONcode respectés entre les sites médicaux et de santé entre l'Afrique francophone sub-saharienne et l'Europe de 2001 à 2006. (Résultats issus de la certification HONcode par la Fondation Health On the Net.)

3.3 La méconnaissance des enjeux liés à la qualité de l'information en ligne

La question de la fiabilité des informations médicales en ligne n'est pas évoquée spontanément dans les entretiens. En outre, les systèmes d'accréditation de l'information médicale en ligne sont peu connus.
Les critères de fiabilité se répartissent en deux groupes égaux : la moitié sont relatifs au contenu de l'information, l'autre moitié à tout ce qui « enveloppe » le contenu informationnel proprement dit, c'est-à-dire des critères « formels ». Ainsi, les professionnels de la santé confrontent l'information collectée sur Internet avec leur propre expérience (19 % des occurrences sémantiques), puis ils se fient à celle d'une tierce personne (10 %). Ensuite, ils examinent le titre (7 %) et lisent les résumés (5 %). Les critères de fiabilité « formels » sont relatifs au pouvoir symbolique de l'institution au sein de laquelle est produite l'information (par exemple une revue, un site ou une institution) (14 %), à l'attribution (mention de l'origine et la date de l'information) (11 %) et à l'autorité (10 %).
Ainsi, nous avons constaté une carence de regard critique sur l'information : sa qualité ou sa fiabilité n'est pas perçue comme une question importante et elle n'apparaît jamais spontanément. Ces résultats sont comparables à ceux de Blumberg et Sparks (1999) [11] qui ont interrogé des étudiants américains à la fin des années quatre-vingt-dix. Il ne paraît donc pas y avoir de différences socioculturelles clairement identifiables dans la construction de la confiance.

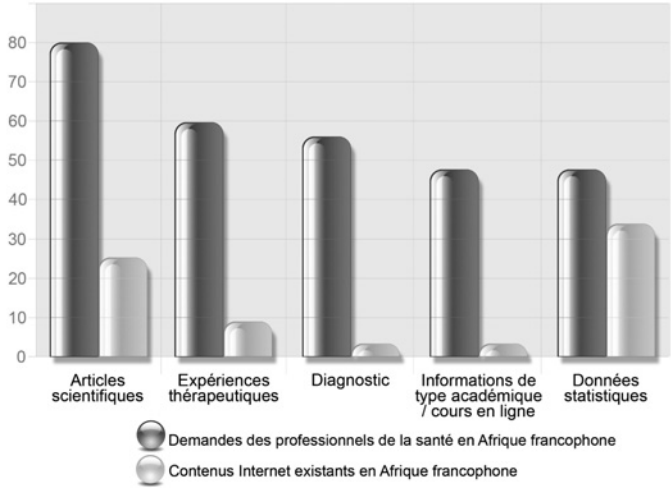


Figure 6 - Comparaison des demandes des professionnels de la santé en Afrique francophone par rapport aux contenus Internet santé existants en Afrique francophone. (Étude du contenu de sites Web en Afrique francophone sub-saharienne par la Fondation Health On the Net, n = 116 ; enquête en ligne sur l'utilisation de l'internet médical et de santé en Afrique francophone par IUED-HON; n = 54.)

3.4 Les barrières à l'accès à l'information médicale en ligne

La carence d'infrastructure informatique dans certains établissements ou services médicaux constitue l'obstacle le plus important pour l'accès à l'information médicale en ligne. Dans les Centres de Santé communautaires, les Centres de Santé régionaux et

certaines cliniques privées, l'informatique est très peu présente, voire inexistante. Dans les deux hôpitaux universitaires, le nombre d'ordinateurs connectés varie sensiblement d'un service à l'autre, de même que la qualité de la connexion. La dimension hiérarchique influence la manière dont le média Internet est employé, tant d'un point de vue quantitatif que qualitatif. Les plus grands utilisateurs sont les médecins spécialistes, les internes ainsi que les jeunes médecins qui viennent de terminer leur internat. Les infirmiers, les sages-femmes et les aides-infirmiers ne consultent pour ainsi dire pas ce média. Les pratiques des médecins diffèrent beaucoup selon leur statut professionnel. Les généralistes se servent le plus souvent d'un moteur de recherche ordinaire (Google) alors que les médecins spécialistes emploient Internet de manière plus précise et privilégient les moteurs de recherches spécialisés (Pubmed, INIST, CISMEF).

Les hommes utilisent davantage Internet que les femmes, tout statut confondu. Cette situation est moins liée à une situation d'inégalité d'accès qu'au phénomène de «double journée ». Alors que les médecins consultent souvent Internet en fin de journée ou lors des week-ends, les femmes mariées doivent gérer les questions domestiques et familiales.
De même, alors que l'emploi d'Internet est généralement corrélé à l'implication du professionnel dans le domaine de la recherche, on observe que les femmes médecins sont plus engagées dans le champ clinique que dans celui de la recherche, ce qui diminue leur chance d'utiliser Internet.

L'usage d'Internet dans les centres de santé est souvent associé à un intermédiaire. Nous en avons recensé trois figures principales :

- Les personnes qui régissent l'accès à l'infrastructure. Les ordinateurs connectés se trouvent le plus souvent dans le bureau du médecin-chef. Les professionnels qui souhaitent se connecter à Internet doivent obtenir préalablement son aval.
- Les personnes qui régissent l'accès à l'information en ligne : certains professionnels passent par une tierce personne pour effectuer leur recherche sur Internet. Les médecins spécialistes mandatent des médecins généralistes ou des internes. Ces derniers font appel à des non-spécialistes (animateurs de cybercafé, responsables de la salle informatique, bibliothécaires, parents, etc.). Ces formations différentes de l'intermédiaire influence la gestion de la fiabilité de l'information en ligne et sa qualité. Les spécialistes – contrairement aux généralistes et aux internes – s'adressent à des professionnels de la santé plus aptes à distinguer une information fiable d'une autre qui ne l'est pas.
- Les personnes qui régissent l'accès aux informations spécialisées. Tous les sites médicaux ne sont pas en libre accès et s'y connecter suppose de détenir des codes. Les professionnels qui connaissent ces codes sont rares et ils les conservent jalousement. Ces intermédiaires sont généralement des médecins-chefs qui ont obtenu les codes grâces aux relations qu'ils entretiennent avec des institutions étrangères (Organisations internationales, institutions, hôpitaux) et à la reconnaissance dont ils disposent, au niveau international.

Doter en priorité les bureaux des médecins-chefs revient à réaffirmer la hiérarchie médicale traditionnelle. La reconnaissance acquise par ces intermédiaires est celle du système de santé, au niveau national. Le pouvoir des jeunes médecins qui effectuent les recherches sur Internet dépend de leur connaissance informatique. Cette fonction d'intermédiaire leur permet de renforcer leur position dans le service. Enfin, la reconnaissance que confère aux médecins-chefs la détention des codes d'accès aux bases de données spécialisées est issue de leurs relations avec les bailleurs de fonds et les pairs au niveau international.

Le média Internet contribue ainsi à reproduire l'organisation hiérarchique. Si les connaissances informatiques des jeunes médecins leur permettent d'acquérir une certaine reconnaissance de la part de leurs aînés, les rapports de pouvoir ne sont pas modifiés de manière déterminante pour autant.

Ce sont les médecins-chefs – et donc les acteurs les plus âgés – qui conservent les deux principales clés, celle du hardware (accès aux ordinateurs connectés) et celle du software (accès aux sites médicaux réservés).

4 Conclusion et recommandation

Ce projet de recherche a permis d'explorer plusieurs domaines liés à la production et à l'accès aux informations médicales en ligne pour les professionnels de la santé en Afrique francophone.

L'étude d'un important échantillon des sites médicaux d'Afrique francophone a précisé la qualité éditoriale et les mécanismes de production ; de plus, elle a révélé le décalage entre les informations publiées et celles qui sont attendues par les professionnels de la santé.

Les quatre-vingts entretiens conduits auprès des professionnels de la santé ont montré que les déterminants de la construction de la confiance ne semblent que peu différer de ceux connus dans les pays occidentaux. Toutefois, des spécificités quant à l'accès aux informations sont apparues. Il s'agit donc de :

- Veiller à l'équité de l'accès (connexions, bases de données, articles scientifiques, etc.) aux informations médicales en ligne, en particulier pour les femmes et pour les infirmiers et les majors.
- Développer l'accès aux informations médicales en ligne chez les médecins en formation, les généralistes, les professionnels paramédicaux expérimentés et les femmes professionnelles de la santé.
- Reconnaître et développer le rôle des intermédiaires dans l'accès à l'information médicale en ligne afin d'offrir des services de qualité.
- Sensibiliser les professionnels à la problématique de la fiabilité de l'information médicale en ligne. Ces derniers n'y accordent que peu d'importance en raison de l'accès souvent difficile à Internet.

Au terme de cette étude, nous formulons les recommandations suivantes :

a) Favoriser la production de contenus médicaux pour une audience locale

- Reconnaître l'existence d'un besoin de formation continue pour les médecins internes et généralistes, les spécialistes notamment les femmes, voire des infirmiers majors.
- Inciter les experts à produire des contenus spécifiques à ce groupe en collaboration avec les professionnels de la santé plus jeune et les professionnels de la gestion de l'information médicale (bibliothécaires, informaticiens). Ceux-ci pourraient en outre développer des compétences spécifiques pour la production et la diffusion de contenus dans le cadre de réseaux collaboratifs structurés.
- Promouvoir l'accréditation de la qualité des sites.

b) Améliorer la production d'information scientifique d'audience internationale

- Diffuser les connaissances produites dans les pays en développement en valorisant la publication en ligne de journaux locaux et de thèses dans le cadre du Global Index Medicus.

- Subventionner les publications en « OpenAccess » pour les auteurs des pays en développement.

c) Améliorer l'accès à l'information de qualité

- Promouvoir l'importance de la qualité et de la pertinence de l'information médicale en ligne auprès des professionnels de la santé et des étudiants.
- Promouvoir les ressources en ligne disponibles gratuitement (MEDLINE, PubmedCentral, BioMedCentral, HINARI, Global Index Medicus, HON, CISMEF [12], Institut de l'Information Scientifique et Technique du CNRS [13], portails médicaux nationaux, etc.).
- Renforcer la formation continue des professionnels de la santé, dont une partie peut être facilitée par l'utilisation du e-learning. Veiller à un accès équitable pour les femmes.
- Développer le rôle des intermédiaires de l'accès à l'information de santé, les positionner dans des situations stratégiques telles que des bibliothèques et les organiser en réseaux collaboratifs.
- Veiller à ce que l'accès à de l'information médicale en ligne de qualité ne se concentre pas entre les mains des seuls hommes et/ou de l'« élite » hospitalo-universitaire.
- Faire connaître les possibilités d'accès à l'information de qualité (HINARI, etc.) au plus grand nombre de professionnels et favoriser l'accès à cette information au moindre coût. Limiter ainsi les tentatives de rétention d'information.
- Développer une infrastructure technique performante et bon marché d'accès à l'Internet. Dans des situations de connectivité insuffisante, diffuser les contenus hors ligne, par exemple par le projet eGranary [14].

Remerciements

Ce projet de recherche a été financé par le Réseau Universitaire International Genevois, avec la collaboration de l'Université de Genève, de l'Organisation mondiale de la santé, de l'Institut Universitaire d'Études du Développement, de la Fondation Health On the Net, et de l'Institut des Sciences Humaines du Mali. Il a bénéficié de la participation volontaire de nombreux professionnels de la santé et de la cybersanté.

Références

[1] Wolton D. *L'Autre Mondialisation.* Flammarion, Paris 2003.

[2] Déclaration de Bamako : www.unige.ch/iued/wsis/DOC/046FR.PDF

[3] Boyer C, Geissbuhler A. A decade devoted to improving online health information quality. *Stud Health Technol Inform* 2005; 116: 891-896.

[4] www.who.int/healthacademy/news/eHealth_EB_Res-fr.pdf

[5] www.who.int/ghl/en/index.html

[6] Geissbuhler A, Ly O, Lovis C, L'Haire JF. Telemedicine in Western Africa: lessons learned from a pilot project in Mali, perspectives and recommendations. *J Amer Med Inform Assoc* 2003 S: 249-253.

[7] Steiner, B. 2006, « Entre bavardages et secrets. L'usage d'Internet dans les cybercafés de Bamako », *Tsantsa*, vol. n° 11 pp. 119-123. *La fin des outils : technologie et domination*, Cahiers de l'IUED, n° 5, 1977, Genève : IUED, Paris : PUF. Corminboeuf C. « Internet : l'illusion de la connaissance pour tous » In: *La mondialisation des anti-sociétés.* - Paris : Presses universitaires de France ; Genève : Institut universitaire

d'études du développement, 1997 ; 173-183. *Annuaire Suisse de politique de développement : Société de l'information et coopération internationale, development.com*, Genève : IUED, 2003.

[8] www.who.int/ghl/medicus/en/

[9] www.hon.ch/HONcode/Conduct_f.html

[10] www.who.int/workforcealliance/en/

[11] Blumberg P, Sparks J. Tracing the evolution of critical evaluation skills in students' use of the Internet. *Bull Med Libr Assoc* 1999; 87(2): 200-5.

[12] HINARI: www.who.int/hinari/en/ dernière visite le 18 novembre 2008.

[13] CISMEF: www.cismef.org, dernière visite le 18 novembre 2008.

[14] INIST: www.inist.fr, dernière visite le 18 novembre 2008.

[15] www.widernet.org/digitallibrary, dernière visite le 18 novembre 2008.

Adresse de correspondance

Dr Cheick Oumar Bagayoko
Service Informatique Médicale, Hôpitaux Universitaires, Genève
LERTIM, Faculté de Médecine, Université de la Méditerranée, Marseille
E-mail: cob@hcuge.ch

Informatique et Santé, 2009 (17) : 299-312
© Springer – Verlag France

Plate-forme pour la surveillance épidémiologique et l'aide à la décision en santé. Hivern@le – KhiObs : surveillance épidémiologique des pathologies hivernales de la sphère ORL chez l'enfant

Laurent Toubiana [1,2], Thomas Clarisse[1,2], François Colin[1,2], Hugo Alexandre[1,2] et Paul Landais[1]

1) Université Paris-Descartes - Faculté de Médecine AP-HP- UPRES EA 4067 - Service de Biostatistique et d'Informatique Médicale Groupe hospitalier Necker - Enfants Malades
2) Inserm, SCEPID « Système Complexe et Épidémiologie », France

Abstract

Background: An observatory of ENT diseases in children was set up. It aimed at estimating the incidence of acute episodes of rhinopharyngitis, otitis media, tonsillitis, sinusitis and laryngitis and evaluating their time-space evolution in winter in France. In addition, this study evaluated the ability of our surveillance platform generator to implement and manage such an observatory.

Methods: A network of practitioners, "observers", declares the cases cared during their consultation. The incidence of the observed diseases is estimated. Results are available on line by interactive and direct access to data, or by newsletters including weekly national and regional reports.

Results: The observatory operated during the winter seasons 2005-2006 and 2006-2007 with a network of general practitioners, ENT specialists and paediatricians. A cumulative number of 388,685 consultations have been collected during the two seasons. The mean number of consultations was then estimated at 18.6 millions for all of France during each winter for these diseases which corresponded to more than 1.65 consultation per child per season.

Conclusion: Over the last 2 winters, the Hivern@le-KhiObs observatory showed the very high incidence of ENT infections. This health platform helps to better understand the spread of these diseases during winter in order to contribute to a rationalization of the public health approach of these diseases.

Keyword

Communicable diseases; Acute ENT diseases in children; Multi-Source Information System; Health Networks; Public Health Informatics.

1 Introduction

Dans un article publié en 2006, Doherty *et al.* rapportent les travaux de la commission canadienne des maladies à déclaration obligatoire [1]. Celle-ci recommande que les maladies transmissibles à haute fréquence et faible gravité soient placées sous surveillance

en particulier lorsque les autorités sanitaires ne disposent pas ou peu de sources de données les concernant. Les pathologies ORL de l'enfant correspondent exactement à ces conditions. Ces pathologies sont très fréquentes et pourtant leur surveillance épidémiologique est quasiment inexistante. En France, les infections ORL de l'enfant surviennent essentiellement en hiver. Elles sont alors le principal motif de consultation chez les pédiatres et les médecins généralistes [2]. Après avoir exploré la littérature internationale dans ce domaine, il s'avère que les données d'incidence sur ces affections sont rares, peu précises et le plus souvent fondées sur des enquêtes restreintes.

Dans ce contexte, il apparaissait opportun de mettre en place un observatoire national pour la surveillance épidémiologique de ces pathologies pendant la saison hivernale. L'objectif était d'estimer la fréquence hebdomadaire, par département français métropolitain et par tranche d'âge des épisodes infectieux aigus de la sphère ORL et de fournir au corps médical et aux décideurs de santé publique les moyens d'information sur l'évolution spatio-temporelle de ces pathologies.

L'objectif secondaire de cette étude était de démontrer la faisabilité de la mise en place d'un tel observatoire sur la base du générateur de plate-forme de surveillance temps réel que nous avons élaboré dans le cadre du programme de recherche intitulé « MetaSurv » [3 , 4]. Ce programme bénéficie d'une expérience de plus de 15 ans dans le domaine de la surveillance en temps réel des maladies transmissibles [5] et a pour but de diffuser auprès des acteurs de santé des moyens pour la mise en place de systèmes de surveillance complets, à coût réduit et dans des délais courts.

Cet article relate l'expérience de la mise en place de l'observatoire « Hivern@le-KhiObs » grâce à la plate-forme MetaSurv. Il s'agit d'une étude observationnelle descriptive nationale. La durée annuelle de l'étude est de 4 mois de décembre à mars correspondant aux saisons hivernales des années 2005-2006 et 2006-2007. L'étude concerne les populations d'enfants des 2 sexes âgés de 6 mois à 15 ans résidant en France métropolitaine au moment de l'étude.

2 Matériel et méthodes

2.1 Mise en œuvre et utilisation de la plate-forme

L'observatoire des pathologies hivernales de la sphère ORL chez l'enfant repose sur un système d'information permettant le recueil en temps réel de données épidémiologiques.

Le système d'information, fondé sur une architecture de type n-tier, est une plate-forme LAMP[1] accessible via internet. L'interface graphique utilise les standards web, feuille de style (CSS), XHTML (HyperText Markup Language), Ajax (Asynchronous JavaScript And XML).

2.1.1 Divers modes de fonctionnement en fonction des types d'utilisateurs

Sur cette plate-forme, toute entité possédant un accès sur le site, c'est-à-dire ayant un code et un mot de passe est appelée dans la suite du texte « utilisateur ». La notion de « type d'utilisateurs » est définie par le mode d'interaction avec la plate-forme. Pour chaque type d'utilisateur correspond un ensemble de droits d'accès à certaines fonctionnalités de la

[1] Linux debian, Apache_2, MySql_5 and PHP_5

plate-forme. L'ensemble de ces droits d'accès satisfait à un mode de fonctionnement qui lui-même correspond au final à une interface utilisateur spécifique. Par exemple, le type prédéfini « administrateur » correspond au mode ayant le plus large accès aux fonctionnalités alors que le type « visiteur » est celui ayant les droits d'accès les plus restreints. Entre ces 2 types extrêmes, il est potentiellement possible de créer une infinité de types d'utilisateurs en définissant les fonctions du site auxquelles ils auront accès. Lors de la mise en place de l'observatoire « Hivern@le-KhiObs », nous avons généré une dizaine de type d'utilisateurs parmi lesquels le type « observateur » qui constitue la grande majorité des utilisateurs dans la mesure où ce type est dédié aux médecins du réseau afin que ceux-ci puissent effectuer la saisie des cas. Nous décrivons ci-dessous quelques fonctionnalités principales relatives au mode « administrateur » afin de montrer un panorama des fonctions disponibles sur cet observatoire.

L'administrateur a pour rôle de superviser l'ensemble de l'activité de l'observatoire. Globalement, cette supervision peut se subdiviser en deux catégories principales d'interactions ; d'une part la gestion des utilisateurs du site et l'animation du réseau et d'autre part la gestion des données recueillies par les médecins-observateurs (« observateurs »).

2.1.2 Interface générique aux données ou « fouille de données »

Quel que soit le type d'utilisateur, l'essentiel des interactions avec l'observatoire consiste dans l'absolu à accéder à la base de donnée du système d'information. En d'autres termes et schématiquement, il est possible de considérer l'observatoire comme une interface en ligne à une base de données. Bien sûr, selon le type d'utilisateur (tel qu'ils ont été évoqués plus haut) et donc selon son niveau supposé de compétence vis-à-vis des langages d'interrogation des bases de données, nous avons développé des interfaces adaptées de façon à rendre ces interrogations transparentes. Ainsi nous montrerons comment le formulaire de saisie permet de saisir rapidement l'information épidémiologique spécifique au protocole de l'étude. Néanmoins, dans la plus grande majorité des cas, l'interaction avec les données peut s'effectuer par l'intermédiaire de tableaux (données tabulées) qui se révèlent être un mode à la fois intuitif et efficace d'interrogation.

Nous avons donc créé une interface générique fondée sur ce principe. Il s'agit d'un outil qui permet la présentation des données contenues dans une table d'une base de données mais également leur manipulation interactive. La combinaison des filtres et des tris permet d'accéder instantanément aux enregistrements contenant l'information recherchée. Il est possible d'effectuer des opérations sur un ou plusieurs champs disposés en colonnes et/ou sur un ou plusieurs enregistrements disposés en lignes. De très nombreuses fonctions ont été développées : navigation par page, affichage du nombre total d'enregistrements et de pages obtenus suite à une recherche, choix du nombre d'enregistrements par page, tris croissants et décroissants sur le contenu des colonnes, sélection des champs (colonnes) à afficher, filtres sur les colonnes pour sélectionner l'information à afficher, changement de l'apparence du tableau (couleur, espacement...), édition des valeurs d'une cellule ou d'une colonne avec une aide à la saisie, suppression d'enregistrements, adjonction de liens vers des outils afin d'effectuer des traitements complexes sur le (ou les) champ(s) des enregistrements sélectionnés et visibles dans le tableau.

Cet outil se présente sous la forme d'une librairie qu'il est possible d'intégrer sur n'importe quel serveur web. Cette librairie développée en Ajax, améliore considérablement le confort de l'interface utilisateur par la réactivité due à l'accès asynchrone aux données tout en évitant le rechargement de page sur le navigateur pendant son utilisation. Cette interface

est largement utilisée dans le site notamment en mode administrateur pour gérer les utilisateurs, leur activité, les données saisies ainsi que les résultats obtenus par le traitement de ces données.

2.1.3 Authentification

L'accès à la plate-forme est sécurisé et protégé par un identifiant et un mot de passe. Ces identifiants doivent correspondre à un compte activé selon un certain mode d'utilisation. C'est l'administrateur qui attribue au moment de l'inscription le mode d'utilisation en affectant le nouvel inscrit à un type d'utilisateur.

Pour sécuriser la navigation sur le site, une vérification de l'identité du visiteur est effectuée sur chaque page. Si une incohérence (changement des droits d'accès, changement inattendu du mot de passe, inactivité pendant plus de 10 minutes) est détectée, le visiteur doit se reconnecter. Chaque visite sur le site est enregistrée dans la base de données avec des informations concernant la date, la provenance des utilisateurs, le navigateur utilisé et le système d'exploitation. Ces données permettent de connaître plus en détails les utilisateurs de la plate-forme et leurs habitudes de navigation. En accord avec la loi « Informatique et Liberté » du 6 janvier 1978, les médecins ont la possibilité d'accéder et de rectifier les informations personnelles les concernant.

2.1.4 Tableau de bord et statistiques de connexions

L'interface du mode administrateur permet d'obtenir une vision globale de l'activité et les interactions de tous les types d'utilisateurs avec le site. Pour cela, il dispose d'un outil de synthèse sous forme d'un tableau généré par le module présenté ci-dessus. À partir de ce « tableau de bord » qui affiche des statistiques en temps réel, il accède à un grand nombre de fonction de visualisation et de gestion de l'observatoire parmi lesquelles : ajouter des utilisateurs, gérer les demandes d'inscriptions, modifier les informations concernant les observateurs, gérer (avec tact) les relances des observateurs n'ayant pas suivi le protocole, envoyer rapidement des courriels à un ou plusieurs utilisateurs, obtenir le détail des connexions de chaque utilisateur, obtenir le détail des saisies de chaque observateur, obtenir des statistiques concernant l'utilisation de la plate-forme, gérer les listes d'événements pour assurer un suivi de chaque utilisateur, créer des types d'utilisateurs auxquels il peut déléguer tout ou partie de ces tâches.

Par l'intermédiaire d'un code couleur, il met en évidence le temps écoulé depuis la dernière saisie des observateurs. L'administrateur peut alors sélectionner les observateurs les moins réguliers et leur envoyer un message de relance qui leur rappelle le fonctionnement de l'étude et les invite à se connecter rapidement sur le site pour rapporter les cas observés.

Chaque utilisateur est lié à une liste d'événements le concernant. Cette liste est maintenue à jour soit par l'administrateur qui y reporte les contacts qu'il a eu avec l'utilisateur concerné, soit par le système qui ajoute automatiquement les principaux événements comme par exemple les dates de connexions ou de saisies, l'envoi d'un courriel ou un appel téléphonique.

Ainsi, quand un observateur appelle la hotline de l'observatoire ou envoie un courriel pour obtenir des informations, l'opérateur a un accès instantané à l'historique de l'observateur en récupérant le contexte des étapes déjà accomplies et les problèmes auxquels il a été confronté.

Les gestionnaires du site (utilisateurs auxquels l'administrateur a délégué certaines tâches de gestion) peuvent rechercher rapidement des utilisateurs en fonction de critères de

filtrage (nom, prénom, login, courriel, sexe, département, etc.), modifier les données, accéder au suivi de chacun d'entre eux, lister les saisies effectuées et réaliser éventuellement des corrections et d'autres opérations plus complexes programmables et ajoutées en lien comme outil dans l'interface lui-même. Un exemple d'outil accessible directement par le tableau est un module permettant l'accès aux données via des cartes interactives. Nous utilisons l'API Google Maps pour générer des cartes afin d'accéder rapidement à l'information sur le nombre de médecins présents dans chaque département et les informations principales les concernant. Ces cartes contiennent plusieurs couches d'informations : le fond de carte qui permet de repérer les médecins avec une géo-localisation sur l'adresse de leur cabinet (20 000 médecins). Les observateurs sont quant à eux repérés par des marqueurs codés en couleur en fonction de leur date de dernière connexion afin de visualiser instantanément la localisation des médecins « en retard » par rapport au protocole de saisie ; des plages départementales de couleurs, du vert au rouge (choroplèthe), indiquent la densité d'observateurs dans les différents départements ; des marqueurs de préfectures de département, des infos-bulles affichent, lors d'un clic sur une préfecture, la liste des médecins participant dans ce département ; une fiche médecin s'affiche quand on clique sur le nom d'un médecin dans l'info-bulle. Ce module intégré sous forme d'outil accessible par le tableau dispose d'une fonction de recherche (Ajax) qui permet de centrer la carte sur le département et de géo-localiser à l'adresse du cabinet du médecin recherché afin d'accéder plus rapidement aux informations le concernant.

2.1.5 Envoi de courriels

La gestion des courriels est un module important du site. Celle-ci est elle aussi intégrée en tant qu'outil associé du tableau. Tous les courriels sont liés à des événements automatiquement ajoutés à la liste des événements de chaque destinataire. Les courriels envoyés via le site sont généralement interactifs. Ils sont conçus de manière à induire des clics (par exemple questions posées au destinataire). Chaque clic sur un courriel provoque des réactions automatiques du serveur programmables facilement par l'administrateur du site. Ceci permet d'effectuer des statistiques sur les envois électroniques mais aussi d'automatiser un grand nombre de communication entre l'observatoire et ses utilisateurs au sens large (cf. supra). Ce système s'est révélé être d'une grande efficacité pendant la phase de déploiement de l'application.

2.1.6 Téléchargement de document et animation du réseau d'observateurs

L'espace document a été séparé en deux parties : les documents officiels de l'étude, les autres documents en partage, mis à jour par l'administrateur.

Le système permet de protéger les documents confidentiels mis en ligne sur le site. Pour pallier le problème de téléchargement illicite, nous utilisons la base de données pour stocker non seulement le type de fichier, sa taille et son nom, mais aussi son contenu. Une page spéciale d'accès aux documents vérifie l'identité et les droits de l'utilisateur puis reconstitue le document et l'envoie directement au client. Ainsi les documents ne sont plus stockés physiquement sur le serveur et ne peuvent donc pas être téléchargés sans autorisation.

Nous avons évoqué ci-dessus un certain nombre d'outils permettant une gestion rigoureuse de l'activité du réseau, ce qui constitue un élément nécessaire pour une adhésion au projet. Cependant notre expérience dans le domaine montre que l'animation (au sens large) du réseau d'observateurs est l'un des aspects fondamentaux du bon fonctionnement d'un observatoire, notamment en termes de compliance au protocole de l'étude. Il serait trop long dans le cadre de cet article, de développer les nombreux autres outils disponibles pour animer le réseau des médecins observateurs. Citons que le serveur dispose de modules de

partage de documents, de widget, de news, de micro-sondage et d'une manière générale privilégie le retour d'information.

2.2 Modalité de recueil des données

Les médecins ORL, généralistes et pédiatres (« observateurs ») reçoivent les enfants dans le cadre habituel de leur consultation ou des visites à domicile sans modification de leurs habitudes diagnostiques et thérapeutiques.

2.2.1 Formulaire électronique de saisie

Dans un esprit d'efficacité, nous avons estimé que pour un observateur, la saisie des cas devait prendre le moins de temps possible. Dans le cadre de cette étude, étant donné le faible nombre d'information nécessaire pour décrire chaque cas, nous avons conçu un mode de saisie agrégé. Sur la base de cette saisie, un programme prend en charge de manière transparente la description individuelle de chaque cas qui permettra ensuite des calculs d'incidence à des résolutions de temps et d'espace très fins (cf. ci-après : Méthodes d'estimation des taux d'incidence). La figure 1 montre une capture d'écran du formulaire électronique de saisie accessible sur site web dédié à l'observatoire. Au moins une fois par semaine, le jour de la semaine de leur choix, les médecins observateurs déclarent sur le côté droit de ce formulaire, pour chaque catégorie (maladie, sexe et tranche d'âge), le nombre de cas examinés au cours de la dernière période (depuis leur dernière connexion).

Afin d'améliorer la qualité et la cohérence de la saisie, les médecins disposent d'une visualisation instantanée de statistiques synthétiques concernant leur propre saisie ainsi que celle des autres membres du réseau, agrégée par niveau géographique (département, région, France entière). Le système d'information effectue automatiquement des contrôles de cohérence au moment de la saisie. À la fin de la saisie, les médecins observateurs sont automatiquement dirigés vers un rapport de synthèse sur les pathologies observées de la semaine courante.

	Masculin			Féminin				Résultats précédents			
	6mois - 2ans	2ans - 8ans	8ans - 15ans	6mois - 2ans	2ans - 8ans	8ans - 15ans	TOTAUX	Totaux de votre dernière déclaration	Totaux de toutes vos déclarations	Totaux sur votre région	Totaux sur la France
OMA	0	0	0	0	0	0	0	0	13	132	2781
Rhino	0	0	0	0	0	0	0	0	26	493	11050
Angine	0	0	0	0	0	0	0	0	5	122	2684
Sinusite	0	0	0	0	0	0	0	0	0	26	827
Laryngite	0	0	0	0	0	0	0	0	0	57	1669

Valider | Sortie sans valider

Figure 1 - Capture d'écran du formulaire électronique.

2.2.2 Contrôle automatique des saisies

L'intervalle de temps entre deux saisies ne doit pas être trop faible afin de ne pas fausser les résultats de l'étude. Pour éviter ce problème, une vérification automatique interdit de saisir deux fois dans la même journée. Si un observateur retourne sur la page de saisie, une alerte lui annonce que la validation de la nouvelle saisie remplacera la précédente. De même, si un observateur revient moins d'une semaine après avoir effectué une saisie, une alerte s'affiche lui proposant soit d'ajouter une nouvelle saisie soit de remplacer la

précédente. Ainsi les observateurs ont un droit à l'erreur sans pour autant pouvoir modifier toutes leurs saisies.

2.2.3 Gestion des saisies

Seul l'administrateur a accès à la liste de toutes les saisies effectuées sur le site et peut les modifier (date de la saisie et nombre de cas observés) ou les supprimer. Il peut obtenir une liste des saisies accompagnées du nombre de cas déclarés pour chaque maladie. Il a la possibilité de modifier le nombre de cas ainsi que la date de la saisie suite à la demande motivée d'un observateur. Cette possibilité a une influence notoire sur l'ensemble des calculs permettant notamment d'estimer les incidences. Elle implique que ceux-ci soient régénérés en temps réels. Dans la pratique, elle est donc utilisée avec parcimonie et beaucoup de précaution.

2.3 Méthodes d'estimation des taux d'incidence

Le taux d'incidence T pour une pathologie, pour un groupe de population (tranche d'âge, sexe) en un lieu donné au cours d'une période est estimé ainsi :

$$\mathbf{T = (C/P)*(N/Pop)}$$

Avec :
C : somme des contributions quotidiennes
P : somme des participations quotidiennes
N : effectif total des médecins, et Pop : effectif total du groupe de population observé

Cette estimation se fonde sur les tables de populations de l'Insee pour les années concernées, par département et tranche d'âge de 1 an. La contribution quotidienne d'un médecin est obtenue en distribuant sur chaque jour d'une période définie par l'intervalle entre ses 2 dernières connexions, le nombre de cas déclarés lors de sa dernière connexion, rapporté au nombre de jours de cette période. La participation quotidienne correspond à une pondération de l'activité du médecin en fonction de la fréquence de ses connexions valides. Il est ainsi possible d'estimer par agrégation toutes les incidences dans l'espace d'observation (département, région ou France entière) selon la résolution temporelle désirée (jour, semaine, mois, saison) rapportées à la population de la tranche d'âge et/ou sexe et en fonction de la spécialité des médecins (généraliste, pédiatre, ORL). Tous les calculs sont mis à jour automatiquement à fréquence régulière (en l'occurrence chaque jour) et stockés dans un entrepôt de données afin d'optimiser les temps d'accès pour la génération des synthèses (tableaux/cartes/graphes).

2.4 Retour d'information

Un bulletin hebdomadaire est disponible en ligne et donne les résultats essentiels obtenus au niveau national par l'observatoire au cours de la semaine précédant la semaine en cours sous forme de cartes, courbes et tableaux. Un flux d'information RSS [2] ainsi que des lettres d'information nationales et régionales sont ainsi générés. Ces supports sont diffusés chaque semaine à toute personne inscrite pour la recevoir. Ils permettent de tenir les internautes informés des nouveaux bilans mis en ligne sur « Hivern@le-KhiObs ». Par ailleurs, des outils accessibles pour certain types d'utilisateurs permettent de générer en temps réel et à la demande des tableaux, des cartes et des courbes pour représenter les résultats de l'étude en fonction des différentes variables (sexe, maladies, tranches d'âges, spécialité des observateurs, départements et régions).

---

[2] RSS (*Really Simple Syndication*) désigne une famille de formats XML utilisés pour la syndication de contenu Web. La syndication consiste à diffuser un programme à plusieurs diffuseurs.

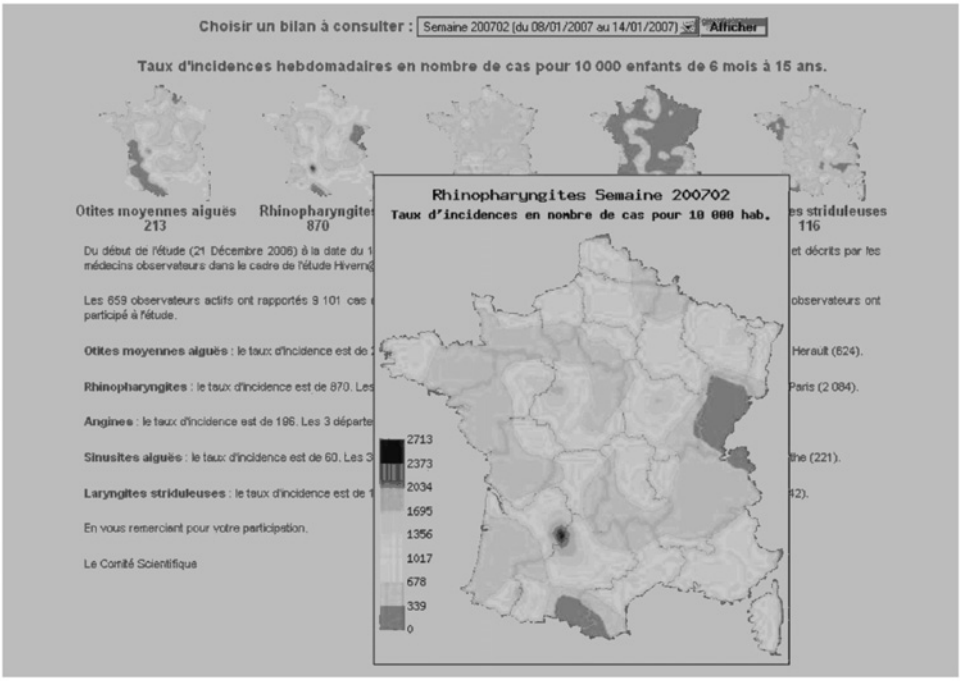


Figure 2 - Bilan hebdomadaire pour la France entière.

2.5 Mise en place du réseau de médecins observateurs

En fusionnant plusieurs bases de données commerciales concernant l'activité des médecins (domaine d'activité, pourcentages de patients traités dans les différentes populations, les attitudes de prescription, etc.), nous avons extrait un fichier d'environ un tiers de l'effectif total de médecins en France métropolitaine. Ce sous-ensemble de près de 20 000 médecins contient les informations concernant des médecins généralistes ayant une forte proportion d'enfants parmi leur patientèle, des ORL et des pédiatres.

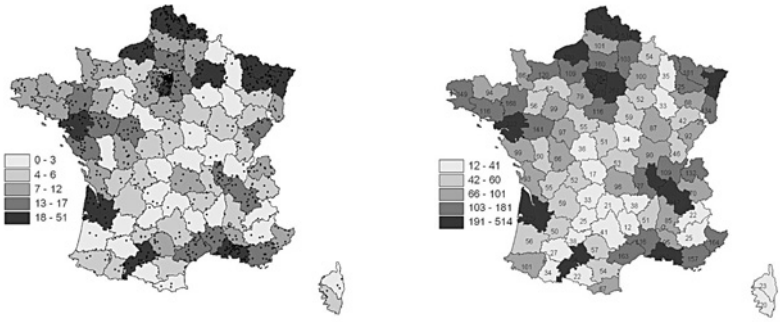


Figure 3 - a) Distribution départementale des médecins observateurs participants (à gauche). b) distribution départementale des populations départementales d'enfants de 15 ans et moins (à droite).

Un courrier papier comprenant un dossier d'information sur la mise en place de cet observatoire et un coupon retour pour acceptation de participation a été adressé aux médecins de ce fichier. Le protocole de l'étude prévoyait au moins une connexion par semaine pour saisir les cas. Chaque connexion était rémunérée à raison de 20 euros par semaine de surveillance pour 12 connexions au maximum. Plus de 40 % des médecins contactés ont répondu favorablement pour une participation à cet observatoire.

3 Résultats

3.1 Saison 2005-2006

La première saison, hiver 2005-06, a été celle de l'initialisation du programme. L'échantillonnage des médecins a été réalisé par un sondage en grappe de degré 1 au niveau départemental en tirant aléatoirement parmi les médecins dont nous avions reçu l'acceptation par écrit pour une participation. Nous nous sommes assurés de la représentativité selon les critères de sexe, d'âge selon les données de l'Insee et de l'Irdes (figure 3). Nous avons contacté par téléphone tous les médecins sélectionnés en utilisant les services d'une entreprise spécialisée externe. Cette mise en place téléphonique avait pour but de valider les identifiants de connexion sur le site et de familiariser les médecins observateurs avec l'interface en effectuant une première saisie, assistée en ligne avec un opérateur. Un réseau de 1 500 médecins a été opérationnel en moins de 3 semaines à partir du 2 décembre 2005. En moyenne sur cette saison, la participation a été de 799 médecins par semaine.

Le protocole prévoyait le recueil hebdomadaire des cas de 3 pathologies ORL de l'enfant : les rhinopharyngites, les otites moyennes aiguës (OMA) et les angines. Au cours de la période comprise entre le 12 décembre 2005 et le 26 mars 2006, 217 432 cas ont été enregistrés dont 150 889 cas de rhinopharyngites (69 %), 36 140 cas d'OMA (17 %), et 30 403 cas d'angines (14 %) ; 30 % étaient âgés de 6 mois à 2 ans ; 44 % de 2 à 8 ans et 26 % de 8 à 15 ans. Le sex-ratio était de 1,08. Pour cette saison, les taux d'incidence hebdomadaires moyens estimés de consultations en France métropolitaine (pour 100 enfants de moins de 15 ans) étaient de : 7,75 [IC95 % : 7,59 – 7,91] pour les rhinopharyngites, 1,86 [IC95 % : 1,79 – 1,94] pour les OMA et 1,53 [IC95 % : 1,46 – 1,61] pour les angines.

3.2 Saison 2006-2007

Pour la saison 2006-2007, la procédure de mise en place du réseau a été légèrement différente. Elle a été fondée sur le sous-ensemble de la base de donnée de 7 000 médecins ayant accepté d'être contacté via un envoi électronique. Cette procédure a permis de constituer un réseau échantillonné de la même manière que l'année précédente de 1000 observateurs (limité à ce nombre en raison de contrainte budgétaire), opérationnels en 1 semaine, sans mise en place téléphonique. La moitié des médecins observateurs pour cette nouvelle saison avait déjà participé à la saison précédente. En moyenne sur cette deuxième saison, la participation a été de 564 médecins par semaine.

En plus des 3 pathologies surveillées lors de la première saison, nous avons mis en place la surveillance de 2 pathologies supplémentaires : la sinusite aiguë et la laryngite aiguë striduleuse. Au cours de la période comprise entre le 11 décembre 2006 et le 25 mars 2007, 151 826 cas ont été enregistrés dont 101 870 cas de rhinopharyngites (67 %), 26 396 cas d'OMA (17 %), 23 560 cas d'angines (16 %). De plus nous avons enregistré, 12 726 laryngites et 6 701 cas de sinusites. Le nombre total de cas enregistrés sur la période est de

171 253. La distribution d'âge et de sexe était superposable à celle de la saison précédente. Pour cette saison, les taux d'incidence hebdomadaires moyens estimés de consultations en France métropolitaine (pour 100 enfants de moins de 15 ans) étaient de : 7,16 [IC95 % : 6,96 – 7,36] pour les rhinopharyngites, 1,81 [IC95 % : 1,71 – 1,91] pour les OMA, 1,67 [IC95 % : 1,57 – 1,76] pour les angines, 0,92 [IC95 % : 0,85 – 1,00] pour les laryngites et 0,49 [IC95 % : 0,43 – 0,54] pour les sinusites.

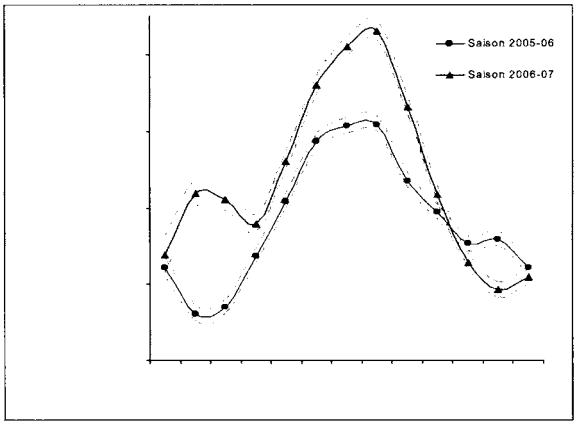


Figure 4 - Évolution temporelle du nombre moyen hebdomadaire estimé de consultations pour une affection ORL en France métropolitaine au cours des saisons 2005-06 et 2006-07.

La figure 4 illustre l'évolution temporelle comparée pour les 2 années de surveillance du nombre hebdomadaire estimé de consultations toutes pathologies ORL confondues. Chaque courbe est accompagnée de son intervalle de confiance à 95 %. Le nombre de cas estimé est plus important pour la saison 2006-07, en revanche, il atteint son maximum au cours de la 6e semaine de l'année pour les 2 saisons et l'évolution semble montrer une certaine similitude de dynamique. Des analyses plus précises sont effectuées pour chaque pathologie. Elles sont disponibles sur demande sur le site de l'observatoire aux adresses : http://www.hivernale.com ou http://www.khiobs.org.

4 Discussion

4.1 Limite de l'étude : contrôle qualité des données et représentativité du réseau

Les critères diagnostiques des affections étudiées figuraient au protocole. Toutefois, un biais de classification des affections déclarées a pu être introduit mais son amplitude n'a pu être estimée. De la même manière, il n'a pas été réalisé de contrôle de qualité des diagnostics posés par les médecins observateurs pour des raisons de complexité d'organisation d'un tel contrôle.

La question de la participation et de la représentativité du réseau est un élément majeur d'une telle approche. La littérature dans ce domaine documente largement la question de la représentativité des réseaux de surveillance [6 , 7]. En France, le réseau Sentinelle, étant donné sa notoriété et son antériorité, peut être considéré comme une référence dans le domaine de la surveillance en temps réel. Il surveille depuis près de 24 ans, 8 maladies transmissibles et notamment les syndromes grippaux qui surviennent en période hivernale.

Celui-ci fonctionne officiellement avec un réseau de 500 médecins généralistes. La participation hebdomadaire réelle pour des périodes équivalentes à celles que nous présentons est nettement inférieure à cet effectif. À titre de comparaison, l'observatoire Hivern@le-KhiObs dispose d'un réseau dont l'effectif se situait entre 1 000 et 1 500 médecins au cours des saisons présentées. En moyenne sur les 2 saisons, la participation hebdomadaire a été supérieure à 650 médecins par semaine.

Contrairement à d'autres observatoires dont la participation médicale est basée sur le bénévolat, Hivern@le-KhiObs a indemnisé ces observateurs à raison d'un montant maximum de 240 € par saison. Cette contrepartie équivaut approximativement au montant d'une consultation pour chaque semaine de surveillance. Il s'agit donc plus d'un dédommagement que d'une véritable rémunération. Dans la pratique, nous avons pu constater qu'une part importante des médecins interrogés sur ce point était sensible à cette compensation même si celle-ci est considérée comme modeste au regard de l'effort consenti en termes d'assiduité sur 12 semaines de surveillance. De ce fait, il est probable que par rapport à l'indemnisation pratiquée dans le cadre de cette étude, le bénévolat introduise un biais notablement plus important dans le recrutement et donc dans la représentativité de l'échantillon des médecins participants. Nous avons réalisé une étude spécifique (à paraître) relative à la question de l'échantillonnage et de la représentativité dans le cadre des réseaux de surveillance où nous montrons comment la taille et les caractéristiques de l'échantillon de médecins observateurs ont une répercussion notoire sur l'intervalle de confiance des estimations.

4.2 Confirmation de l'importance de ces pathologies

Au cours des deux saisons hivernales 2005-06 et 2006-07, l'observatoire Hivern@le-KhiObs a confirmé l'importance des consultations pour 5 pathologies ORL chez l'enfant. Il a permis d'estimer que le taux d'incidence hebdomadaire moyen de consultations pour ces pathologies en France métropolitaine (pour 100 enfants de moins de 15 ans) s'élevait à 11,85 [IC95 % : 11,63 – 12,08]. À titre de comparaison, ces valeurs sont plus de 15 fois supérieures à celles estimées pour les consultations de syndromes grippaux pour l'ensemble de la population en France métropolitaine et pour les mêmes périodes [8, 9]. Ainsi, à partir des données collectées sur cet observatoire, il a été possible d'estimer que le nombre moyen de consultations par saison pour ces pathologies ORL de l'enfant était de 18,6 millions ce qui correspond à une moyenne de 1,65 consultation par enfant au cours de chaque saison hivernale. Étant donné les coûts d'exploitation de l'observatoire, nous nous sommes limités à 12 semaines de recueil. Notre objectif de réaliser un observatoire en médecine de ville à l'échelon national pour disposer d'estimations fiables et précises des affections ORL de l'enfant est à notre connaissance le premier travail du genre. Il n'était pas prévu, a priori, de détecter d'éventuelles épidémies qui auraient nécessité une période de surveillance plus longue encadrant cette fenêtre.

4.3 Défis et enjeux pour la santé publique

Alors que ces pathologies peuvent être considérées comme relativement bénignes, l'intérêt d'une telle surveillance est multiple à la fois d'un point de vue individuel, économique, social, organisationnel et technologique.

Au niveau individuel, au-delà de la récurrence des épisodes pouvant déboucher sur d'éventuelles complications cliniques, elles sont source de traitements antibiotiques répétés connus pour leur implication dans l'émergence de résistances des germes [10-13].

D'un point de vue économique, l'observatoire donne une estimation précise sur le nombre de consultations pour ce type de pathologies qui révèlent l'ampleur des coûts représentés par ces affections. En première approximation en se fondant uniquement sur ces coûts directs [9, 14], nous avons estimé à partir du résultat de l'observatoire que la charge des affections ORL représenterait par saison plus d'un milliard d'euros.

Au niveau social, ces affections sont une source d'absentéisme scolaire et en crèche mais aussi d'arrêts de travail pour les parents afin d'assurer la garde des enfants ; une approche économique plus précise est en cours. L'intérêt porté à une telle surveillance est attesté par l'écho de cette activité, au 15 janvier 2009, 2 190 citations (recherche Google, mot clef « khiobs »).

Sur le plan organisationnel, il s'agit de mettre en place chaque saison un réseau de professionnels de santé afin que celui-ci soit opérationnel dans un laps de temps très court d'une semaine. L'animation du réseau doit être la moins coûteuse possible en temps/homme ce qui signifie une organisation de la saisie, du contrôle et de la restitution de l'information épidémiologique la plus automatisée possible.

Sur le plan technologique, cette plate-forme est adaptée pour répondre à de tels objectifs. De plus, l'entrepôt de données s'inscrit dans une architecture de type n-tier avec un retour d'information via un système d'information géographique en ligne. L'originalité de cette plate-forme réside dans sa souplesse d'adaptation à différents types de surveillances compte tenu de sa modularité.

5 Conclusion

Lors de son introduction dans le domaine de la santé publique voici une vingtaine d'année, la téléinformatique a montré son extraordinaire potentiel. L'utilisation d'outils pour la mise en place dans de brefs délais de systèmes permettant la collecte de données à distance, l'analyse en temps réel et la redistribution de l'information démontre leur efficacité en cas de situation de crise [15] mais aussi selon le groupe de travail international sur la notification des maladies transmissibles pour les maladies à incidence élevée et sévérité faible [1]. Ce dernier point s'applique à l'observatoire des pathologies ORL « Hivern@le-KhiObs ».

Jusqu'à la mise en place de l'observatoire national « Hivern@le-KhiObs » pour la surveillance épidémiologique des pathologies ORL chez l'enfant pendant la saison hivernale, aucun suivi de ces pathologies n'était effectué et aucune quantification précise à un niveau national n'était disponible. Étant donné les nombreuses contraintes liées à la mise en place d'un observatoire de ce type, cette étude a été l'occasion de démontrer la faisabilité du projet sur la base du générateur de la plate-forme de surveillance « MetaSurv ». Les résultats obtenus au cours de cette étude ont démontré les capacités d'une telle plate-forme qui jusqu'alors avait été utilisée à titre expérimental sur des projets de moindre envergure. Elle est désormais utilisée pour plusieurs projets d'études en santé et en recherche clinique.

L'expérience de la mise en place de l'observatoire Hivern@le-KhiObs souligne l'importance que revêt en santé publique la surveillance hivernale des infections ORL de l'enfant. Cet observatoire et sa plate-forme MetaSurv, dont l'objectif est de diffuser une information sur ces pathologies, contribuent naturellement à relever le défi de cet enjeu de santé publique.

Remerciements

Ce travail a été mené dans le cadre d'un partenariat entre l'Inserm et Sanofi-Aventis-France. Nous remercions le Dr Mourad Behar, toute l'équipe de Sanofi-Aventis-France et les 1 500 médecins qui ont participé au recueil dont la liste est consultable sur le site de l'observatoire.

Références

[1] Doherty JA. Final report and recommendations from the National Notifiable Diseases Working Group. *Can Commun Dis Rep* 2006; 32(19): 211-25.

[2] Franc C. *et al.* La prise en charge des enfants en médecine générale : une typologie des consultations et visites. Etudes et Résultats Dress 2007; 588.

[3] Toubiana L, Moreau S, Bonnard G. MetaSurv: Web-Platform Generator for the Monitoring of Health Indicators and Interactive Geographical Information System. *Stud Health Technol Inform* 2005; 116: 989-93.

[4] Toubiana L, *et al.* MetaSurv : Générateur de plates-formes de surveillance d'indicateurs de santé sur Internet et système d'information géographique interactif. In *JFIM2005* ; XI Journées Francophones Informatique Médicale. 2005. Lille - France.

[5] Flahault A, *et al.* Virtual surveillance of communicable diseases: a 20-year experience in France. *Stat Methods Med Res* 2006; 15(5): 413-21.

[6] Chretien JP, *et al.* Syndromic surveillance: Adapting innovations to developing settings. *Plos Medicine* 2008; 5(3): 367-372.

[7] Choi BCK, *et al.* Enhancing global capacity in the surveillance, prevention, and control of chronic diseases: seven themes to consider and build upon. *Journal of Epidemiology and Community Health* 2008; 62(5): 391-397.

[8] Vaux S, *et al.* Epidemiological and virological influenza surveillance in France: season 2005-2006. *BEH* 2006; 51-52: 403-406.

[9] Devaux MGN, SC. Déremboursement des mucolytiques et des expectorants : quel impact sur la prescription des généralistes ? Questions d'économie de la santé IRDES, 2007; 128.

[10] Coenen S, *et al.* European Surveillance of Antimicrobial Consumption (ESAC): outpatient antibiotic use in Europe. *J. Antimicrob. Chemoth* 2006; 58: 401-407.

[11] Fleming DM, *et al.* Respiratory illness associated with influenza and respiratory syncytial virus infection. *Archives of Disease in Childhood* 2005; 90(7): 741-746.

[12] Goossens HFM, Vander Stichele R. Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. *Lancet* 2005; 365: 579-87.

[13] Magne D, et al., Antibiotic consumption at the Assistance publique-Hopitaux de Paris (AP-HP): trends over 15 years, France, 1990-2004. *BEH,* 2007. **5**: p. 40-42.

[14] Kobari, T., *et al.* Factor analysis of drug supply time series at pharmacies. *Journal of Health Science* 2008; 54(1): 107-111.

[15] Garshnek V, *et al.* Telecommunications systems in support of disaster medicine: applications of basic information pathways. *Ann Emerg Med* 1999; 34(2): 213-8.

Adresse de correspondance

Laurent Toubiana
Service de Biostatistique et d'Informatique Médicale
Groupe hospitalier Necker-Enfants Malades
161, rue de Sèvres 75743 Paris cedex 15
E-mail: Laurent.toubiana@inserm.fr

Informatique et Santé, 2009 (17) : 313-324
© Springer – Verlag France

Méthodes et évolution de la certification de sites de santé sur le Web à travers l'expérience du HONcode

Arnaud Gaudinat[1], Celia Boyer, Vincent Baujard, Antoine Geissbühler[1,2]

[1]*Fondation Health on the Net, Genève, Suisse*
[2]*Service d'Informatique Médicale, Hôpital universitaire de Genève, Suisse*

Abstract

Today, the Web is a media with increasing pervasiveness around the world. Its use is constantly growing and the medical field is no exception. According to the Pew Internet & American Life Project, 80% of Internet users have searched for American medical information in 2006. However, the problem is no longer about finding information but assessing the credibility of the publishers as well as the relevance and accuracy of the documents retrieved from the web. This problem is particularly relevant in the medical area which has a direct impact on the well-being of citizens and in the Web 2.0 context where information publishing is easier than ever. To address the quality of the medical Internet, the HONcode certification proposed by the Health On the Net Foundation (HON) is certainly the most successful initiative. The aims of this paper are to present certification activity through the HONcode experience and to show that certification is more complex than a simple code of conduct. Therefore, we first present the HONcode, its application and its current evolutions. Following that, we give some quantitative results and describe how the final user can access the certified information.

Keywords

Certification; Natural Language Processing; Internet; Information Storage and Retrieval; Public Health Informatics.

1 Introduction

Internet est devenu un des media de communication le plus largement utilisé. Avec la démocratisation de la disponibilité de logiciel et de serveur Web, tout un chacun peut facilement publier tout type d'information qui sera alors disponible et accessible sur la planète entière. Ces dernières années, cette facilité c'est encore accrue avec l'avènement du phénomène Web 2.0. Ainsi malgré la profusion de contenu disponible, la question n'est plus seulement de trouver l'information mais surtout de savoir si les rédacteurs d'information sont crédibles. Le problème est particulièrement critique en ce qui concerne le domaine médical où l'information trouvée (explication des maladies, traitements recommandés, type de médecine...) a un impact direct sur la santé du citoyen. Comme le démontre la dernière étude de *Pew internet & American Life project* [1] conduite en 2007 où les patients ayant une maladie chronique sont plus à même d'utiliser Internet pour prendre leur décision sur leur traitement (E-patients invalides ou souffrant de maladie chronique), 53 % des personnes qui cherchent de l'information santé déclarent que l'information trouvée a un impact sur leur propre soin de santé ou la façon dont il l'utilise pour les autres [2]. En fait dans de nombreux cas, les sites de santé sur le Web ne

fournissent pas d'information idoine concernant la conception des études scientifiques ou même de références à des études pour confirmer leurs allégations médicales.

Pour répondre à l'insuffisance de clarté, de nombreuses initiatives pratiques et théoriques ont jalonné la courte histoire du Web [3]. L'histoire de la qualité de l'information sur le Web est intimement liée à la recherche d'information. En effet une réponse pertinente du point de vue sémantique n'a d'intérêt que si elle est « juste » (donc crédible, de qualité et adaptée). Les initiatives suivantes (non exhaustives) représentent les méthodes et les courants les plus significatifs qui ont été appliqués au Web concernant la qualité de l'information (médicale ou non) :

- ***La sélection*** (ou le référencement explicite) de pages dans des répertoires est certainement la première initiative qui a essayé de mettre en exergue les meilleures pages du Web (Yahoo). Dans le domaine médical, l'américain MedlinePlus[4] et le français CISMeF [5]sont des références incontournables. En plus des répertoires, cette sélection peut être aussi mise à disposition dans des moteurs de recherche comme dans MedHunt [6] par exemple.
- ***L'auto-régulation*** des sites Web, pour les producteurs de contenus, consiste en leur propre certification en respectant une liste de critères de qualités bien établis [7].
- ***La popularité*** des pages Web, liée à la structure hypertextuelle du Web, peut être un bon indicateur de qualité [8, 9]. La popularité est en fait corrélée avec le jugement implicite des producteurs d'information lorsqu'ils décident de mettre un lien vers un autre site Web. Google et son fameux PageRank [10] est certainement l'initiative la plus représentative.
- ***La certification*** de pages Web a pour but la promotion de la qualité de l'information. Les pages d'un site Web doivent se soumettre à l'évaluation d'une organisation tierce (a contrario de l'auto-régulation), tel que la Fondation HON avec sa charte le HONcode [11] et l'URAC [12], organisation États-Unienne. La certification est corrélative au respect d'une charte éthique et de qualité clairement définie.
- ***L'éducation*** des utilisateurs d'Internet consiste à les aider en leur donnant des règles de bonne pratique pour mieux sélectionner les sites visités [13-15].
- ***L'initiative collaborative, sociale ou encore nommée communautaire*** consiste en l'utilisation d'une grande quantité de jugements d'utilisateur pour mieux caractériser le Web (Google Co-op [16]) et/ou en créant des réseaux sociaux de confiance, sur le principe « comme je te fais confiance, je fais confiance aux personnes à qui tu fais confiance » (Lijit [17]).

L'objectif de la fondation HON est de donner accès à de l'information médicale de qualité sur le Web pour le citoyen qu'il soit patient ou professionnel de santé. Pour aider à l'éducation des responsables de sites Web de santé et des cyber-citoyens, la fondation Health On the Net a établi une charte de déontologie en huit points, le HONcode. La création de cette charte a été motivée par les responsables de sites réclamant des directives que tous pourraient respecter et qui garantiraient un comportement éthique et la qualité de l'information de santé présente sur ses sites. Ainsi, depuis 1995, HON s'est engagé dans la voie de la certification avec succès qui a mené son code de conduite, le HONcode, à devenir un standard de facto.

Simple à première vue, la certification HONcode s'articule sur 4 éléments fondamentaux et indispensables. En effet, l'expérience nous a montré que 1/ ***la charte éthique*** bien qu'essentielle n'était qu'un des éléments visibles de la certification, d'autres éléments du

HONcode sont tout aussi critiques. Ainsi 2/ ***les procédures d'évaluation selon le HONcode*** sont capitales car elles reflètent la mise en œuvre pratique et équitable du HONcode dans le travail quotidien des collaborateurs de la Fondation. Ensuite, 3/ ***La tâche de certification*** apparaît comme l'atout essentiel de la certification, en effet d'autres codes de santé n'ont pas fait long feu car peu de personnes étaient volontaires pour les appliquer. Et finalement, le dernier élément souvent sous-estimé est 4/ ***les moyens mis en œuvre pour donner accès à cette information*** sur Internet.

Dans cet article nous présenterons le HONcode, et sa méthodologie d'implémentation. Ensuite nous indiquerons qu'elles sont les évolutions actuelles du HONcode. Dans un second temps, nous exposerons quelques résultats quantitatifs ainsi que les initiatives réalisées dans le cadre de l'accès aux résultats de la certification HONcode.

2 Méthode et évolution du HONcode

2.1 La certification

Le HONcode n'est pas le seul système basé sur l'évaluation de critères, ainsi d'autres initiatives importantes ont été proposées. Cependant leurs applications concrètes (auto-régulation/tierce-partie-certification, simple/complexe, accessibilité du résultat, gratuite/ payante...) peuvent être très différentes et ont déterminé leur pérennité.
Il est possible de citer les initiatives les plus marquantes avec leurs caractéristiques les plus saillants (voir aussi [18]) : Omni en 1995, HONcode (tierce-partie-certification, 8 principes vérifiés annuellement, multilingue) en 1996, Discern (auto-régulation) en 1998, WebMedica en 1998 (tierce-partie-certification, espagnol), MedCertain (précis mais compliqué) et Hi-ethics (auto-régulation) en 2000, eHealth Code of Ethics (auto-validation) en 2001, URAC en 2001 (détaillé mais aussi couteux), European Guideline en 2002 (eq. aux principes du HONcode dont HON a participé à l'élaboration) et Afgis en 2003 (dédié aux sites allemands). Alors que certaines initiatives ont disparu ou que d'autres sont en forte perte de candidats, le HONcode est quant à lui traduit dans 34 langues, dispose de plus 6 500 sites certifiés en 2009, est présent dans 118 pays, et à été sélectionné fin 2007 par la France comme organisme de certification officiel des sites de santé Français [19]. Sur le Web, la certification, proposée par HON, n'est pas qu'un « simple » instrument de notation de qualité, elle a aussi pour effet de sensibiliser l'esprit critique de l'utilisateur ainsi que le producteur de l'information en leur donnant une méthode simple à mettre en œuvre dans le travail au quotidien.

2.2 Code éthique de bonne conduite HONcode

Le HONcode [20] est un ensemble de huit principes éthiques définis par la fondation HON dans le but de proposer un standard de transparence des pratiques opérationnelles et visant à améliorer la qualité de l'information de santé sur le Web. Ces principes sont appliqués durant une évaluation des sites Web qui demandent la certification HONcode. Chaque principe existe en 2 versions, une courte (tableau 1) pour l'internaute et une plus longue pour le professionnel de santé et/ou responsable de sites.

La description longue des principes est plus explicite et enlève toute ambiguïté quand aux possibles mauvaises interprétations. Par exemple, la déclaration complète du premier principe d'autorité est : *Tout avis médical fourni sur le site sera donné uniquement par du personnel spécialisé (diplômé) du domaine médical et des professionnels qualifiés, à moins qu'une déclaration explicite ne précise que certains avis proviennent de personnes ou d'organisations non médicales.* Il indique que l'information médicale devra être fournie par

des personnes formées dans le domaine médical. Les principes du HONcode peuvent être consultés en ligne dans 34 langues [20]. De plus, la fondation HON fournit un règlement du processus de certification.

Tableau 1 - Version courte des principes du HONcode.

Principes	Version courte
1. Autorité	Indiquer la qualification des rédacteurs.
2. Complémentarité	Compléter et non remplacer la relation patient-médecin.
3. Confidentialité	Préserver la confidentialité des informations personnelles soumises par les visiteurs du site.
4. Attribution et datation	Citer la/les source(s) des informations publiées et dater les pages de santé.
5. Justification	Justifier toute affirmation sur les bienfaits ou les inconvénients de produits ou traitements.
6. Professionnalisme	Rendre l'information la plus accessible possible, identifier le webmestre, et fournir une adresse de contact.
7. Transparence du financement	Présenter les sources de financements.
8. Honnêteté dans la publicité et la politique éditoriale	Séparer la politique publicitaire de la politique éditoriale.

2.3 Processus de certification

La certification étant un acte volontaire de la part des requérants, la première étape consiste en la soumission du formulaire de demande par Internet sur le site de HON (Le volontarisme permet une plus grande implication du demandeur). Une pré-évaluation est proposée au webmestre dans le but d'identifier les principes manquant de façon autonome. Ensuite une fois la requête de certification soumise, des experts humains de la Fondation évaluent les sites Web qui en ont fait la demande. Chaque information éthique manquante est indiquée et doit être ajoutée ou justifiée au contenu des pages Web du site. Une fois les modifications réalisées, un sceau de certification et un identificateur unique sont délivrés. Tous les sites HONcode sont certifiés pour une période de 1 an et donc révisés annuellement et ceci de façon systématique. Si un site Web ne respecte plus les principes HONcode, le webmestre reçoit un avertissement et, si les modifications requises ne sont pas effectuées, le site peut perdre sa certification. Comme on peut le voir, le processus de certification est un processus interactif qui permet d'établir un contact constructif entre les réviseurs HONcode et le webmestre. En effet le but n'est pas seulement de séparer le bon grain de l'ivraie mais aussi d'aider au maximum à amener les sites à un certain niveau de qualité et de transparence.

2.4 Évolution du HONcode au sein du Web 2.0

Malgré l'utilisation des forums avant l'avènement du Web, ce n'est que récemment que HON a été confronté à une demande croissante de certification de sites collaboratifs dont la nature même interroge sur l'efficacité du HONcode tel qu'il a été originellement défini. En effet que devient le principe d'autorité, dans le contexte du Web 2.0, lorsque tout un chacun peut contribuer à l'ajout de contenu sur un site.

Tableau 2 - Spécificité Web 2.0 du HONcode.

Principes	Règlement d'application aux principes
1. Autorité	Il doit être clairement établi si la plate-forme est modérée ou non (plate-forme modérée par réaction soit aux alertes des lecteurs ou par vérification aléatoire des messages). Les qualifications des modérateurs ainsi que leur taux de fréquentation de la plate-forme doivent figurer. Les noms des modérateurs ne sont pas obligatoires, ils peuvent utiliser des pseudonymes. Il doit être stipulé que par défaut, tous les utilisateurs de la plate-forme ne doivent pas être considérés comme des professionnels de santé. S'il en est autrement, cela doit être précisé et la spécialité médicale doit être indiquée. Il doit être clairement stipulé que tous les utilisateurs de la plate-forme ainsi que les modérateurs doivent se comporter en tout temps avec respect et honnêteté. Les fonctions et les responsabilités des modérateurs doivent être clairement établies (le fait de pouvoir effacer un contenu jugé non approprié, par ex.). La manière d'effacer un message, de bannir (à discrétion, avec ou sans explication immédiate, mise en garde, notification) un utilisateur de la plate-forme doit également être clairement établie.
2. Complémentarité	Le principe de complémentarité doit être rappelé sur la page d'accueil de la plate-forme.
3. Confidentialité	La plate-forme doit avoir une politique de confidentialité. Le fait que tout message peut être lu et utilisé par tous doit être souligné ainsi que le fait que tout message peut être repris et cité dans d'autres messages. La possibilité ou non d'effacer ou modifier ses propres messages en temps qu'utilisateur de la plate-forme ainsi que la possibilité d'effacer des citations dans d'autres messages doit être établie.
4. Attribution et datation	Une phrase demandant aux utilisateurs de la plate-forme d'indiquer les sources (références, liens, etc.) des données médicales diffusées si elles ne proviennent pas de l'expérience personnelle doit figurer. Tous les messages ou contributions des utilisateurs doivent être automatiquement datés.
5. Justification	Toute affirmation relative au bénéfice ou à la performance d'un traitement donné, d'un produit ou d'un service commercial est considérée comme une allégation (...).
6. Professionnalisme	Aucun ajout pour ce principe.
7. Transparence du financement	Il doit être stipulé si les modérateurs sont bénévoles ou pas.
8. Honnêteté dans la publicité et la politique éditoriale	Il doit être notamment établi si les utilisateurs de la plate-forme ont le droit de publier de la publicité, sous forme de contenu, bannière ou liens par exemples.

Par ailleurs un des avantages du HONcode est sa simplicité puisqu'il tient en huit principes. Ainsi, afin de répondre à l'évolution continue du Web, certains ajouts ont été effectués dans le règlement d'application des principes du HONcode pour répondre aux particularités du Web 2.0 tout en gardant la même ligne de conduite. Un site collaboratif désirant être certifié doit respecter les directives ou règlements d'application du HONcode comme présenté dans le tableau 2. Ainsi le HONcode montre qu'il est totalement adapté à

l'évolution du Web et utilisable dans le contexte du Web 2.0. De plus il montre qu'il n'est pas figé et que les directives d'implémentations évoluent au fur et à mesure de l'évolution du Web.

2.5 La certification des sites français

La France a été le pays précurseur dans le domaine de la eSanté en votant une loi en 2004 concernant la certification des sites informatiques dédiés à la santé. La Haute Autorité de Santé a été mandatée pour mettre en œuvre cette loi. Après une période d'évaluation des différentes alternatives existant dans le monde, le HONcode a été choisi comme modèle de certification. Cependant ce n'est pas uniquement la certification HONcode qui a été approuvée. Les procédures d'application de la Fondation HON, son équipe ainsi que son infrastructure informatique ont participé à son choix. Enfin, suite à l'évaluation annuelle par des qualiticiens indépendants de la démarche qualité au sein de la Fondation, HON a été officiellement accréditée par la HAS pour la certification des sites français. Cette accréditation, renouvelée chaque année, a permis d'améliorer la pratique quotidienne de la certification afin de garantir un meilleur service. Cette démarche qualité garantit notamment une formation adéquate des réviseurs HONcode, une équité dans le traitement des demandes de certification, un service respectant des délais, une infrastructure fiable et la mise en place d'indicateurs pour de futures améliorations. Cette évolution fondamentale met bien en exergue l'idée que la certification n'est pas simplement qu'un code de bonne conduite et que la mise en œuvre concrète de ce dernier est tout aussi importante.

2.6 La détection automatique des principes

Au vue de la dynamique du Web, la certification peut s'apparenter de plus en plus à la tâche des Danaïdes, condamnées à remplir sans fin un tonneau percé, qui ici serait de plus en plus grand. Ainsi, récemment, des initiatives basées sur des algorithmes de reconnaissances de critères basés sur des règles (originellement proposé par Price [21]) ou par apprentissage ont été présentées afin de donner automatiquement des indices de qualité et d'éthique aux pages de santé du Web. Alors que le modèle d'apprentissage supervisé de Aphinyanaphongs [22] est basé sur des exemples statiques de bonnes et mauvaises pages et donc dépendant du domaine traité, celui de HON [23] est quant à lui basé sur un modèle plus générique : le HONcode. Actuellement le système de détection de HON obtient une micro-précision de 0,78 et 0,73 de F-mesure avec 0,06 d'erreur en moyenne sur tous les principes. Des principes tels que la confidentialité sont très facilement reconnus alors que d'autres comme les justifications sont difficiles à être identifiés automatiquement. Dans notre cas, les informations d'apprentissage proviennent des déclarations des principes identifiées par les réviseurs durant le processus de certification. Il a aussi été montré que l'utilisation des URLs permettait d'améliorer la détection des critères de qualité sur les sites Web [24] (e.g. www.nlm.nih.gov/medlineplus/privacy.html). Ce genre d'approche en plus d'être intéressante pour assister les réviseurs peut s'avérer utile pour caractériser automatiquement les sites Web qui ne demandent pas forcément la certification.

3 Résultat de la certification et son accès

La certification est pertinente à condition que les utilisateurs d'Internet puissent y avoir accès. Cette section présente quelques résultats quantitatifs ainsi que les moyens d'accès mis en place par HON pour que l'utilisateur final puisse avoir accès aux informations certifiées de la manière la plus naturelle possible (c'est-à-dire dans son usage quotidien).

3.1 Base de données de sites certifiés HONcode

Actuellement, la base de données comporte plus de 6 500 sites Web certifiés HONcode à travers 118 pays qui représentent plus de 1 million de pages indexées dans Google. 52 % des sites certifiés sont en anglais et environ 11 % en français, suivis par de nombreux sites en espagnol, italien, allemand et portugais. Pour chaque site évalué sont collectées les informations suivantes : 1) respects des principes HONcode ; 2) extrait de texte correspondant, aux 8 principes ; 3) URLs des principes sur le site ; 4) mots clés du site sélectionnés parmi les termes du MeSH [25] ; et 5) étiquette sémantique général (*i.e.*, santé pour les femmes/hommes, information pour les patients/professionnels de santé).

Tableau 3 - Nombres de demande d'inscriptions au HONcode de 2005 – 2008.

Année	Nombre d'inscriptions en France	Nombre d'inscriptions au total
2005	57	895
2006	42	898
2007	217	1142
2008	442	1409

Le tableau 3 indique le nombre de demandes d'inscriptions pour la période allant de 2005 à 2008. La première colonne donne cette information pour les sites français alors que la seconde colonne est réservée à la totalité des demandes d'inscription (y compris les sites français).

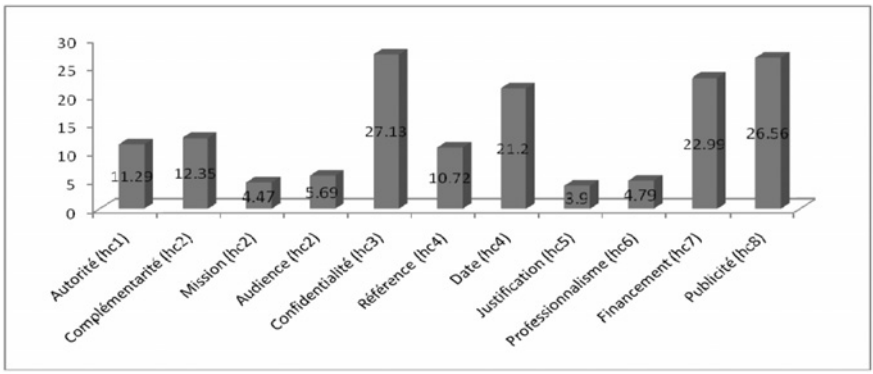


Figure 1 - Pourcentage des principes manquants pour les 1231 sites ayant demandé la certification du 1 janvier 2007 au 25 aout 2008.

Ainsi, dans les deux cas, on constate une croissance dans le nombre de demande ces 2 dernières années. La croissance est bien entendue beaucoup plus forte pour les sites français et cela est notamment le résultat de l'accréditation de HON en 2007 par la HAS qui a favorisé les nouvelles demandes pour ces sites. À noter aussi qu'en 2008, près d'1 site sur 3 qui fait la demande est un site français alors qu'en 2005 on avait seulement 1 site français sur 16. La figure 1 présente le pourcentage de principes manquants sur les 1 231 sites ayant fait une demande de certification pendant la période du 1er janvier 2007 au 25 aout 2008.

Une première lecture montre qu'il manque souvent des principes et que certains principes tels que la confidentialité, la date, le financement et la publicité sont fréquemment insuffisamment décrits sur le site. Cependant, de l'avis de l'équipe HONcode, cela n'indique pas forcément que le site ne respecte pas la confidentialité de ses utilisateurs mais qu'il n'a pas forcément jugé utile de faire une déclaration explicite. En revanche, un principe tel que la date est plus objectif car il met en exergue un certain manque de rigueur et de transparence.

3.2 Les sceaux dynamiques

Au début en 1996, un simple logo permettait aux internautes d'identifier un site certifié d'un site non certifié. Cependant, le sceau HONcode est très vite devenu un moyen de garantie supplémentaire pour l'internaute en obligeant les sites à lier le sceau HONcode de leur site vers le site HON et son certificat unique. L'idée étant de limiter toute utilisation abusive du logo HONcode, puisque la vérification ultime se fait sur le site de HON. À noter que ce genre d'abus est très rare et plutôt le fait de négligence que d'acte volontaire.

Figure 2 - Sceau HONcode dynamique suivant l'état courant du processus de certification.

Cependant pour garantir une réactivité plus grande entre l'utilisateur, le webmestre et HON, les sceaux datés et dynamiques ont été proposés en 2006. Le principe de base est que HON garde au final le contrôle de l'affichage du HONcode (l'image unique générée pour un site donné est ainsi aussi hébergée chez HON). La figure 2 présente les différents états que peut afficher le sceau HONcode suivant son statut actuel, de gauche à droite, 1) le site est certifié avec la date de dernière certification indiquée, 2) le site est en court de réévaluation lors de l'inspection annuelle ou lors de contrôle inopiné et 3) enfin, la certification du site est retirée lorsque le site ne veut plus respecter le HONcode.

3.3 Google pour promouvoir la qualité de l'information

Google étant le moteur de recherche le plus utilisé par l'Internaute, il devient l'outil par excellence pour la promotion et la prise de conscience de la qualité de l'information médicale sur Internet. C'est en 2006 que Google a lancé sa plate-forme collaborative Google Co-op. En collaboration avec Google, HON a mis au point un ensemble de 60 labels permettant de mieux spécifier l'information de santé suivant les besoins de l'utilisateur. Un des principes du Google Co-op est de proposer 3 « topics » dont celui de la santé, où les principaux acteurs du domaine peuvent partager leurs informations afin de mieux qualifier les sites ou les pages. Ensuite, Google propose dans sa version par défaut (google.com seulement) des reformulations de requêtes plus pertinentes pour le domaine de la santé à condition que l'utilisateur soumette une requête liée à ce domaine. Concrètement dans notre cas, la procédure consiste à proposer à Google des annotations de pages Web par rapport à des labels prédéterminés tels que « traitement », type d'audience etc. Ainsi les dernières informations soumises à Google par HON étaient constituées de 83 243 annotations pour un total de 49 305 URLs (ou « patterns »).

La figure 3 illustre cette fonctionnalité pour la requête « asthma » dans la version originale de Google.com (donc seulement en anglais). L'utilisateur se voit dans un premier temps proposé une liste de labels pour mieux spécifier sa requête (Treatment, Symptoms…).

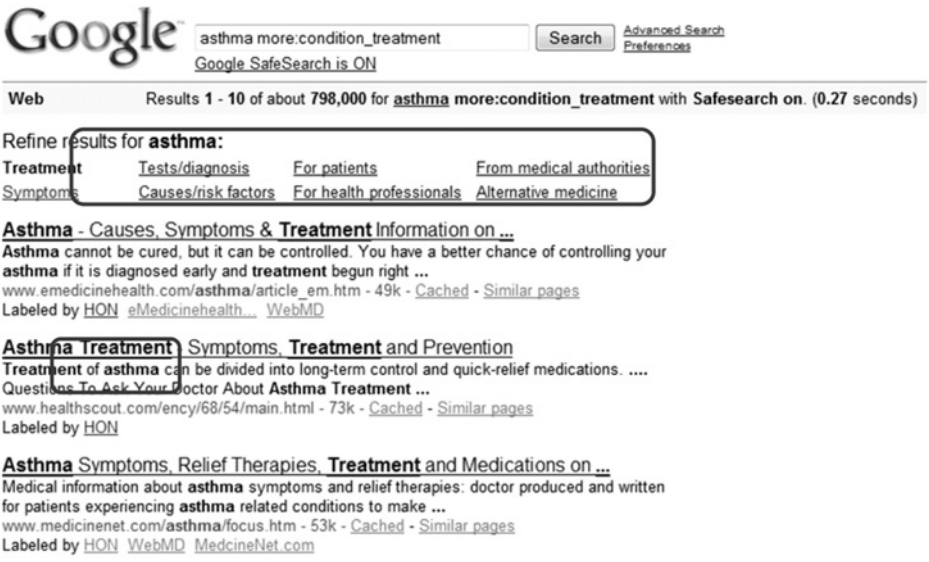


Figure 3 - Google Co-op pour la requête « asthma ».

Toujours dans la figure 3, l'utilisateur a sélectionné le label « Treatment », et a donc obtenu les résultats en rapport avec le traitement de l'asthme. Ici, les organisations ayant fourni des labels sont mentionnées pour chaque résultat, dont HON pour les trois premiers résultats. Ce système bien qu'intéressant pour mieux préciser sa requête reste insuffisant pour identifier la totalité des sites de qualité, car les sites HONcode qui n'abordent pas les sujets des labels Google n'y apparaissent pas. De plus cette annotation et sélection n'influent aucunement sur le classement des résultats de Google.

3.4 La barre d'outils HON

La barre d'outil d'HON [26] est certainement le moyen le plus intégré concernant l'accès aux sites certifiés HON. Elle est composée de 3 fonctionnalités qui sont 1/ **l'identification de l'appartenance au HONcode** du site courant, en temps réel, lors de la navigation. 2/ L'outil de recherche, **HONcodeHunt** exclusivement dédié aux sites certifiés HONcode et accessible depuis la barre de recherche du navigateur. 3/ La **mise en exergue des liens vers des sites certifiés** dans des outils de recherches populaires tel que Google et Yahoo ou des plates-formes de santé incontournables tel que MedlinePlus, ou encore des sites Web2.0 populaire tel que Wikipedia.

La figure 4 illustre la fonctionnalité d'identification des sites certifiés pour la requête « cancer du poumon » dans Google. Ici, 2 sites certifiés ont été trouvés dans les 3 premiers résultats (ils sont identifiés par le logo HONcode à gauche du résultat). L'utilisateur peut donc facilement d'un coup d'œil avoir une information sur la qualité des résultats en plus du classement usuel de Google.

Figure 4 : La barre d'outils de HON en action dans Google

À noter que le moteur de recherche HONcodeHunt intégré dans la barre d'outils est aussi accessible directement sur le site de HON. Dans les deux cas, il permet de rechercher uniquement dans les sites certifiés HONcode et de pouvoir mieux préciser sa recherche par rapport au profil recherché (homme/femme, professionnel/patient, nouveau né/enfant/adulte/senior…). Cette recherche spécifique est rendue possible par le travail des réviseurs de HON qui identifient, par rapport aux types d'audience, les pages, les sections ou les sites lors de l'inspection sites durant le processus de certification.

4 Conclusion et perspectives

L'objectif de ce papier était de montrer les multiples facettes du HONcode à travers son histoire, ses évolutions, son implémentation et son utilisation. Durant ces 12 dernières années, la fondation HON s'est efforcée de promouvoir la qualité de l'information médicale sur le Web à une échelle mondiale. Pour répondre aux exigences quantitatives du Web, l'expertise humaine (nécessaire) est assistée par de nombreux systèmes automatiques permettant une évaluation systématique et plus rapide des sites Web. L'internaute peut avoir accès à l'information de santé certifiée grâce à l'utilisation de barre d'outils HON intégrée dans son navigateur Web.

Toutes les méthodes et outils développés par la fondation HON peuvent être aussi utilisés à d'autres domaines, où la qualité et la transparence sont des composants essentiels (e.g. sécurité du Web pour les enfants).

Il est très important de multiplier les canaux de diffusion pour atteindre le plus d'utilisateurs possibles. Ainsi la réalisation de collaboration, pour partager nos informations, notre philosophie et notre vision, avec des acteurs majeurs tels que la

National Library of Medicine (États-Unis) ou Google sont indispensables pour le plus grand bénéfice de l'internaute.

L'approche conduite par la fondation HON est globale et couvre plus de 34 langues de par le monde. Cependant, HON doit répondre aux besoins locaux, à la variété des langues, aux différences culturelles et aux différentes réglementations. La création de représentations locales dans différentes parties du monde, telles qu'initiées en Afrique du Nord, en Afrique du Sud ou en Espagne, devrait nous permettre de penser localement et d'agir globalement pour améliorer la qualité de l'information médicale sur Internet.

La France a été précurseur dans le domaine de la qualité eSanté en légiférant sur la problématique de la qualité des sites santé. Une approche similaire dans les autres pays européen serait bienvenue pour continuer à promouvoir la qualité de l'information médicale sur Internet au bénéfice des utilisateurs finaux d'Internet.

Remerciements

L'activité de la fondation Health On the Net est soutenue par l'État de Genève et plus particulièrement par le département de l'économie et de la santé de Genève (DES), par la commission européenne à travers les projets de recherche WRAPIN (IST-2001-33260), ACTIVEHealth (IST-2001-37857) et PIPS (IST-2002-507019) ainsi que par la collaboration avec la Haute Autorité de Santé Française.

Références

[1] Fox S. *E*-patients with a Disability or Chronic Disease. *Pew Internet & American Life Project* 2007.

[2] Fox S. Online Health Search 2006, Online Health Search 2006. *Washington, DC: Pew Internet & American Life Project*, October 29, 2006.

[3] Wilson P. "How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the Internet", *British Medical Journal* 2002; 324 (7337): 598-602.

[4] Miller N, Lacroix EM, Backus JE. MedlinePlus: building and maintaining the National Library of Medicine consumer health Web service. *Bull Med Libr Assoc* 2000 ; 88 (1): 11–7.

[5] Darmoni SJ, Leroy JP, Baudic F, Douyere M, Piot J, Thirion B. CISMeF: cataloque and index of French speaking health resources. *Stud Health Technol Inform* 1999 ; 493-6.

[6] Boyer C, Baujard O, Baujard V, Appel R. New developments of MedHunt medical search engine. *In : MedNet 98*, Brighton, UK, 1998.

[7] Sheldon T. Trust mark launched as a guarantee of safety in the Netherlands. *BMJ* 2002; 324(7337): 567.

[8] Amento B, Terveen L, Hill W. Does"Authority" Mean Quality? Predicting Expert Quality Ratings of Web Documents. In: *Proceedings of the Twenty-Third Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2000.

[9] Borges H, Cervi M, Alvarez de Arcaya PT, Guardado G, Rabaza R, Sosa J. Rate of compliance with the HON code of conduct versus number of inbound links as quality markers of pediatric web sites. In: *Proceedings of the Sixth World Congress on the Internet in Medicine*, Udine, Italy, 29 November-2 December 2001.

[10] Page L, Brin S, Motwani R, Winograd T. The pagerank citation ranking: Bringing order to the web. *Technical report, Stanford Digital Library Technologies Project*, 1998.

[11] Boyer C, Baujard V, Scherrer J. HONcode: a standard to improve the quality of medical/health information on the internet and HON's 5th survey on the use of internet for medical and health purposes. In *6th Internet World* Congress for Biomedical Sciences (INABIS 2000), 1999.

[12] URAC: http://www.urac.org/MMandQualityChasm.asp, Nov 2008.

[13] HON - Conseils aux patients pour l'utilisation d'Internet en toute confiance : http://www.hon.ch/visitor_safeUse_f.html.

[14] Charnock D. *The DISCERN Handbook.* 1998. Radcliffe Medical Press.

[15] OMNI: omni.ac.uk, Nov 2008.

[16] Google Co-op: http://www.google.com/coop, Nov 2008.

[17] Lijit: http://www.lijit.com, Nov 2008.

[18] Risk A. Dzenowagis J. Review of Internet information quality initiatives. *Journal of Medical Internet Research* 2001; 3(4): e28.

[19] HAS, La certification des sites Internet santé: http://www.has-sante.fr/portail/jcms/c_334538/la-certification-des-sites-internet -sante, nov 2008.

[20] HONcode Guidelines: http://www.hon.ch/HONcode/Guidelines/guidelines.html, Nov 2008.

[21] Price S.L., Hersh W.R. Filtering web pages for quality indicators: an empirical approach to finding high quality consumer health information on the world wide web. *Proc AMIA Symp* 1999 ; 911-5.

[22] Aphinyanaphongs Y, Aliferis C. Text categorization models for identifying unproven cancer treatments on the web. *Stud Health Technol Inform*, 2007.

[23] Gaudinat A., Grabar N., Boyer C.: Machine learning approach for automatic quality criteria detection of health web pages. *Medinfo* 2007; 129: 705-9.

[24] Gaudinat A, Grabar N, Boyer C. Automatic retrieval of web pages with standards of ethics and trustworthiness within a medical portal: What a name page tells us - *11th Conference on Artificial Intelligence in Medicine (AIME 07)* - 07-11 July 2007 Amsterdam, The Netherlands.

[25] National Library of Medicine, Bethesda, Maryland. Medical Subject Headings, 2001. http://www.nlm.nih.gov/mesh/meshhome.html.

[26] HON Toolbar: http://www.hon.ch/HONcode/Plugin/Plugins.html, nov 2008.

Adresse de correspondance

Arnaud Gaudinat,

Health on the Net Foundation, HUG/SIM

24, rue Micheli-du-Crest,

1211 Genève 14 Suisse

Index des auteurs

A

Abdoune H., 79 ,
Alexandre H., 299
Avillach P., 251
Aymé S., 277

B

Bagayoko CO., 163 , 287
Bandon D.,15
Barlesi F., 203
Baujard V., 313
Baujat G., 277
Bayat S., 125 , 239
Ben Said M., 277
Bodemer C., 277
Bouaud J., 191
Bourdon-Lanoy E., 277
Bousquet C., 91 , 213
Boussadi A., 213
Boyer C., 287 , 313
Brizon A., 191
Buemi A., 57

C

Cauvin JM., 263
Chinot O., 203
Chmielik J., 139
Clarisse T., 299
Clavel L., 91
Colin F., 299
Cuggia M., 125 , 239
Culty T., 191

D

Dahamna B., 47, 57
Dameron O., 23
Daniel C., 227
Darmoni SJ., 47 , 57 , 79 , 101
Degoulet P., 3, 213
Delahousse J., 47
Despond M., 185
Dirieh Dibad AD., 101

Dufour E., 277
Dufour JC., 203 , 251
Duvauferrier R., 125

E

El Fadly N., 227

F

Fabry P., 69
Fieschi D., 203
Fieschi M., 47 , 203 , 251
Frimat L., 239

G

Garreau PJ., 263
Gasperina L., 91
Gaudinat A., 313
Geissbühler A., 15 , 163 , 287 , 313
Gerard-Blanluet M., 277
Grabar N., 139
Grant A., 69
Grauser D., 185
Grouin C., 23

H

Hery V., 263

J

Jais JP., 277
Jaulent MC., 113
Jilani I., 113
Jonquet C., 151
Joubert M., 47 , 57 , 79 , 101, 251
Jullien S., 171

L

Landais P., 277 , 299
Lastic PY., 227
Le Guillou C., 263
Le Merrer M., 277

Le Mignot L., 277
Lecornu L., 263
Lecroq T., 79
Lemieux R., 69
Lovis C., 15, 35, 185
Lucas N., 227

M

Massari P., 57
Mentre F., 191
Merabti T., 79
Messiaen C., 277
Meyer R., 3
Mougin F., 251
Musen MA., 151

N

Niang M., 163

P

Pariente A., 251
Pereira S., 57 , 101
Poulain P., 125
Pladys P., 125
Prieur E., 101

R

Rath A., 277
Ravery V., 191
Richard JB., 277
Robert H., 125
Rodrigues JM., 91
Rosier A., 23
Rossille D., 239
Rouger P., 171

S

Sakji S., 101
Salomon R., 277
Séroussi B., 191
Serrot E., 57
Shah N., 151
Staccini P., 171
Staub JC., 185

T

Thiessard F., 251
Thillay G., 263
Thomas C., 299
Torre JP., 203
Toubiana L., 299
Traore AK., 163
Traore ST., 163
Trolliard P., 15
Trombert B., 91
Tschopp M., 185

V

Verloes A., 277
Verplancke P., 227
Vessière S., 171

Z

Zweigenbaum P., 23