

JFIM 2014

**15^{es} Journées francophones
d'informatique médicale**
*15th Francophone
Medical Informatics Conference*

**Actes de la conférence
*Proceedings of the Conference***

12–13 Juin 2014
Faculté de médecine et de pharmacie
Fès, Maroc



Production et réalisation par / *Production and manufacturing by*
Pierre Zweigenbaum
SIG francophone de l'IMIA



Avec le soutien de / *Sponsors*



Copyright ©2014 par les auteurs / held by the authors

Également publié en ligne par CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073) :
Also published online by CEUR Workshop Proceedings (CEUR-WS.org, ISSN 1613-0073) :

<http://ceur-ws.org/Vol-1379/>

Préface

La quinzième édition des Journées francophones d'informatique médicale s'est tenue à la Faculté de médecine et de pharmacie de Fès, au Maroc, les 12 et 13 Juin 2014. Ces 15^{es} JFIM étaient organisées par la Société Marocaine d'Informatique Médicale & Santé (SMIMS) et ont eu lieu à la suite du Congrès National d'Informatique Médicale (CNIM) qui s'est tenu les 11 et 12 juin 2014. Le comité scientifique des JFIM était placé sous l'égide du SIG francophone de l'IMIA, en collaboration avec les autres sociétés francophones d'informatique médicale.

Les thèmes de l'appel concernaient les connaissances et informations médicales (ontologies ; standards ; terminologies ; fouille de données et de textes ; traitement automatique des langues ; imagerie et simulation ; aide à la décision et guides de bonnes pratiques ; diagnostic, traitement, pronostic, stratification des risques), les systèmes d'information en santé (développement et évaluation des systèmes d'information cliniques ; infrastructures et systèmes d'information de santé régionaux et nationaux), ainsi que d'autres thèmes associés à l'informatique médicale (facteurs humains ; enseignement et formation en informatique ; qualité, caractère privé et sécurité de l'information ; évaluation des systèmes et de leur utilisabilité cognitive ; questions éthiques et mesures réglementaires en informatique).

Nous remercions l'IMIA qui a donné naissance au SIG francophone, la Société Marocaine d'Informatique Médicale & Santé et le comité d'organisation qui ont pris en charge l'organisation matérielle de la conférence, les membres du comité scientifique qui ont assuré la relecture et la sélection des articles publiés, les auteurs et présentateurs de communications qui ont apporté la matière scientifique à la conférence, et le professeur Francis Roger-France qui a accepté de donner une présentation invitée en ouverture du programme. Tous ont concouru à la réalisation d'une conférence originale et chaleureuse, dans la tradition des JFIM précédentes, ouverte sur la diversité des travaux actuels en informatique médicale.

Pierre Zweigenbaum (LIMSI-CNRS, Paris, France)
Cheick Oumar Bagayoko (FMOS, Bamako, Mali)

Preface

The fifteenth edition of the Journées francophones d'informatique médicale took place at the Faculté de médecine et de pharmacie de Fès, in Morocco, on June 12–13, 2014. These 15th JFIM were organized by the Moroccan Society for Health and Medical Informatics (SMIMS) and took place after the Congrès National d'Informatique Médicale (CNIM) which was held on June 11-12, 2014. The JFIM scientific committee was acting under the auspices of the IMIA Francophone SIG, in collaboration with the other francophone medical informatics societies.

The topics of the call for papers included medical knowledge and information (ontologies ; standards ; terminologies ; data and text mining ; natural language processing ; imaging and simulation ; decision support and practice guidelines ; diagnosis, treatment, prognosis, risk stratification), health information systems (development and evaluation of clinical information systems ; regional and national health infrastructures and information systems), and other topics associated with medical informatics (human factors ; teaching and learning in informatics ; information quality, privacy and security ; evaluation of systems and of their cognitive usability ; ethical questions and regulatory policies in informatics).

We thank IMIA which gave birth to the Francophone SIG, the Société Marocaine d'Informatique Médicale & Santé and the organising committee who carried out the material organisation of the conference, the members of the scientific committee who reviewed and selected the published articles, the authors and presenters of communications who brought scientific content to the conference, and Professor Francis Roger-France who accepted to give an invited talk at the opening of the program. They all concurred to achieve an original and hearty conference, in the tradition of previous JFIMs, open on the diversity of current work in medical informatics.

Pierre Zweigenbaum (LIMSI-CNRS, Paris, France)
Cheick Oumar Bagayoko (FMOS, Bamako, Mali)

Présidents :

Pierre Zweigenbaum (LIMSI-CNRS, Paris, France)

Cheick Oumar Bagayoko (FMOS, Bamako, Mali)

Conférencier invité :

Francis Roger-France

(Université Catholique de Louvain, Faculté de Médecine, Bruxelles, Belgique)

Comité scientifique :

Cheick Oumar Bagayoko (Bamako, Mali)

Mohammed Bennani Othmani (Casablanca, Maroc)

Omar Bouhaddou (San Diego, États-Unis)

Anita Burgun (Paris, France)

Jean Charlet (Paris, France)

Marc Cuggia (Rennes, France)

Stéfan Darmoni (Rouen, France)

Étienne De Clercq (Bruxelles, Belgique)

Somian Francis Ehua (Abibjan, Côte d'Ivoire)

Marius Fieschi (Marseille, France)

Jean-Paul Fortin (Québec, Canada)

Marie-Pierre Gagnon (Québec, Canada)

Antoine Geissbuhler (Genève, Suisse)

Riadh Gouider (Tunis, Tunisie)

Andrew Grant (Sherbrooke, Canada)

Ali Harmel (Tunis, Tunisie)

Marie-Christine Jaulent (Paris, France)

Gustave Karara (Bujumbura, Burundi)

Christian Lovis (Genève, Suisse)

Fleur Mougin (Bordeaux, France)

Innocent Nanan (Abibjan, Côte d'Ivoire)

Aurélie Névéol (Orsay, France)

Ignace Nodjibougada (N'Djaména, Tchad)

Choua Ouchemi (N'Djaména, Tchad)

Sylvia Pelayo (Lille, France)

Jean-Marie Rodrigues (Saint-Étienne, France)

Francis Roger France (Bruxelles, Belgique)

Zineb Serhier (Fès, Maroc)

Brigitte Séroussi (Paris, France)

Pascal Staccini (Nice, France)

Frank Verbeke (Bujumbura, Burundi)

Pierre Zweigenbaum (Orsay, France)

Les JFIM ont été organisées sous l'égide du SIG francophone de l'IMIA, avec des représentants des sociétés francophones d'informatique médicale : MIM (Belgique), ABIM (Burundi), SOCIM (Cameroun), SIBIM (Côte d'Ivoire), AIM (France), SSIM (Suisse), SOMIBS (Mali), SMIMS (Maroc), SOQIBS (Québec), ATIBS (Tchad), ATIM-TELEMED (Togo), STIM (Tunisie).

Comité d'organisation :

Mohammed Bennani Othmani (Casablanca, Maroc, président de la SMIMS)

Chakib Nejjari (Fès, Maroc)

Zineb Sehrier (Fès, Maroc)

ainsi que :

Cheick Oumar Bagayoko (Bamako, Mali)

Omar Bouhaddou (San Diego, États-Unis)

Antoine Geissbuhler (Genève, Suisse)

Brigitte Séroussi (Paris, France)

Pierre Zweigenbaum (Orsay, France)

Actes assemblés avec \LaTeX à l'aide du paquet ACLPUB.

Optimisés pour lecture sur petit écran et impression au format livret.

Proceedings assembled with \LaTeX using the ACLPUB package.

Optimized for small-screen reading and pamphlet printing.

Table des matières

Table of contents

Présentation invitée [*Invited talk*]

Mondialisation de la Télémédecine : utopie ou instrument d'équité ?

[Worldwide Telemedicine: Utopia or a Tool Toward Equity?]

Francis Roger France 1

Articles longs [*Long papers*]

La participation des usagers aux processus de mise en œuvre des programmes de télésanté pour favoriser l'adoption et l'utilisation des services : Une revue de littérature

[Participation of users in the update of telehealth programs to facilitate service adoption and use: a literature review]

Hassane Alami, Marie-Pierre Gagnon & Jean-Paul Fortin 9

Recommandations de pratique clinique sous la forme d'un document textuel et d'un système d'aide à la décision : une expérience de production et diffusion simultanées des deux types de format pour la prise en charge du diabète de type 2

[Clinical practice guidelines as a textual document and a clinical decision support system: simultaneous development and publication of both formats in the management of type 2 diabetes]

Jacques Bouaud, Hector Falcoff, Valérie Ertel-Pau, Véronique Raimond & Brigitte Séroussi 21

Mapping de terminologies diagnostiques en oncologie par l'intermédiaire du NCI Metathesaurus

[Mapping of diagnostic terminologies in oncology using the NCI Metathesaurus]

Bérénice Brechat, Fleur Mouglin, Frantz Thiessard & Vianney Jouhet 34

Terminologies et référentiels d'interopérabilité sémantique en Santé

[Reference terminologies and templates for semantic interoperability in healthcare]

Amina Chniti, Nicolas Griffon, Lamine Traore, Sajjad Hussain, Jean-Marie Rodrigues, Stéfan J. Darmoni, Jean Charlet, Éric Sadou, David Ouagne, Éric Lepage & Christel Daniel 44

Intégration des données de prescription dans un entrepôt de données biomédicales

[Integration of prescription data in a clinical data warehouse]

Kevin Dalleau, Denis Delamarre, Thibault Ledieu & Marc Cuggia 59

Utilisation des techniques de modélisation des connaissances pour la comparaison de recommandations de pratique clinique : application à la prise en charge de l'hypertension
[Using knowledge modeling to compare clinical practice guidelines: a case study with the management of hypertension]

Alexandre Galopin, Jacques Bouaud, Suzanne Pereira & Brigitte Séroussi 71

Enrichissement sémantique associé à la détection de la négation et des antécédents familiaux dans un entrepôt de données hospitalier
[Semantic enrichment based on detection of negation and history family in a French hospital data warehouse]

Nicolas Garcelon, Rémi Salomon & Anita Burgun 83

Avantages, effets et préoccupations reliés au dossier de santé personnel électronique (DSPé) : Une mise à jour de la littérature
[Benefits, Effects and Concerns Related to the Electronic Personal Health Record: Update of the Literature]

El Kebir Ghandour, Marie-Pierre Gagnon & Jean-Paul Fortin 94

La satisfaction des utilisateurs des Systèmes d'Information Clinique du Groupe Hospitalier Paris Ouest (HUPO)
[Clinical Information System User Satisfaction in the Paris-West University Hospital Group (HUPO)]

Brahim Hadji, Isabelle Dupuis, Laurence Leneveut, Didier Heudes, Jean-François Wagner & Patrice Degoulet 115

Indicateurs en biologie et en imagerie au sein des systèmes d'information de santé
[Impact indicators for radiology and biology information systems]

Mher B. Joulakian, Nicolas Griffon, Matthieu Schuers, Éric Lepage, Céline Savoy-Collet, Soumaya Skalli, Philippe Massari & Stéfan J. Darmoni 127

Moteur de recherche sémantique au sein du dossier du patient informatisé : langage de requêtes spécifique
[Semantic search engine in the computerized patient record: specific query language]

Romain Lelong, Tayeb Merabti, Julien Grosjean, Mher B. Joulakian, Nicolas Griffon, Badisse Dahamna, Marc Cuggia, Suzanne Pereira, Natalia Grabar, Frantz Thiessard, Philippe Massari & Stéfan J. Darmoni 139

Usage secondaire du dossier médical informatisé à des fins épidémiologiques et d'évaluation de la qualité des soins : le projet SYNODOS
[A secondary use of the electronic health record for epidemiological studies and evaluation of health care quality: the SYNODOS project]

Marie-Hélène Metzger, Quentin Gicquel, Domoïna Rabarijaona, Frédérique Segond, André Bittar, Luca Dini, Lina F. Soualmia & Stéfan J. Darmoni 152

Protection des données de santé lors de leur réutilisation à des fins de recherche au sein des entrepôts de données biomédicaux (EDBM)

[Health data protection for secondary use in clinical data warehouse]

Christine Riou & Marc Cuggia 162

Articles courts [Short papers]

Place de la technologie Smartphone en santé mentale

[Smartphone technology in mental health]

Dalal Ben Loubir, Samy Housbane, Zeineb Serhier & Mohammed Bennani

Othmani 188

Un outil de visualisation de classifications et d'intégration de données phénotypiques et génétiques pour faciliter le codage des maladies rares

[Coding rare diseases in health information systems: a tool for visualizing classifications and integrating phenotypic and genetic data]

Rémy Choquet, Yannick Fonjallaz, Albane de Carrara, Meriem Maaroufi, Pierre-Yves Vandebussche, Ferdinand Dhombres & Paul Landais 198

Les technologies de l'information et de la communication (TIC) et la promotion de la santé sexuelle auprès des adolescents et jeunes adultes : au-delà de l'efficacité

[Information and communication technologies (ICT) and the promotion of sexual health among adolescents and young adults: beyond the effectiveness]

Serge Djossa Adoun & Marie-Pierre Gagnon 204

Transition vers la mobilité dans le monde médical : n'oubliez pas l'utilisabilité

[Switching to handheld tools in healthcare: don't forget usability]

Frédéric Ehrler, Évelyne Sarrey, Magali Walesa, Rolf Wipfli & Christian Lovis .. 211

Importance de la caractérisation des situations de travail dans l'évaluation de l'impact des Systèmes d'Information en Santé

[Importance of the work system characterization for impact evaluation of Health Information Technology]

Ludivine Watbled, Sandra Guerlinger, Marie-Catherine Beuscart, Stéphanie Bernonville, Régis Beuscart, Philippe Massari, Éric Lepage & Sylvia Pelayo 219

Index

Index des auteurs [Author index] 229

Mondialisation de la Télémédecine : utopie ou instrument d'équité ?

Worldwide Telemedicine: Utopia or a Tool Toward Equity?

Francis Roger France

Université Catholique de Louvain, Faculté de Médecine, 1200 Bruxelles, Belgique

Résumé

La télémédecine est une pratique de la médecine à distance à l'aide des technologies de l'information et de la communication. Elle autorise une relation médecin-malade en ligne ainsi qu'un dialogue entre professionnels de la santé par un échange de sons, d'images et de données dans le monde entier. Est-ce une utopie ? Cette médecine futuriste pourrait ne pas réussir en raison du coût important de sa technologie et, surtout, par inadéquation du dialogue entre personnes, bref, en évoluant vers une déshumanisation de la pratique médicale. Par contre, la télémédecine pourrait, à l'inverse, être une opportunité extraordinaire d'économies par réduction des hospitalisations et des consultations, par un retour des généralistes dans des territoires qu'ils ont déserté, en leur donnant un accès à distance à des spécialistes. Par ailleurs, elle ouvre la voie aux soins à distance de populations isolées en cas de désastre majeur. Elle offre aussi la possibilité au sein de pays émergents de recourir à des centres de diagnostics ou de traitements dans le monde entier. Partout, elle permet de « passer la main » là où la médecine est limitée, ce qui augmente la qualité des soins et autorise un accès plus équitable aux meilleurs soins. Pour atteindre de tels objectifs, il faut prévoir les obstacles à surmonter (technologiques, économiques, juridiques, éthiques, politiques et sociaux) en bâtissant des modèles réalistes, sécurisés, acceptables à l'échelle internationale, basés sur des expériences solides.

Abstract

Telemedicine is a medical practice at distance, using information and communication technologies. It enables on-line physician-patient relationship as well as a partnership between health professionals by enabling a worldwide exchange of sounds, images and data. Is it an utopia? This futuristic medicine could not succeed, due to important technologic costs and, mainly, because human dialogue might be inappropriate, leading to dehumanize medical practice. On the contrary, Telemedicine might offer an exceptional opportunity to reduce health costs, by diminishing hospital stays and outpatients visits as well as by giving incentives to general practitioners who deserted wide regions to come back, because of integrated care with specialized physicians from far away. Furthermore, it opens the way to medical care from abroad to isolated populations in case of major disaster. It enables also emerging countries to obtain access to reference diagnostic and treatment centers. It allows everywhere to "make way to someone else" when health care possibilities are limited, resulting in an improvement of quality of care as well as in obtaining more equity for access to better care. In order to reach these objectives, several obstacles (technology, economy, laws, ethics, sociology

and politics) have to be foreseen and overcome by building realistic models, secure, internationally acceptable, based on strong experiments.

Mots-clés : télémédecine ; économie des soins de santé ; équité d'accès ; qualité des soins ; sécurité.

Keywords: telemedicine; health economics; equity of access; quality of care; security.

1 Caractère universel de la télémédecine

Télé signifie à *distance*. C'est en 1920 que le téléphone a été reconnu officiellement aux États-Unis comme un moyen de traitement de patients sur des bateaux. Depuis lors, la télévision a ajouté l'image au son. Actuellement, une miniaturisation des capteurs et des émetteurs permet de faire des ECG ou des EEG reliés à un portable du patient à tout endroit dans le monde. De plus, l'informatique et la télématique donnent la possibilité d'aide à la décision en utilisant la robotique et l'intelligence artificielle. Les *télécommunications* se sont largement développées. Les informations peuvent passer par des satellites tournant autour du globe, captées par des paraboles sur terre. L'alternative la plus utilisée est la communication d'informations digitalisées par l'intermédiaire de réseaux de câbles intercontinentaux, sous terre et au sein des océans.

Les applications à distance couvrent dès à présent *la plupart des domaines de la médecine* [1–3]. Citons en premier lieu la téléconsultation. Elle permet un télédiagnostic par l'utilisation d'un robot captant les images et les paroles du malade, tout en dialoguant avec le médecin traitant local qui peut procéder à un examen clinique ou explorer des organes par endoscopie ou laparoscopie et enregistrer les bruits obtenus par un stéthoscope. La *téléradiologie* donne accès à l'avis de spécialistes dans le monde entier, surtout pour la résonance magnétique. La télépathologie est une des branches les plus utiles, car il n'y a pas assez d'anatomopathologues dans de nombreux pays.

Plus impressionnant pour le grand public, il est possible aussi de pratiquer de la *téléchirurgie*, surtout si le médecin sur place utilise la laparoscopie (chirurgie digestive, ophtalmologique, neurochirurgie, chirurgie gynécologique,...).

Les capteurs (mini-ordinateurs portables sur le patient) sont lancés par l'industrie [4] pour le monitoring (*télesurveillance*) ainsi que pour la téléadaptation. La plupart des spécialités sont concernées : la télécardiologie, la téléorthopédie, la télépsychiatrie, la télédentisterie, la téléneurologie, la télépharmacie, le télé nursing sont toutes des branches qui ne demandent pas une spécialité télé mais bien une spécialité médicale qui peut être pratiquée de loin.

2 Plaidoyer pour le développement d'une télémédecine à caractère humain

La télémédecine implique une adaptation à l'être humain. En effet, la médecine peut être définie comme l'art et la science de traiter des malades. L'art consiste à découvrir chez le patient des signes pertinents pour poser un diagnostic. Le praticien le fera par l'écoute, la vision, la palpation. Il doit aussi sentir, humer, avoir des intuitions. Toutefois, le médecin ne comprendra les plaintes du patient que s'il parvient à situer les signes observés dans le cadre des connaissances actuelles pour pouvoir traiter adéquatement ces informations. Pour cela, il dispose d'une science. Celle-ci a créé des modèles de connaissances : l'anatomie, la physiopathologie, la microbiologie, la biochimie, la génétique. L'expert en médecine doit avoir les connaissances suffisantes de ce qui existe actuellement.

Il doit continuellement s'adapter aux progrès liés au développement des sciences fondamentales et

appliquées. Les nouvelles technologies ont continuellement modifié les paradigmes des sciences. Ainsi, au 19^e siècle, la microscopie a permis de voir des bactéries associées à certaines maladies. La tuberculose était définie antérieurement au microscope de manière très différente de celle qui a tenu compte de la présence du bacille de Koch, non seulement dans les poumons mais aussi dans les os, les reins, etc. Auparavant, ces atteintes étaient considérées comme des maladies séparées.

Une telle médecine, à l'écoute du malade est-elle possible à distance ?

La reconnaissance du diagnostic implique une reconnaissance d'un modèle de connaissance. *La maladie en soi n'existe pas !* Ce qui existe, ce sont des malades. Ceux-ci ont des symptômes caractéristiques d'un modèle de connaissance appelé maladie.

Le médecin procède par inductions et déductions [5]. Il cherche à reconnaître des signes, des symptômes. Il en induit des possibilités de diagnostics. En fonction de celles-ci, il va déduire qu'il devrait trouver un autre symptôme ou un autre élément biologique ou anatomique. Et il va confirmer ou infirmer son hypothèse diagnostique. Ainsi, si un malade est ictérique, il va se poser la question d'une hépatite, d'une lithiase, d'une néoplasie, par induction. Il va procéder ensuite à une déduction en dosant les SGOT, SGPT et les Gamma-GT (des tests biologiques) ainsi que les anti-gènes et anti-corps de l'hépatite. Admettons qu'il trouve qu'il s'agit d'une hépatite C, il va peut-être vouloir procéder à une biopsie du foie mais il ne pourra la faire que dans la mesure où il n'y a pas de contre-indication, comme un nombre trop bas de plaquettes.

À notre sens, le raisonnement médical peut être pratiqué aussi bien à proximité qu'à distance du malade dans la mesure où les données sont bien caractérisées et bien décrites et si le médecin a les connaissances suffisantes pour retrouver les modèles adéquats..

Les signes pathognomoniques sont rares, Par exemple la douleur, dans le gros orteil au chant du coq qui signe la goutte. Mais pour donner un autre exemple de ce qui est difficile à traiter, car mal connu à distance, citons l'histoire d'une tache noire sur le thorax d'un patient, qui venait de St Tropez, dans le midi de la France, où il campait avec son chien. Il avait de la fièvre, des jambes rouges, des taches importantes et l'assistant belge n'avait jamais vu cela. Le spécialiste a de suite demandé : « avait-il une tache noire » quelque part ? « Oui, sur le thorax, décrite dans le dossier ». Il s'agissait d'une piqûre de tique et d'une fièvre méditerranéenne. Celui qui connaissait le modèle a trouvé directement le diagnostic et le traitement.

L'aide au diagnostic par ordinateur est plus simple lorsque les signes sont isolés et caractéristiques comme un souffle au cœur, une insuffisance rénale, une douleur exquise ou un ictère. Par contre s'ils sont intertriqués, cela devient plus difficile.

Les éléments contribuant au diagnostic sont essentiellement l'anamnèse 82 %, l'examen clinique seulement 9 %, et les examens complémentaires 9 % également. Les connaissances requises sont la mémoire à long terme, une bonne notion de la nosologie et une bonne maîtrise de la sémiologie, de l'épidémiologie et de la physiopathologie.

3 La télémédecine : une utopie ?

S'agit-il d'une chimère c'est-à-dire d'un monstre fabuleux, d'un rêve invraisemblable ? Peut-on rêver d'une médecine pouvant se passer de médecin ? De diagnostics établis uniquement par une intelligence artificielle ? D'opérations à distance faites par des robots ?

Les opposants au développement de la télémédecine citent de nombreuses désillusions. Il y a sur-

tout un problème d'interopérabilité, d'incompréhension entre médecins, de manque de connaissance suffisante de la langue du patient, de lacunes d'interfaces, de manque de standards.

Il peut aussi y avoir des défaillances technologiques comme le satellite acheté par la Mayo Clinic qui est tombé en panne, comme la rupture de réseaux câblés par des travaux, comme la difficulté d'installation de paraboles (cela s'est passé au Mali). Il peut y avoir aussi, pour l'informatique, une insuffisance des outils d'intelligence artificielle en cas de pathologies intriquées et complexes.

Il faut citer les problèmes de gouvernance, d'organisation et de financement. Il peut y avoir un risque de monopole de centres d'expertise dans les pays du Nord et pas dans les pays du Sud. Comment payer localement l'acte médical et comment le payer à distance ? Des inégalités de modèles de couverture sociale de la santé dans le monde ne permettent pas facilement de résoudre cette question.

Citons aussi un problème de disponibilité médicale et d'éthique. Sera-ce une efficacité de la médecine du Nord au bénéfice de la médecine du Sud ? Y a-t-il assez de médecins pour la demande ? Par exemple, dans des pays aussi grands que les Indes ou le Maroc y a-t-il suffisamment d'anatomopathologues ? D'autre part, en Europe et aux États-Unis d'Amérique, y a-t-il assez de reconnaissance professionnelle nationale et internationale ? Ne faut-il pas créer des licences européennes, américaines ou asiatiques ?

Enfin, il y a des aspects légaux qui ne sont pas suffisamment complets. Il faut clarifier les responsabilités à proximité et à distance du patient et sécuriser les informations, promouvoir la qualité en pénalisant ceux qui ne passent pas la main en médecine.

4 La télémédecine : un instrument d'équité des soins ?

La télémédecine peut, au contraire, être une opportunité extraordinaire d'une plus grande égalité des chances d'avoir accès aux soins dans tous les pays du monde. Elle a montré dans plusieurs pays un retour des médecins généralistes dans des zones désertées par eux car ils peuvent organiser une aide par télémédecine de centres de soins distants. Elle permet aussi une mondialisation de l'accès aux spécialistes et un travail en commun en équipe de médecins au service du patient.

Il y a des expériences stimulantes. Des régions rurales éloignées et désertées par les médecins ont montré l'utilité d'interactions directes avec des patients par téléaccès à des spécialistes. Les soins à distance existent pour de nombreuses spécialités. La plus fameuse réalisation est sans doute celle liée aux accidents vasculaires cérébraux (indication dans les 3 heures de produits détruisant les caillots) mais de nombreuses possibilités existent en orthopédie, en cardiologie et en réanimation.

Dans les bateaux, certains programmes d'aide au diagnostic ont fait leurs preuves dont ceux développés par de Dombal pour les abdomens aigus. Dans les pays en développement et en Europe de l'Est, l'accès à des centres spécialisés dans des domaines peu disponibles chez eux comme la télépathologie, l'imagerie sophistiquée médicale, peut être un pas en avant important. En cas de désastre majeur, de tsunami, de contamination nucléaire, de tremblements de terre, les médecins ne peuvent plus se déplacer, ils doivent rester sur place et la présence de robots expérimentés auparavant, permet de soigner les patients non seulement pour des diagnostics mais aussi pour des traitements y compris chirurgicaux.

Dans tous les pays la télémédecine peut-être un moyen de suivre les *maladies chroniques*. Cela existe en Belgique pour l'insuffisance rénale, en Hollande pour le diabète, dans plusieurs pays pour la *télépharmacie* pour les doses, les interactions, les contre-indications. La miniaturisation des capteurs

permet une meilleure télé-surveillance et un *télé-nursing*.

Les services considérés comme encore expérimentaux mais qui existent, sont la chirurgie à distance aussi bien en ophtalmologie qu'en chirurgie digestive et en neurochirurgie. La télééducation et la télérevalidation, la neuropsychologie et la médecine du travail, l'audiophonologie sont encore d'autres applications en cours de développement.

Parmi les démonstrations d'efficacité et de mondialisation de la télémédecine citons :

- en Belgique, la téléradiologie à Malines, la téléneurologie à l'UCL Mont-Godinne, la télédialyse un dossier fédéral ;
- en France et en Allemagne, la télépathologie, entre autres à Dijon et Heidelberg ainsi que la télésurveillance ;
- en Norvège, les téléconsultations et la télédialyse ;
- au Danemark un réseau national à larges bandes de télésanté et de télépneumologie ;
- au Pays-Bas de la téléophtalmologie, dermatologie et cardiologie ;
- au Royaume-Uni un programme national de télésanté ;
- aux États-Unis d'Amérique, le téléHealth neuro-cardio-imagerie self management avec comme pionnier la Mayo Clinic à Rochester, Minnesota (qui travaille avec des hôpitaux à Jacksonville en Floride) et la Kaiser Permanente Foundation en Californie. La télédermatologie, la téléorthopédie et la télésurveillance, entre autres à l'université du Texas ;
- au Canada, la télérobotique laparoscopique permet la chirurgie abdominale en Ontario.

Des pays vastes comme l'Australie, ont permis le développement d'un réseau national à larges bandes avec téléconsultations synchrones et enregistrements d'envois différés asynchrones. Les pays demandeurs et en cours de travaux de réalisation internationaux sont en Afrique du Nord et au Moyen-Orient, le Maroc et l'Algérie avec la France et les USA surtout, et l'Égypte. En Asie, le Japon, la Thaïlande, la Corée du Sud, le Pakistan avec les USA, les Indes avec la France, les USA et le Royaume-Uni. Le Bangladesh avec l'Australie et parmi les pays de l'Est demandeurs, citons l'Albanie, la Bulgarie et la Roumanie.

5 Conditions d'une mondialisation ?

5.1 Aspects éthiques

La responsabilité professionnelle doit être bien définie car c'est d'un dialogue entre médecins et autres professionnels de la santé que va dépendre le succès ou l'insuccès de la télémédecine. Concernant la responsabilité professionnelle, le médecin traitant devrait être considéré comme entièrement responsable de son patient y compris de la demande d'expertise. Le spécialiste à distance, donne un avis argumenté, une aide pour traiter mais il le fait dans la limite des informations transmises. Il est responsable des interventions qu'il a entreprises. Enfin le robot est sous la responsabilité de celui qui l'a conçu et de celui qui le téléguide, par exemple pour la chirurgie à distance.

Un deuxième aspect éthique est celui de la sécurité des systèmes d'informations. Sécurité définie par la confidentialité, l'intégrité et l'accessibilité des données.

5.2 Aspects financiers

La question de la rentabilité des applications est posée. Mais il manque des évaluations solides et suffisamment larges à ce sujet. Il existe bien sûr un coût des équipements, des connections à distance.

Mais il faut voir quel est leur bénéfice par rapport aux dépenses. Et là, les réductions des hospitalisations, du transport des patients, des consultations et de l'absentéisme, permettent des économies considérables. Il y a aussi des économies d'échelle et un accroissement industriel. Cette rentabilité doit être évaluée aussi en fonction de la qualité des soins. Malheureusement, il faut constater une insuffisance de critères d'évaluation de la qualité des soins avec ou sans télé-médecine.

5.3 Des expériences pilotes pour convaincre

Plus que tout, c'est probablement le manque de conviction qui a freiné le développement de la télé-médecine pourtant en expansion. Il existe des expériences pilotes mais pas encore de développement à grande échelle. Il n'y a pas assez d'initiatives de centres hospitaliers. Certains expriment leurs craintes de défaillances technologiques. Il faut tenir compte d'un manque d'information des dirigeants. Citons aussi l'opposition des médecins à l'aspect commercial des contrats dans le cadre d'initiatives privées. Enfin, les patients n'ont pas encore compris qu'ils pouvaient mettre la pression pour obtenir de meilleurs soins, c'est-à-dire obtenir un accès à des experts reconnus dans le monde entier.

Les conditions d'une mondialisation sont d'obtenir *une offre fiable de télé-médecine avec disponibilité de services de spécialistes reconnus*, par une multiplication d'expériences pilotes [6–8], dans tous les pays qui peuvent l'obtenir et une mise en place de critères de performance et de qualité des soins. Pour répondre à une demande de population en manque de médecins traitants à la recherche de meilleurs spécialistes, il est important de développer des expériences là où l'isolement, l'éloignement, un tsunami, ou une contamination nucléaire, ont bloqué la possibilité d'accès à certains soins médicaux. Par ailleurs, il faut tenir compte de l'augmentation des pathologies chroniques par le vieillissement de la population demandant un suivi régulier qui pourrait être obtenu par la télé-médecine.

5.4 Couverture sociale

Mais tout cela requiert une universalisation de la couverture sociale des soins de santé, ce qui reste un grand projet non encore suffisamment réalisé.

5.5 Être prêt en cas de désastre majeur

Fukushima représente un exemple de recherche de nouveaux systèmes d'information et de communication. Ainsi, le professeur Kiyomu Ishikawa au moment de Fukushima m'a interrogé concernant les modèles de dossiers électroniques en cas de désastre majeur. Ma réponse a été celle du résumé clinique européen qui a été développé dans de nombreux pays, dont l'Italie. À Udine, il y a eu un tremblement de terre. Un système d'information existant a permis d'adapter le registre hospitalier d'admissions pour enregistrer les personnes vivantes, décédées et disparues et il a été possible d'adapter le résumé d'autres dossiers (MBDS Européen) aux pathologies à suivre (posttraumatiques, dues à des contaminations ou chroniques, exigeant des soins continus, comme le diabète).

Il serait possible, lorsque des médecins ne peuvent plus quitter leur pays, de relier ces informations à des demandes d'applications de télé-médecine comme la téléorthopédie, la télépharmacie, la télésurveillance et de contacter divers réseaux. Dans cette application, le Prof. Guha-Sapir (UCL et OMS) spécialiste en épidémiologie des désastres s'est rendue au Japon et le Prof. John Evans (de Harvard

Cyprus University) nous a transmis toutes les possibilités d'aide humanitaire qui ont été envisagées et obtenues à Haïti, pays beaucoup plus pauvre. Une conférence internationale s'est tenue en Israël à ce sujet, organisée par le Docteur Assa Reichert.

5.6 Position Fédérale Belge

Un rapport (136B) du KCE Kennis Centrum-Centre d'Expertise, datant de 2010, a analysé la surveillance à distance des patients porteurs de défibrillateurs implantés. Il s'agit d'une évaluation de la technologie et du cadre réglementaire général de la télémédecine.

Le KCE est un centre d'expertise qui permet en tant que parastatal de donner des avis au gouvernement pour déterminer quels sont les examens et les traitements qui peuvent être remboursés en Belgique. En matière de télémédecine, le KCE estime que les principes juridiques spécifiques à ces aspects sont insuffisants. Il faut préciser les responsabilités des médecins et aller au-delà des droits du patient. Il faut recommander des bonnes pratiques. Le KCE recommande de ne pas obtenir de remboursement par l'INAMI (Institut National d'Assurance contre la Maladie et l'Invalidité) en l'absence de données probantes sur la sécurité et l'efficacité et le rapport coût-efficacité de la télémédecine. Il propose de faire une convention entre l'INAMI et un centre de suivi à distance où les conditions spécifiées sont remplies et si un système d'enregistrement rigoureux permet des recherches et de réévaluer la situation.

À notre sens cette étude est basée sur une application très scientifique mais très limitée, qui n'autorise pas les recommandations générales émises, car elle ne tient pas compte du vaste champ d'applications de la télémédecine. Le KCE méconnaît les publications internationales entre autres en France, aux États-Unis et en Australie en matière de responsabilité, de sécurité et de code de bonne pratique. Il ignore les possibilités de la télémédecine en cas de désastre majeur en Belgique, risque de condamner à mort des victimes non couvertes par la sécurité sociale pour des soins à distance. Ce blocage du financement de développement de centres de référence utilisant la télémédecine, ne permet pas de maintenir une compétence médicale et une compétitivité. À terme, il s'agit d'une mise en péril des soins de santé dans le pays, alors qu'il y a une opportunité d'engager des médecins, de développer les outils qui autoriseraient une économie de services avec les pays émergents.

6 Conclusion

La télémédecine : une illusion ? Croire qu'un robot aura les mêmes capacités intellectuelles, autant de sensibilité que l'homme est probablement illusoire.

Un renforcement de la demande de collaboration par des applications variées de télémédecine de la part de médecins sur place et à distance utilisant de nouveaux robots pourrait aider les patients et permettre un apprentissage progressif accompagné d'évaluations du coût et de l'efficacité des soins.

Par ailleurs, la télémédecine offre une possibilité unique d'obtenir une meilleure équité dans les soins.

Si les médecins et les patients comprennent la possibilité de cette télémédecine et font confiance aux progrès technologiques sous contrôle, et si nos gouvernants pouvaient avoir une vision de l'avenir, ils mettraient probablement en œuvre une stratégie autorisant une redistribution des professionnels de la santé, un meilleur accès des patients du monde entier à des soins de qualité à un moindre coût.

Dans cette perspective, l'Union Européenne devrait donner les moyens d'une mondialisation des services à des experts par des centres d'excellence à identifier et à proposer à l'extérieur [9].

Il faut clarifier les aspects légaux et éthiques de sécurité et de responsabilité. Le développement de la télémédecine implique un renversement des tendances à l'autodéfense professionnelle comme le numerus clausus aboutissant à une pénurie de spécialistes. Pourquoi ne pas en former plus en utilisant les moyens financiers consacrés à cette demande dans les pays émergents, tout en les associant à une formation similaire à la nôtre ?

Références

- [1] Shortliffe E, Perreault L, Wiederhold G, and Fagan L, eds. *Medical Informatics — Computer Applications in Health Care and Biomedicine*. Health Informatics. Springer, 2001.
- [2] Albert A, Roger-France F, Degoulet F, and Fieschi M, eds. *Santé et réseaux informatiques*, (vol10) of *Informatique et Santé*. Springer, Paris, 1998.
- [3] American Telemedicine Association . <http://www.americantelemed.org/>. Retrieved 21 August 2011.
- [4] Gozian M. Un médecin à portée de clavier. *Le Monde Sciences et Technologies* 3 janvier 2013.
- [5] Roger-France F. A structured clinical presentation of case to medical students. In: Pagès J, Levy A, Grémy F, and Anderson J, eds, Meeting the challenge: informatics and medical education, Amsterdam. North-Holland, 1983:193–9.
- [6] Berman M and Fenaughty A. Technology and managed care: patient benefits of telemedicine in a rural health care network. *Health Economist* June 2005;14(6):559–73.
- [7] Gilbert BK, Mitchell MP, Bengali AR, and Khandheria BK. NASA/DARPA advanced communications technology satellite project for evaluation of telemedicine outreach using next-generation communications satellite technology: Mayo foundation participation. *Mayo Clinic Proceedings* Aug 1999;74(8):753–7.
- [8] Meschia JF. Camera in the emergency department: the evolution of stroke telemedicine. *Mayo Clinic Proceedings* 2009;84(1):3–4.
- [9] Commission européenne . Communication de la Commission au Parlement européen, au Conseil, au Comité économique et social européen et au Comité des régions du 4 novembre 2008 concernant la télémédecine au service des patients, des systèmes de soins de santé et de la société. [COM(2008) 689 final — Non publié au Journal officiel].

Adresse de correspondance

Francis ROGER FRANCE, MD, MS, PhD
Commission Européenne
Service Médical de Bruxelles
Avenue d'Auderghem,19
1040 Bruxelles, Belgique
Courriel : francis@rogerfrance.com

La participation des usagers aux processus de mise en œuvre des programmes de télésanté pour favoriser l'adoption et l'utilisation des services : Une revue de littérature

Hassane Alami^{1,4}

Marie-Pierre Gagnon^{1,2}

Jean-Paul Fortin^{3,4}

¹*Axe Santé Publique et Pratiques Optimales en Santé, CRCHUQ, Hôpital Saint-François d'Assise, Québec, Canada*

²*Faculté des Sciences Infirmières, Université Laval, Québec, Canada*

³*Faculté de Médecine-Département de Médecine Sociale et Préventive, Université Laval, Québec, Canada*

⁴*Centre de Recherche sur les Soins et Services de Première Ligne au CSSS-VC, Québec, Canada*

Résumé :

Contexte : Le défi qu'affronte la mise en place des services de télésanté est celui de réussir à intégrer le système de soins et faire partie intégrante de la routine clinique. En effet, une grande partie des projets de télésanté peine à survivre au-delà de l'étape de projet pilote. Pour la majeure partie de ces projets, il s'avère que les citoyens sont rarement consultés quant à leurs besoins et attentes concernant les services de télésanté. **Objectifs :** Cette revue vise à explorer des méthodes favorisant la participation des citoyens-usagers à la planification et l'implantation des services de télésanté, et l'impact de cette participation sur l'adoption et l'utilisation de cette technologie.

Méthodes : Une revue de littérature a été réalisée, sans limitation temporelle. Les bases de données MEDLINE, EMBASE, Cochrane Library, CINAHL et CAIRN ont été consultées. Les articles ayant abordé la participation du public, des citoyens, des patients ou des usagers dans les décisions concernant les services de télésanté ont été analysés. **Résultats :** Sur les 43 articles identifiés par le biais de la stratégie de recherche, 10 ont satisfait nos critères d'inclusion et ont été retenus dans la présente revue. La participation des citoyens aux processus de planification et d'implantation des services de télésanté constitue une voie prometteuse pour améliorer les conditions d'adoption et d'utilisation de la télésanté par la population.

Abstract

Background: The challenges facing the implementation of telehealth services are to successfully integrate the health care system and make an integral part of clinical routine. Yet, a large proportion of telehealth projects barely survive beyond the pilot stage. For most of these projects, it appears that the citizens are rarely consulted about its needs and expectations regarding telehealth services. **Objectives:** This review aims to explore methods allowing citizens to participate in the planning and implementation of telehealth services, and the impact of their participation in the adoption and use of this technology. **Methods:** We conducted a literature

review using MEDLINE, EMBASE, Cochrane Library, CINAHL and CAIRN databases, without time limitation. We analyzed articles discussing public participation of citizens, patients or users in decisions regarding telehealth services. **Results:** on 43 articles identified through the search strategy, 10 met our inclusion criteria and were thus considered in this review. The participation of citizen-users in planning and implementation of telehealth services represents a promising way to improve the conditions of adoption and use of telehealth by the population.

Mots clés: télésanté, citoyen-usager, empowerment, services de santé.

Keywords: telehealth, citizen-user, empowerment, health services.

1. Introduction

Plusieurs pays sont confrontés à la problématique de l'accès à la prise en charge médicale, surtout au niveau des territoires ruraux et isolés. En effet, l'offre de soins est souvent concentrée dans des zones dites « favorisées », notamment urbaines [1]. Cette disparité territoriale reflète la réalité d'une médecine à deux vitesses qui se traduit par des « inégalités territoriales de santé » importantes, entre autres [2].

Dans le but d'améliorer l'accessibilité et la continuité des services de santé et remédier à la pénurie des médecins dans certaines spécialités, plusieurs pays ont investi dans les nouvelles technologies de l'information et de communication (TIC), notamment la télésanté, pour permettre de répondre aux besoins de services de leurs populations [3].

La télésanté désigne « *les soins et les services de santé ainsi que les services sociaux, préventifs ou curatifs, rendus à distance par le biais d'une télécommunication, incluant les échanges audiovisuels à des fins d'information, d'éducation et de recherche, et le traitement de données cliniques et administratives* » [3].

Malgré l'intérêt qu'elle suscite et les avancées considérables qu'elle a connu ces vingt dernières années, la télésanté se confronte, plus que jamais, à la problématique de l'intégration en routine clinique. En effet, malgré toutes les attentes construites sur cette technologie et son rôle dans l'amélioration de l'accès aux soins, l'état des connaissances et des réalisations actuelles est encore loin de permettre d'aspirer à son intégration immédiate en tant que forme de pratique médicale à part entière [4]. Cette situation s'explique en grande partie par le fait que l'immense majorité des expérimentations ne dépasse pas le stade du projet pilote, pour aspirer à la mise en place de services durables [5]. À titre d'exemple, Sanders et collaborateurs ont souligné, dans le cadre d'une étude sur les télésoins à domicile, que les freins à l'adoption de la télésanté vont au-delà de la seule dimension technologique [6]. À l'occasion, ils concluent sur le fait que les usagers ont besoin de temps et davantage d'informations pour pouvoir accepter d'utiliser cette technologie. Ainsi, les auteurs estiment que pour améliorer l'acceptabilité de la télésanté, et son utilisation par la suite, il faudrait prendre en considération les attentes et les besoins des différents acteurs concernés par la question (ex : patients, professionnels de santé, etc.). Des conclusions

que partagent aussi Abbott et collaborateurs, dans leur revue systématique sur l'utilisation de la télésanté [7]. De leur côté, Gagnon et collaborateurs parlent du critère multidimensionnel et complexe de l'adoption et l'utilisation de la télésanté [5]. En ce sens, ils soulignent l'importance de la compréhension des différents facteurs individuels et organisationnels influençant l'adoption, l'utilisation et la diffusion de la télésanté. Ils mettent, ainsi, l'emphase sur la nécessité d'adopter des stratégies permettant d'associer les usagers (ou futurs usagers) aux différents stades de développement et de mise en œuvre de ces programmes.

En résumé, ces études montrent principalement que la réalité du terrain dépasse le simple enjeu technique de la télésanté, laissant émerger d'autres enjeux, d'ordre économique, organisationnel, juridique, sociopolitique, éthique et humain. Ces questions n'ont pas été suffisamment considérées jusqu'à maintenant et il importe de s'en préoccuper afin de mieux encadrer le développement et la mise en œuvre des programmes de télésanté.

Dans le même esprit, ces travaux ont souligné l'importance de l'adoption de stratégies favorisant la prise en considération de la perspective « usager » dans les projets relatifs à l'implantation des services de télésanté. Ces recommandations viennent appuyer l'attention que reçoit la participation des citoyens à la réflexion et l'établissement des priorités en termes de services de santé dans de nombreux pays [8]. En effet, de nombreux travaux soulignent que l'association des citoyens au processus décisionnel des réseaux de santé facilite l'implantation des innovations et améliore la pertinence des recherches tout en favorisant la promotion de l'expertise citoyenne [9, 10, 11].

Cet ensemble d'indications nous amène à considérer deux éléments majeurs. Le premier, porte sur la problématique de l'adoption et l'utilisation de la télésanté, notamment par les usagers des services de santé. Le deuxième élément concerne l'importance d'associer les citoyens au processus de l'implantation de ces services sur leur territoire. Cette démarche pourrait augmenter l'acceptation de ces services et par conséquent leur utilisation.

L'analyse des expériences internationales montre un intérêt grandissant pour l'association des citoyens au développement et l'implantation des projets de télésanté, notamment au Canada et au Québec. L'objectif de cette revue est d'explorer la littérature scientifique qui traite la question de la participation des citoyens aux processus décisionnels relatifs à la planification et l'implantation de programmes de télésanté. À l'issue de ce travail, nous aspirons à avoir un retour sur l'utilité et la pertinence d'une telle approche dans l'amélioration de l'acceptabilité et l'utilisation des services de télésanté.

2. Méthodes

Une stratégie de recherche a été développée à partir des mots clés pertinents aux deux groupes sémantiques d'intérêt (services de télésanté, participation citoyenne). D'autres mots clés plus spécifiques ont été aussi utilisés : En français: télésanté, télémédecine, e-santé, participation citoyenne, participation sociale, expertise citoyenne. En anglais: *telemedicine, telecare, e-health, telehealth, involving patient, citizen participation, citizen expert*. La stratégie de recherche a été

adaptée et effectuée dans les principales bases de données bibliographiques dans le domaine des services de santé (MEDLINE, EMBASE, Cochrane Library, CINAHL, CAIRN). Une stratégie similaire a été effectuée dans les principaux moteurs de recherche web. Nous avons procédé à la sélection des documents potentiellement pertinents sur la base des titres et résumés en appliquant des critères d'inclusion et d'exclusion prédéfinis. Pour être inclus, un document devait : 1) traiter de la participation de patients, citoyens ou public dans les décisions concernant les services et projets de télésanté; 2) rapporter des résultats empiriques basés sur une approche de recherche qualitative, quantitative ou de méthodes mixtes; 3) être publié en anglais ou en français. Les documents ne satisfaisant pas ces critères étaient exclus. Aucune limite temporelle ou géographique n'a été considérée pendant la recherche documentaire. En effet, compte tenu du nombre limité des études sur le sujet, nous avons voulu avoir, dans le cadre du possible, une vision d'ensemble sur les différentes expériences internationales et leur évolution dans le temps. Les titres et les résumés des études identifiées par la stratégie de recherche décrite ont été lus pour déterminer leur admissibilité potentielle à l'examen. Les articles complets ont ensuite été évalués pour leur pertinence.

La première recherche, à la lecture des titres, a permis de sélectionner quarante-trois références, après avoir éliminé les doublons issus des différentes bases de données interrogées. Puis, à l'issue de la lecture des résumés, nous avons effectué une deuxième sélection qui a permis de ne garder que 18 références. Enfin, la lecture complète des articles nous a amené à ne prendre en considération que 10 articles. Les articles exclus étaient principalement des textes de réflexion, commentaires ou analyses, ce qui ne répondait pas à nos critères d'inclusion.

La sélection finale se compose ainsi de : quatre revues systématiques (Broens & al, 2007; Neubeck & al, 2009; Ekeland & al, 2010; Ekeland & al, 2012), cinq études qualitatives (Clemensen & al, 2011; May & al, 2002; King & al, 2011; Mort & al, 2005; Yu & al, 2009) et une étude mixte (Sanders & al, 2012).

3. Résultats

Beaucoup de travaux ont insisté sur les bienfaits de la télésanté quant à la qualité des soins et services [12] et d'autres recommandent une meilleure prise en compte des attentes et des besoins des usagers (ou futurs usagers) [8, 13, 14]. Ces études soulignent que l'utilisation de tout service de télésanté dépend des motivations et des connaissances des usagers [15, 16]. Selon McLean et collaborateurs, l'un des facteurs de succès des programmes de télésanté résiderait dans une conception plus active et explicite de l'expertise des citoyens-usagers dans le processus de mise en place des services de télésanté [17]. D'autres soulignent qu'une meilleure intégration des services de télésanté dépend fortement des contextes locaux des territoires desservis [18, 19]. Ainsi, l'approfondissement du point de vue des usagers potentiels, ou actuels, mérite une plus grande attention.

Certaines études ont fait état de la réticence exprimée par les patients/usagers à l'idée d'être consultés, ou suivis, via des applications de télésanté [20]. Au Royaume-Uni, Sanders et

collaborateurs ont fait appel à un groupe de personnes ayant refusé d'utiliser des services de télésanté (ex : télésoins à domicile, téléassistance) pour essayer d'en comprendre les raisons [6]. A l'issue de cette étude, les auteurs concluent que ce refus est dû principalement aux questions de l'identité, au rapport à la technologie, à la littératie, à l'intrusion dans la vie privée ainsi que le risque de déshumanisation des services en faisant appel à cette technologie. En ce sens, les auteurs estiment qu'il est important de s'ouvrir aux attentes et de prendre en considération les opinions et les craintes des personnes concernées par la question, cela pour nourrir les étapes de conception et de mise en œuvre de ces services en vue d'en améliorer l'acceptation et l'utilisation par la suite.

Sur la base de cette idée, quelques travaux ont fait état d'expériences impliquant des citoyens dans les prises de décisions concernant les services de télésanté à travers des mécanismes de jurys et de panels de citoyens [21, 22]. Ainsi, au Royaume-Uni, King et collaborateurs ont fait appel à des jurys de citoyens pour explorer les obstacles et les facilitateurs à la mise en place des services de télésanté. Les jurys de citoyens étaient appelés à discuter et à délibérer sur un ensemble de questions spécifiques à leur santé et la place qu'occupe désormais la télésanté dans un certain nombre de spécialités médicales. L'étude a montré que les participants ont exprimé une grande envie d'être inclus dans le développement des programmes de télésanté (et de l'e-santé en général) sur leur territoire et que leurs propositions soient prises en compte par les décideurs [21].

Toujours au Royaume-Uni, Mort et collaborateurs ont recruté des délégués bénévoles pour former un panel de citoyens afin de débattre sur la question de l'intégration des nouvelles technologies dans les services de santé, notamment la télésanté [22]. Suite à ce travail, les chercheurs sont sortis avec un ensemble de recommandations. Celles-ci couvrent divers points touchants aux patients, usagers et personnels soignants, aux approches de prestation de services, ainsi qu'à la recherche et au partage des connaissances. Les résultats de cette étude montrent que les citoyens ont une grande connaissance des valeurs fondatrices du système de santé de leur pays et estiment que ces valeurs ne sont pas traduites adéquatement dans la réalité. Les participants ont montré aussi une bonne connaissance des enjeux liés à l'utilisation des nouvelles technologies dans le domaine de la santé, qu'ils ont un recul suffisant sur les questions relatives à leur santé et qu'il serait très éclairant qu'ils soient associés au développement des projets technologiques, dont la télésanté, sur leurs territoires.

Au Danemark, Clemensen et collaborateurs ont exposé le potentiel d'associer les citoyens à la mise en place de projets de télésanté sur leur territoire. À l'issue de cette étude, les auteurs ont constaté que la participation des citoyens, combinée au partage de l'information pertinente sur leur santé, permet d'augmenter leur autonomie, notamment les personnes âgées et celles vivant avec une pathologie chronique. Ils concluent ainsi que l'adoption et l'utilisation des programmes de télésanté dépend fortement de l'inclusion des citoyens dans le processus de planification et de développement des applications de télésanté qui leur sont destinées afin de mieux répondre à leurs attentes ou besoins [14].

Enfin, à Taiwan, Yu et collaborateurs ont associé des citoyens bénévoles à la mise en place et l'évaluation d'un système de télésurveillance et gestion des maladies chroniques, pour la mesure et le suivi des paramètres physiologiques des patients à distance (pression artérielle,

rythme cardiaque, etc.) [23]. Les résultats de cette étude ont montré que le fait de faire participer des citoyens à cette expérience leur a permis d'exprimer une grande motivation à l'utiliser. En effet, cette expérience était l'occasion de mieux expliquer les avantages de cette technologie et de répondre aux interrogations des participants, ce qui a amené à voir émerger des attentes positives à son égard. Pour les auteurs, cette démarche s'avère très pertinente pour continuer à vérifier l'acceptabilité et la satisfaction envers la technologie, par transition la pérennité de son utilisation.

4. Discussion

Sur un aspect plus large, l'idée de la participation des citoyens à l'élaboration des politiques publiques a longtemps fait débat et a même divisé, surtout sur la façon dont les pouvoirs publics doivent considérer ce nouvel arrivant sur le paysage décisionnel. Ainsi, la question de la légitimité de la population à participer à la définition des priorités en matière de santé a largement polarisé le paysage politique [24, 25, 26,27]. Ce qui a fait qu'elle a été souvent utilisée d'une manière informelle [28].

En télésanté, beaucoup de travaux ont été consacrés à la question de la planification et de la mise en œuvre des projets de télésanté, mais peu ont abordé la question de leur pérennité. Ce constat explique en partie pourquoi de nombreux projets peinent à survivre au-delà de l'étape du projet pilote. Cette situation a fait que les pouvoirs publics et les organismes de financement exigent la mise en place d'approches permettant d'avoir une vision d'ensemble des différents facteurs conditionnant la réussite et la survie des programmes mis en œuvre. Ces constats ont amené à s'interroger sur la place des citoyens dans le processus décisionnel de l'élaboration et l'implantation des politiques de télésanté et son impact sur l'acceptation et l'utilisation des services.

L'analyse de la littérature a montré que les citoyens réclament de plus en plus le droit à se prononcer sur les politiques et les programmes ayant un potentiel impact sur leur santé, puisque ce sont eux qui auront à vivre les conséquences de ces décisions dans leur quotidien. Pour cause, ils estiment qu'ils connaissent mieux leurs besoins et les dynamiques animant leur communauté, et que leurs connaissances viennent directement du contact quotidien qu'ils ont avec leurs réalités locales. Donc, ils peuvent apporter un regard éclairé sur la pertinence et les impacts des politiques mises en œuvre sur leur territoire. Ainsi, plusieurs auteurs estiment que cette démarche permettrait de renforcer le processus décisionnel en mettant à contribution les connaissances des individus et des groupes pour pointer les éléments de priorisation ou de modification à apporter à une politique ou un programme afin que celui-ci puisse répondre à leurs besoins [5, 6, 7, 9, 10, 11, 14, 17]. Cela rejoint ce qui a été déjà constaté dans le système de santé où on observe une présence de plus en plus croissante du patient dans le processus décisionnel en s'impliquant dans le diagnostic, le traitement et la quête d'information [29]. Cette perspective pourrait être porteuse dans le cadre des recherches sur les stratégies permettant d'améliorer l'utilisation des TIC en santé, notamment la télésanté.

D'après la littérature, les difficultés liées à l'intégration de la télésanté en routine clinique s'expliquent en grande partie par le fait qu'on s'est historiquement concentré sur les aspects techniques et cliniques de cette technologie, au risque de sous-estimer les autres aspects, notamment ceux qui touchent aux choix de l'utilisateur-patient. En effet, plusieurs auteurs estiment que les usagers-patients ont besoin de temps et davantage d'informations pour pouvoir accepter d'utiliser cette technologie. De fait, pour améliorer l'acceptabilité de la télésanté, et son utilisation par la suite, il faudrait prendre en considération les attentes et les besoins des différents acteurs concernés par la question (ex : patients, professionnels de santé, etc.) [6].

À la lumière de cette revue synthétique, et sur la base de l'ensemble des travaux analysés, la prise en considération de la perspective citoyenne revient comme une approche pertinente et prometteuse pour améliorer les conditions d'adoption et d'utilisation des programmes et services de télésanté par la population, en favorisant le partage de l'information, le recueil des besoins et la réflexion sur les possibilités d'utilisation de cette technologie. Cela en plus du fait qu'elle constituerait une opportunité d'accroître l'empowerment des citoyens-usagers. Ces conclusions sont en accord avec les résultats obtenus à l'issue de travaux qui ont essayé d'évaluer l'intérêt de la participation des citoyens à la réflexion sur des questions de santé les concernant. À titre d'exemple, au Canada, Martin et collaborateurs ont mené une étude consistant en des entrevues avec des décideurs, des professionnels de la santé et des citoyens [30]. Les auteurs ont ainsi recueilli les perceptions des différentes parties prenantes concernant l'implication des citoyens dans la définition des priorités en matière de soins et services de santé. Les résultats de cette initiative ont montré que les citoyens pourraient avoir un rôle important à jouer dans le processus décisionnel relatif à la définition des priorités du système de santé. En Australie, les travaux de Molster et collaborateurs relatent le déroulement d'une initiative publique, basée sur des forums citoyens, qui a été entreprise dans la perspective d'élaborer une politique au sujet de la mise en place et la gestion des banques biologiques (ou biobanques) [31]. Les recommandations formulées, à l'issue des délibérations, ont été traduites par des actions et directives gouvernementales. Les auteurs soulignent que cette initiative a permis de répondre aux préoccupations des différentes parties, tout en aboutissant à des décisions éclairées et consensuelles. Selon les auteurs, cette initiative aurait permis d'augmenter la légitimité et la portée des politiques qui en suivent.

Il reste que l'idée d'associer les citoyens-usagers au processus décisionnel relatif à la planification et l'implantation des services de télésanté n'en est qu'à ses débuts. En effet, nous avons constaté qu'il y a très peu de littérature ayant traité ce sujet à ce jour. De fait, davantage de travaux et d'expérimentations sont nécessaires pour mieux cerner cette démarche et arriver à produire un cadre de référence qui permettrait de guider les décideurs et les utilisateurs de connaissances dans l'élaboration de stratégies efficaces pour associer les citoyens-usagers aux démarches de planification et d'implantation des services de télésanté. Pour ce faire, des leçons pourraient être tirées des réussites issues de l'implication du public en évaluation des technologies médicales et les interventions en santé (ETMIS) déjà à l'œuvre dans plusieurs pays [32, 33].

En ce sens, il serait pertinent d'explorer encore davantage les points de vue des groupes de citoyens-usagers (associations de patients, groupes communautaires, etc.) et des décideurs (gestionnaires, professionnels de santé, décideurs politiques) au sujet de l'implication des citoyens dans les différentes phases de mise en œuvre des programmes de télésanté, les stratégies efficaces pour favoriser cette implication, ainsi que les barrières et facilitateurs perçus. Confronter les positions et les opinions des groupes d'acteurs concernés permettrait de mieux cerner les différentes perceptions au regard de la participation des citoyens-usagers à l'organisation des services de télésanté et l'impact de cette implication sur l'utilisation de ces services. Ceci permettra aussi d'identifier le (ou les) type(s) de participation, le degré d'implication et le moment à privilégier pour impliquer les citoyens dans les activités de télésanté. Mais avant d'en arriver là, le fait d'associer des citoyens au processus décisionnel relatif à l'implantation des services de télésanté sur leur territoire suggère de répondre à un certain nombre de questions qui émergent en perspective : 1) Ces populations sont-elles considérées suffisamment dans les différentes étapes de recueil des besoins et de mise en œuvre ? 2) Comment participent-elles à la mise en place et au suivi des services de télésanté sur leurs territoires ? 3) Quelle place occupent-elles dans la hiérarchie organisationnelle de ces services ? 4) Quels sont les enjeux que cela présente pour les décideurs politiques, les gestionnaires et les professionnels de santé ?

5. Conclusion :

Sur la base des données de la littérature en notre possession, nous estimons qu'il est nécessaire d'explorer, encore davantage, la question de la pertinence d'associer les citoyens-usagers à la planification et l'implantation des programmes de télésanté sur leurs territoires, et les retombées de celle-ci sur l'utilisation des services. En effet, l'état actuel des connaissances est encore loin de nous permettre d'avoir une vision complète des conditions pouvant favoriser l'adoption et l'utilisation de la télésanté. Le fait de s'inspirer de l'idée de la participation des citoyens, comme démarche d'implication accrue des utilisateurs de services de santé dans les prises de décisions en santé, pourrait constituer une voie prometteuse afin d'offrir des services de télésanté davantage centrés sur les besoins des patients et adaptés à leurs contextes.

Remerciements

Nous tenons à remercier, tout particulièrement, le Programme stratégique de formation des Instituts de recherche en santé du Canada (IRSC) et du Réseau de recherche en santé des populations du Québec (RRSPQ) en recherche transdisciplinaire sur les interventions en santé publique: Promotion, Prévention et Politiques Publiques (4P) pour sa contribution au financement du projet de doctorat de l'un des auteurs. Nos remerciements vont aussi à toutes les personnes et les institutions qui ont contribué directement ou indirectement à la réalisation de ce travail soumis aux Journées francophones d'informatique médicale 2014.

Références

- [1] Lombrail P. Inégalités de santé: le temps de l'action! *Santé Publique*. 2011;23(3):167-8.
- [2] Gauthier J, Haggerty J, Lamarche P, Entre adaptabilité et fragilité: les conditions d'accès aux services de santé pour les communautés rurales et éloignées. *Rapport de Recherche, Institut National de Santé Publique du Québec*, Québec, Canada; 2009. p34. Projet No RC2-0985-05.
- [3] Ministère de la santé et des services sociaux du Québec : «*Vision, orientations et stratégies de développement de la télésanté au Québec*». Document d'orientation officielle du MSSS, Québec, mars 2001.
- 4- Zanaboni P, Wootton R. Adoption of telemedicine: from pilot stage to routine delivery. *BMC medical informatics and decision making*. 2012;12(1):1.
- [5] Gagnon MP, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *Journal of medical systems*. 2012;36(1):241-77.
- [6] Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. *BMC health services research*. 2012;12(1):220.
- [7] Abbott P, Liu Y. A Scoping Review of Telehealth. *Yearbook of medical informatics*. 2013;8(1):51.
- [8] Levasseur M, Richard L, Gauvin L, Raymond É. Inventory and analysis of definitions of social participation found in the aging literature: Proposed taxonomy of social activities. *Social science & medicine*. 2010;71(12):2141-9.

- [9] Greig S, Parry N, Rimmington B. Promoting sustainable regeneration: learning from a case study in participatory HIA. *Environmental Impact Assessment Review*. 2004;24(2):255-67.
- [10] Joffe M. How do we make health impact assessment fit for purpose? *Public Health*. 2003;117(5):301-4.
- [11] Wallerstein N, Duran B. Community-based participatory research contributions to intervention research: the intersection of science and practice to improve health equity. *American Journal of Public Health*. 2010;100.
- [12] Neubeck L, Redfern J, Fernandez R, Briffa T, Bauman A, Freedman SB. Telehealth interventions for the secondary prevention of coronary heart disease: a systematic review. *European Journal of Cardiovascular Prevention & Rehabilitation*. 2009;16(3):281-9.
- [13] Ekeland AG, Bowes A, Flottorp S. Effectiveness of telemedicine: a systematic review of reviews. *International journal of medical informatics*. 2010;79(11):736-71.
- [14] Clemensen J, Rasmussen J, Denning A, Craggs MA, editors. Patient Empowerment and New Citizen Roles through Telehealth Technologies-The Early Stage. eTELEMED 2011, *The Third International Conference on eHealth, Telemedicine, and Social Medicine*; 2011. p. 114-119.
- [15] Ekeland AG, Bowes A, Flottorp S. Methodologies for assessing telemedicine: a systematic review of reviews. *International journal of medical informatics*. 2012;81(1):1-11.
- [16] Broens TH, Vollenbroek-Hutten MM, Hermens HJ, van Halteren AT, Nieuwenhuis LJ. Determinants of successful telemedicine implementations: a literature study. *Journal of telemedicine and telecare*. 2007;13(6):303-9.
- [17] McLean S, Protti D, Sheikh A. Telehealthcare for long term conditions. *Bmj*. 2011;342:d120.
- [18] Ohinmaa A, Hailey D, Roine R. Elements for assessment of telemedicine applications. *International Journal of Technology Assessment in Health Care*. 2001;17(02):190-202.
- [19] May CR, Williams T, Mair F, Mort M, Shaw N, Gask L. Factors influencing the evaluation of telehealth interventions: preliminary results from a qualitative study of evaluation projects in the UK. *Journal of Telemedicine & Telecare*. 2002;8.

- [20] Miller EA. The technical and interpersonal aspects of telemedicine: effects on doctor-patient communication. *Journal of telemedicine and telecare*. 2003;9(1):1-7.
- [21] King G, Heaney DJ, Boddy D, O'Donnell CA, Clark JS, Mair FS. Exploring public perspectives on e-health: findings from two citizen juries. *Health Expectations*. 2011;14(4):351-60.
- [22] Mort M, Finch T. Principles for telemedicine and telecare: the perspective of a citizens' panel. *Journal of telemedicine and telecare*. 2005;11(suppl 1):66-8.
- [23] Yu C, Yang JJ, Chen JC, Liu CS, Chen CC, Lin ML, et al., editors. The development and evaluation of the Citizen Telehealth Care service System: case study in Taipei. Engineering in Medicine and Biology Society, 2009. *Annual International Conference of the IEEE; 2009: IEEE*.
- [24] Wiseman V, Mooney G, Berry G, Tang KC. Involving the general public in priority setting: experiences from Australia. *Social science & medicine*. 2003;56(5):1001-12.
- [25] Church J, Saunders D, Wanke M, Pong R, Spooner C, Dorgan M. Citizen participation in health decision-making: past experience and future prospects. *Journal of Public Health Policy*. 2002;23(1):12-32.
- [26] Thompson J, Bissell P, Cooper CL, Armitage CJ, Barber R. Exploring the Impact of Patient and Public Involvement in a Cancer Research Setting. *Qualitative health research*. 2014;24(1):46-54.
- [27] Carman KL, Dardess P, Maurer M, Sofaer S, Adams K, Bechtel C, et al. Patient and family engagement: a framework for understanding the elements and developing interventions and policies. *Health Affairs*. 2013;32(2):223-31.
- [28] Green C, Gerard K. Exploring the social value of health-care interventions: a stated preference discrete choice experiment. *Health economics*. 2009;18(8):951-76.
- [29] Street Jr R, L, Gordon H S, Ward M. M, Krupat E, Kravitz RL. Patient participation in medical consultations: why some patients are more involved than others. *Medical care*, 2005, vol. 43, no 10, p. 960-969.
- [30] Martin D, Abelson J, Singer P. Participation in health care priority-setting through the eyes of the participants. *Journal of Health Services Research & Policy*. 2002;7(4):222-9.

[31] Molster C, Maxwell S, Youngs L, Potts A, Kyne G, Hope F, et al. An Australian approach to the policy translation of deliberated citizen perspectives on biobanking. *Public health genomics*. 2011;15(2):82-91.

[32] Gagnon MP, Desmartis M, Lepage-Savary D, Gagnon J, St-Pierre M, Rhainds M, et al. Introducing patients' and the public's perspectives to health technology assessment: A systematic review of international experiences. *International journal of technology assessment in health care*. 2011;27(01):31-42.

[33] Royle J, Oliver S. Consumer involvement in the health technology assessment program. *International journal of technology assessment in health care*. 2004;20(04):493-7.

Adresse de correspondance :

Hassane ALAMI

Axe santé publique et pratiques optimales en santé

CRCHUQ, Hôpital Saint-François D'Assise, Édifice D

45, rue Leclerc, porte D6-724 Québec, (QC), Canada, G1L 2G1.

Téléphone : 418 525-4444, Poste 53530

Fax : 418-525-4194

Courriel : hassane.alami.1@ulaval.ca

Recommandations de pratique clinique sous la forme d'un document textuel et d'un système d'aide à la décision : une expérience de production et diffusion simultanées des deux types de format pour la prise en charge du diabète de type 2

Clinical practice guidelines as a textual document and a clinical decision support system: simultaneous development and publication of both formats in the management of type 2 diabetes

Jacques Bouaud^{1,2,3,4} Hector Falcoff⁵ Valérie Ertel-Pau⁶ Véronique Raimond⁶ Brigitte Séroussi^{2,3,4,7}

¹AP-HP, DRCD, Paris, France

²Sorbonne Universités, UPMC Université Paris 06, UMR_S 1142, LIMICS, Paris, France

³INSERM, U1142, LIMICS, Paris, France

⁴Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), Bobigny, France

⁵Paris Descartes, Sorbonne Paris Cité, Département de médecine générale, Paris, France ; SFTG, Paris, France

⁶Haute Autorité de Santé, Saint-Denis-La-Plaine, France

⁷AP-HP, Hôpital Tenon, DSP, Paris, France ; APREC, Paris, France

Résumé

Les systèmes d'aide à la décision clinique (SADM) ont le potentiel d'améliorer la mise en œuvre des recommandations de pratique clinique (RPC) et de favoriser ainsi la qualité des soins. Cependant, les RPC sont difficiles à formaliser car la traduction de textes souvent incomplets et ambigus est une tâche complexe. Par ailleurs, la modélisation du contenu des RPC est habituellement postérieure à l'élaboration des RPC, sans qu'aucune relation entre les auteurs des RPC et les modélisateurs de leur contenu ne soit organisée. En France, en parallèle du développement habituel, à l'initiative de la Haute Autorité de Santé (HAS), des RPC sur la prise en charge thérapeutique du diabète de type 2, nous avons développé un homologue formalisé sous la forme du SADM RecosDoc-diabète. Le système permet de construire de façon interactive un profil centré-patient et d'obtenir les recommandations appropriées. Les RPC textuelles et RecosDoc-diabète ont été publiés et mis en ligne en même temps, en février 2013. Entre le 12 février et le 31 décembre 2013, 55 203 visiteurs ont accédé aux RPC textuelles alors que 10 565 ont utilisé le système. 186 visiteurs du système (2 %) ont répondu au questionnaire proposé. Le système a été globalement positivement évalué, même si les appréciations, contrastées, ont montré que les besoins des utilisateurs sont variés et que « la taille unique » n'est pas la solution en termes de SADM.

Abstract

Clinical decision support systems (CDSSs) are expected to promote the implementation of clinical practice guidelines (CPGs) and improve the quality of care. However, originally produced as texts often incomplete and ambiguous, CPGs are difficult to formalize. In addition, CPG modellers work a posteriori on textual CPGs with no connexion with CPG authors. In conjunction with the development of type 2 diabetes CPGs, conducted in France by the HAS (Haute autorité de Santé), we have simultaneously developed a formal counterpart of the text as a CDSS named RecosDoc-diabète. The system allows to interactively build a patient-centered pathway, and get the appropriate recommendations. Textual CPGs and RecosDoc-diabète were published together in February 2013 and made available online. Between February 12th and December 31st, 2013, 55,203 visitors accessed the narrative CPGs whereas 10,565 accessed the system. 186 visitors of RecosDoc-diabète (2%) responded to the survey provided, evidencing that the system was globally positively evaluated, although assessments were mixed illustrating that users' needs are varied and that "one size does not fit all" in terms of CDSSs.

Mots-clés : Recommandations de pratique clinique, Publication de recommandations de pratique clinique, Système d'aide à la décision clinique, Diabète de type 2.

Keywords: Practice Guidelines as Topic, Guideline publication, Clinical Decision Support Systems, Type 2 Diabetes Mellitus.

1 Introduction

Aujourd'hui, en France, plus de 3,5 millions personnes sont atteintes de diabète, soit près de 5 % de la population, ce qui fait de cette maladie un problème majeur de santé publique et l'objet d'une recherche pharmaceutique accélérée. De nouvelles molécules anti-diabétiques ont ainsi été mises sur le marché au cours des dix dernières années. Si elles sont efficaces, on manque de recul pour en connaître tous les effets à long terme, contrairement aux "anciennes" molécules dont les bénéfices, et les effets secondaires, sont bien connus. Dans ce contexte, la Haute autorité de santé (HAS), agence nationale publique française responsable de la qualité et de la sécurité des soins, et à ce titre, de l'évaluation et de l'amélioration des pratiques professionnelles, a décidé en 2011 de mettre à jour ses recommandations de pratique clinique (RPC) pour la prise en charge thérapeutique du diabète de type 2, dont la dernière version datait de 2006.

Les RPC sont classiquement développées par des sociétés savantes ou des agences nationales, comme la HAS en France. L'objectif est de promouvoir les bonnes pratiques auprès des professionnels de santé afin d'améliorer la qualité des soins et de réduire les coûts [1]. Les RPC sont élaborées sur la base d'une méthodologie standard s'appuyant sur une revue de la littérature scientifique [2]. Elles constituent ainsi une synthèse de l'état des connaissances scientifiques sur une problématique médicale donnée. Pour un ensemble de situations cliniques répertoriées, elles fournissent des recommandations de prise en charge avec un grade, noté A, B, ou C, fonction de la qualité des études qui ont été à l'origine des recommandations. Néanmoins, de nombreuses propositions thérapeutiques figurant dans les RPC ne reposent pas sur le résultat d'études scientifiques. Dans ce cas, il n'y a pas de "preuve", et la proposition thérapeutique, issue d'un consensus, repose sur un accord professionnel (AP), parfois noté grade D. Les RPC se présentent classiquement dans un format textuel narratif.

Des études ont montré que la diffusion des RPC dans leur format textuel d'origine n'avait qu'un impact limité sur les pratiques des professionnels de santé [3]. Au contraire, les systèmes informati-

des d'aide à la décision médicale (SADM) pourraient être efficaces pour favoriser la mise en œuvre effective des RPC en fournissant des recommandations adaptées aux caractéristiques particulières des patients [4–6]. Mais la construction des SADM suppose une étape préalable de traduction des RPC sous la forme de bases de connaissances structurées, ce qui reste une étape complexe. Par ailleurs, cette étape est toujours postérieure à celle de l'élaboration des RPC et sans relation avec cette dernière, ces deux étapes n'impliquant pas les mêmes acteurs, et sans collaboration envisagée.

En collaboration avec la HAS, et à titre d'expérimentation, nous avons proposé de développer une version formalisée et structurée des RPC au fur et à mesure de leur élaboration sous la forme d'un système d'aide à la décision [7]. Ce travail a été réalisé sur les nouvelles recommandations françaises sur la « Stratégie médicamenteuse du contrôle glycémique du diabète de type 2 » [8]. L'objectif de cet article est d'une part, de décrire la construction, en interaction avec le processus de rédaction, de la version structurée des RPC, base de connaissances du système d'aide à la décision RecosDoc-diabète et, d'autre part, de présenter les résultats obtenus lors de la diffusion simultanée sur le site web de la HAS des RPC textuelles et du système RecosDoc-diabète.

2 État de l'art

Les RPC sont largement diffusées dans un format textuel, en dépit du faible impact de ce mode de diffusion. Les RPC sont ainsi accessibles sur des portails internet comme celui de la *National Guideline Clearinghouse* (NGC), de l'*Agency for Healthcare Research and Quality* (AHRQ) aux États-Unis, de la *Canadian Medical Association infobase* au Canada, du *National Institute for Health and Care Excellence* (NICE) au Royaume-Uni, et de la HAS en France. En raison de récentes réglementations dans la plupart des pays industrialisés imposant de garantir la sécurité des patients et d'assurer un haut niveau de qualité des soins, des entreprises privées commerciales ont également investi le domaine de la diffusion des connaissances médicales, dont les RPC, et proposent des solutions pour aider les professionnels de santé dans leur prise de décisions (UpToDate®, MD Consult, Vidal, etc).

De nombreux obstacles s'opposent néanmoins à la mise en œuvre des RPC sur le terrain [9]. Les RPC sont en effet souvent critiquées, jugées trop abstraites, ambiguës, toujours incomplètes, en retard sur l'état de l'art du fait des délais de production, ou encore trop simples ou trop complexes pour pouvoir être appliquées à des patients réels. Cependant, leur variété et leur grand nombre ne facilitent pas la recherche d'une réponse à la question posée par le patient. Il est en effet, tout d'abord, difficile de trouver le « bon » guide de RPC, c.-à-d. celui qui est susceptible de répondre à la question posée par un patient donné. Ainsi, lorsque que l'on interroge les bases du NGC, CMA infobase, NICE ou MD Consult avec le mot-clé *diabetes*, on obtient respectivement 609, 71, 821 et 86 retours. Par ailleurs, quand on trouve le bon guide de RPC, il n'est pas évident de trouver la « bonne » page ou le « bon » paragraphe, car la forme textuelle ne permet pas d'opérationnaliser facilement le contenu des RPC pour un patient donné, et de guider de façon efficace le médecin dans sa prise de décision. On notera néanmoins une évolution récente du format textuel adoptée par le NICE qui fournit une version semi-structurée de ses recommandations dans une forme algorithmique consultable en ligne¹ proche des scénarios du système Prodigy et des CKS² implémentés avec EON [10].

Si les SADM peuvent être des outils efficaces pour améliorer la qualité des décisions et favoriser le suivi des RPC [6], leur développement reste un processus difficile. Il suppose une modélisation formelle des connaissances contenues dans les textes des RPC, ce qui nécessite une interprétation du

1. <http://pathways.nice.org.uk/pathways/diabetes>

2. <http://cks.nice.org.uk>

texte qui peut être sujette à variabilité [11]. Des outils et des méthodes ont été proposés pour faciliter la formalisation des RPC et leur traduction dans un format informatisable [12, 13]. Ainsi, Shiffman et al. [12] ont proposé une série d'étapes successives permettant de modéliser les RPC. L'étape d'atomisation permet d'identifier les notions pertinentes qui seront les critères décisionnels ou les actions recommandées. L'étape de désabstraction permet de reformuler les notions identifiées dans une interprétation opérationnalisable, par exemple la quantification de variables qualitatives, ou la définition de seuils précis. L'étape de désambiguïsation vise à éliminer toute ambiguïté des recommandations textuelles. Elle repose sur un choix d'interprétation du texte, a priori arbitraire (remplacement des formes conditionnelles par des formes indicatives). Enfin, une étape de vérification de la complétude permet d'envisager des situations cliniques qui ne sont pas explicitement mentionnées dans les RPC.

Par ailleurs, du fait du mode de production actuel des RPC textuelles, le processus de modélisation des contenus est nécessairement postérieur à la publication des RPC, provoquant un délai entre la sortie des RPC et du SADM correspondant. Il en résulte que les ambiguïtés et l'incomplétude du texte des RPC ne peuvent pas être résolues de façon satisfaisante lors de la modélisation des connaissances puisque les auteurs des RPC ne sont plus disponibles en tant que groupe de travail pour répondre aux questions que se posent les développeurs des SADM. En particulier, la vérification de la complétude peut faire apparaître des situations cliniques non explicites dans les RPC textuelles pour lesquelles il est nécessaire d'avoir un accord professionnel difficile à établir a posteriori. On a alors recours à des « experts du domaine » pour lever ces difficultés mais les solutions proposées perdent leur statut d'AP pour prendre celui d'« avis d'experts » ce qui n'est pas en adéquation avec la philosophie de l'*evidence-base medicine*, ni avec le paradigme des RPC. Ainsi, les formats structurés sous la forme de bases de connaissances des RPC devraient être construits en parallèle des versions textuelles dans un processus collaboratif où la levée des ambiguïtés révélées par la formalisation améliore les bases de connaissances et la qualité des SADM [14, 15], tout en permettant en retour de clarifier le texte des RPC. C'est cette expérience qui a été conduite en collaboration avec la HAS sur les RPC pour la prise en charge thérapeutique du diabète de type 2.

3 Matériel et méthodes

3.1 Paradigme de modélisation des connaissances

Lors de travaux antérieurs, nous avons proposé un paradigme documentaire de l'aide à la décision où les connaissances sont formalisées sous la forme d'un arbre décisionnel [16]. Les nœuds correspondent à des critères décisionnels C_i , les chemins dans l'arbre à des profils cliniques de patients auxquels correspondent au niveau des feuilles les actions recommandées R_j . L'expansion de l'arbre génère un jeu de règles de production classiques sur le modèle SI C_1 ET C_2 ET ... ET C_n ALORS R_1 OU R_2 OU ... OU R_p . Si la base de connaissances peut-être exécutée automatiquement à partir de données patient codées, un utilisateur peut également parcourir l'arbre décisionnel lors d'une « navigation » interactive où il répond à des questions, ce qui lui permet de caractériser le profil du patient à résoudre. De cette manière, il va sélectionner un chemin de l'arbre décisionnel et obtenir les recommandations adaptées. Le respect lors de la modélisation de contraintes d'exhaustivité et d'exclusivité mutuelle des valeurs des critères décisionnels permet de garantir que, pour une situation clinique donnée, il existe un unique parcours dans l'arbre. Cette approche a été utilisée pour modéliser la prise en charge thérapeutique du cancer du sein (OncoDoc), de la prise en charge du risque cardio-vasculaire, hypertension artérielle, dyslipidémie, sevrage tabagique (ASTI), ou encore la démarche diagnostique des pneumopathies médicamenteuses (PneumoDoc).

Dans le contexte de la prise en charge des patients diabétiques, la figure 1 montre un sous-arbre de l'arbre décisionnel, qui décrit la stratégie thérapeutique recommandée pour les patients traités par metformine, mais dont l'HbA1c n'est pas normalisée, c.-à-d. n'est pas à l'objectif thérapeutique cible. En effet, en dehors de contre-indications et de cas particuliers, la metformine est le traitement de première intention dans la prise en charge du diabète de type 2. Ensuite, si l'objectif cible d'HbA1c n'est pas atteint, l'ajout de sulfamides, puis d'insuline est recommandé. Si les sulfamides sont contre-indiqués, il est recommandé d'associer à la metformine un inhibiteur de alpha-glucosidases ou un inhibiteur de la DPP-4 lorsque l'HbA1c est proche de l'objectif cible, et l'insuline ou les analogues du GLP-1 lorsque l'HbA1c est loin de l'objectif cible. L'ordre des propositions est déterminé par des critères faisant intervenir une balance bénéfico-risque basée sur l'efficacité thérapeutique et le manque de recul pour les nouvelles molécules.

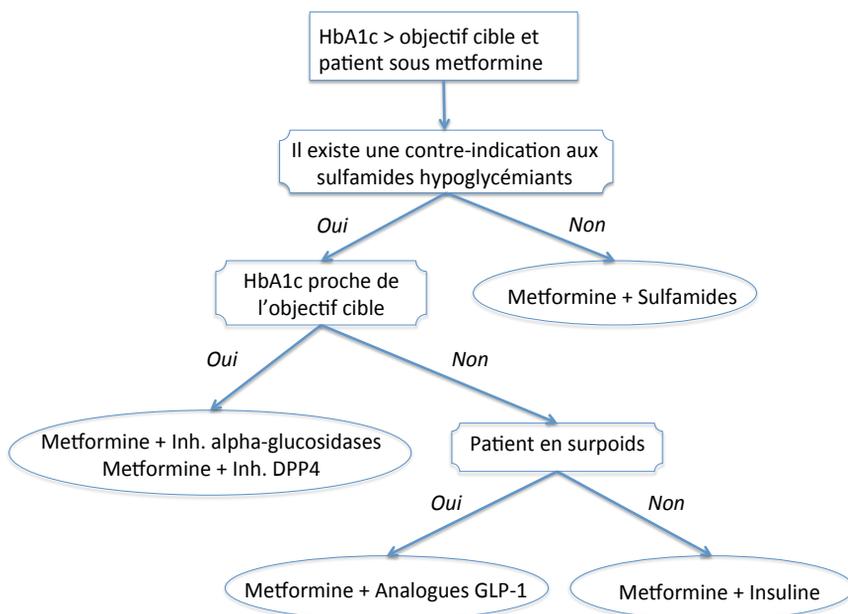


Figure 1 : Stratégie thérapeutique de prise en charge des patients dont l'HbA1c n'est pas normalisée par une monothérapie par metformine.

3.2 Développement synchrone des versions textuelle et structurée des RPC

Une des missions de la HAS consiste à assurer le développement des RPC nationales. Pour chaque RPC, il y a trois coordinateurs, un chef de projet de la HAS et deux chefs de projet médicaux. Les RPC sont élaborées par un groupe de travail (GT) représentant les différentes parties concernées par le thème des RPC. En général, il est prévu 5 à 6 réunions du GT. La première version des RPC est

rédigée par les 3 coordinateurs sur la base d'une revue de la littérature. Cette première version est discutée lors de la première réunion du GT. Le document est ensuite progressivement révisé par les 3 coordinateurs entre les réunions du GT, et les recommandations proprement dites sont élaborées lors des réunions successives du GT. Le document est ensuite transmis à un groupe de lecture plus large qui émet des remarques sur le contenu du texte. Le document final tient compte des remarques et doit ensuite être validé par les instances scientifiques de la HAS.

Le travail de formalisation des RPC s'est inséré dans cette organisation sans pour autant perturber les réunions du GT. Les membres du GT étaient néanmoins au courant du travail expérimental qui était conduit en parallèle. Le principe avait été présenté lors de la première réunion du GT et le système élaboré, dénommé RecosDoc-diabète a été présenté lors de la dernière réunion du GT. La méthode a consisté à travailler avec l'un des deux chefs de projet médicaux entre deux réunions successives du GT. Le travail de modélisation était conduit sur la version des RPC textuelles issue de la réunion précédente du GT et les problèmes d'ambiguïté, d'abstraction ou d'incomplétude étaient transmis par le chef de projet médical lors de la réunion suivante du GT. Ainsi, la formalisation des RPC, faite au fur et à mesure de la rédaction de la version textuelle des RPC, a permis d'identifier au plus tôt les faiblesses du texte et la résolution de ces difficultés a permis d'améliorer la lisibilité du texte et la qualité de la formalisation des RPC.

Pour la modélisation, nous avons utilisé le cadre proposé par Shiffman et al [12]. pour représenter les connaissances contenues dans les RPC sous la forme d'un arbre décisionnel : (i) l'atomisation pour identifier les critères patients nécessaires (p. ex. HbA1c, mode de vie erratique, symptômes d'hypoglycémie, etc.) ainsi que les actions thérapeutiques possibles (p. ex. metformine, sulfamides, mesures hygiéno-diététiques, etc.), (ii) la désabstraction pour quantifier des notions qualitatives, (iii) la désambiguïtation pour adopter un point de vue déterministe dans la formalisation (p. ex. interpréter « considérer les sulfamides » par « recommander les sulfamides ») et (iv) la vérification de la complétude (c.-à-d. identifier les situations cliniques implicites dans le texte des RPC, les rendre explicites et leur associer des prises en charge thérapeutiques). Cette dernière étape met en évidence l'existence, connue, de trous de connaissances dans les RPC. Ainsi, les RPC sont constituées de recommandations *evidence-based* sur la base d'un niveau de preuve, et de propositions thérapeutiques issues d'accords professionnels au sein du GT (AP), mais il existe une troisième catégorie de situations cliniques pour lesquelles les RPC ne proposent aucune solution. Elle sont identifiées par l'étape de vérification de la complétude. La question de les objectiver dans un système d'aide à la décision est non résolue : faut-il proposer un avis d'experts ou laisser ces situations cliniques sans solution ? Nous pensons que les professionnels de santé préfèrent avoir une solution pour toutes les situations cliniques qu'ils rencontrent dans leur pratique quotidienne, à condition que ce qui relève de l'*evidence-based medicine* soit bien différencié de ce qui relève de l'accord professionnel, et de l'avis d'experts, leur liberté de choix étant clairement maintenue pour les options thérapeutiques issues de l'avis d'experts.

3.3 Diffusion synchrone des versions textuelle et structurée des RPC

À l'issue de l'étape de modélisation des RPC textuelles, nous avons développé un système d'aide à la décision, le système RecosDoc-diabète, permettant d'opérationnaliser les RPC textuelles. Il était important que les RPC textuelles et le système RecosDoc-diabète soient publiés le même jour sur le portail de la HAS.

Un sondage anonyme sous la forme d'un questionnaire en ligne a été proposé afin de recueillir la per-

ception des internautes sur le système RecosDoc-diabète, ainsi que sur la démarche. Cinq questions portaient sur le contexte d'utilisation du système lors de leur visite, la facilité d'utilisation du système, l'utilité d'une version structurée des RPC, leur intention d'utiliser le système, l'apport d'une version structurée pour mieux comprendre les recommandations. Un champ de commentaires libres était disponible pour suggérer des pistes d'amélioration. Les répondants étaient invités à caractériser leur profil (généraliste, diabétologue, autre spécialiste, autre professionnel de santé, ou même patient diabétique) et pour les professionnels à préciser leur mode d'exercice (établissement de santé, cabinet individuel, cabinet médical de groupe, centre de santé, maison de santé) et leur nombre d'années de pratique.

Les visites du site ont été enregistrées par le biais des logs des serveurs web et les résultats ont été analysés avec awstats v0.7. Les réponses au questionnaire ont été stockées sur le serveur et exportées vers MS Excel pour être exploitées. Les commentaires libres ont été analysés par les auteurs.

4 Résultats

4.1 Première modélisation des RPC

L'étape d'atomisation a permis d'identifier les atomes de connaissance (variables de décision et d'action). L'étape de désabstraction a pu être réalisée pour certains critères, p. ex. « Surpoids » explicité en « Indice de masse corporelle > 30 », « HbA1c proche de la cible » explicité en « HbA1c à moins de 1 % de l'objectif cible ». Cependant, certaines notions sont restées sous-spécifiées dans le texte comme dans la base de connaissances, p. ex. « Hypoglycémies nocturnes préoccupantes », « Prise de poids sous insuline ». Pour l'étape de désambiguïsation, une attention particulière a été portée à la verbalisation des recommandations. Une formule non ambiguë, comme « Il est recommandé de ... », a été utilisée pour toutes les recommandations, rendant le texte plus explicite. Toutes les recommandations figurant dans le texte des RPC ont été intégrées avec leur grade dans la modélisation. La plupart des recommandations du texte des RPC concernaient des situations cliniques pour lesquelles des prises en charge étaient proposées. La formalisation a permis leur ré-écriture sous la forme de règles de décision SI-ALORS. La partie alternative (SINON), systématiquement manquante dans les premières versions du texte a été ajoutée par l'étape de vérification de la complétude.

À l'issue de la modélisation du texte des RPC, on comptait 805 situations cliniques pour lesquelles une proposition thérapeutique était fournie, 89 critères décisionnels, et une profondeur des chemins de l'arbre de décision allant de 2 à 17 nœuds (10 en moyenne). Il y avait 2 recommandations *evidence-based*, « avec niveau de preuve », 37 recommandations avec « accord professionnel » issues d'un consensus au sein du GT, et 766 situations cliniques, résultant de l'étape de vérification de la complétude, avec « avis d'experts ». Ainsi, dans la base de connaissances, les situations cliniques conduisant à des recommandations explicitement mentionnées dans les RPC représentaient moins de 5 %.

4.2 Seconde modélisation des RPC

La présence dans le système RecosDoc-diabète de propositions correspondant à des « avis d'experts » non validés par le GT a posé un problème à la HAS. Il a donc été décidé de construire une nouvelle base de connaissances, sans appliquer l'étape de vérification de la complétude lors de la modélisation, afin de se limiter aux recommandations explicitement contenues dans le texte.

La seconde modélisation s'appuie sur 30 critères décisionnels. Elle décrit 88 situations cliniques, parmi lesquelles 2 sont gradées « avec niveau de preuve » et 77 avec « accord professionnel ». Il est à noter qu'en se restreignant au texte des RPC, des situations cliniques non opératoires, pour lesquelles la recommandation ne correspond pas à une prescription de soins ont été ajoutées, ce qui explique la différence entre les 77 et 37 recommandations avec accord professionnel. La figure 2 montre une copie d'écran de RecosDoc-diabète (où AE représente l'« accord professionnel » au sein du GT).

The screenshot displays the RecosDoc-diabète website interface. At the top, the title is "RecosDoc-diabète [RBP HAS - 2013]". The main header includes the RecosDoc Diabète logo and the text "Accès structuré aux Recommandations de Bonne Pratique Stratégie médicamenteuse du contrôle glycémique du diabète de type 2". A navigation menu contains "Accueil", "Objectif", "Stratégie", "Questionnaire", and "Contact". The main content area is titled "Stratégie médicamenteuse du contrôle glycémique du diabète de type 2 (v1.2)".

Tableau clinique récapitulatif

1. Femme enceinte ou envisageant de l'être = Non
2. Patient âgé = Non
3. Profil clinique du patient = Patient sans comorbidité connue
4. Patient actuellement sous traitement médicamenteux = Non
5. Symptômes cliniques ou diabète très déséquilibré = Oui

Propositions de prise en charge

- Mesures hygiéno-diététiques + Metformine + Sulfamide hypoglycémiant.
- Mesures hygiéno-diététiques + Metformine + Schéma insulinique intensifié.
- Mesures hygiéno-diététiques + Schéma insulinique intensifié.

Recommandation 11
AE En l'absence de signes cliniques (syndrome polyuro-polydipsique, amaigrissement), le traitement est débuté par une monothérapie.
 En cas de symptômes ou de diabète très déséquilibré avec des glycémies répétées supérieures à 9g/l ou un taux d'HbA1c supérieur à 10 %, une bithérapie voire une insulinothérapie peuvent être instaurées d'emblée. Dans cette situation, lors de la réévaluation du traitement, si le diabète est bien contrôlé, on pourra être amené à passer d'une bithérapie à une monothérapie voire à l'arrêt du traitement, ou d'une insulinothérapie à un traitement oral.

Recommandation 26
AE En cas de diabète très déséquilibré, avec des glycémies supérieures à 3 g/l répétées et/ou une HbA1c > 10 %, un schéma insulinique intensifié pourra être instauré d'emblée après avis d'un endocrinologue.

Figure 2 : Copie d'écran de RecosDoc-diabète.

4.3 Publication des deux formats de diffusion des RPC

Les deux productions dans les deux formats de diffusion ont été présentées lors d'une conférence de presse officielle organisée par la HAS le 13 février 2013. Au cours de la conférence de presse, une démonstration du système RecosDoc-diabète a été faite par le chef de projet médical avec lequel nous avons travaillé³.

Le portail de la HAS a été mis à jour avec une page présentant les nouvelles RPC et les liens vers les différents documents d'accompagnement (texte des recommandations, argumentaires, vidéo de présentation, etc.). Un paragraphe décrivant l'expérimentation RecosDos-diabète a été inclus avec un lien vers le système. Sur la page d'accueil du système, des avertissements indiquaient aux utilisateurs comment utiliser le système, fournissaient un lien vers un mode d'emploi, et en particulier rappelaient que les profils cliniques manquants dans les RPC (trous de connaissances) manquaient également dans la base de connaissances de RecosDoc-diabète. L'accès au questionnaire en ligne était également proposé.

3. RecosDoc-diabète est accessible à cette adresse : <http://ics.upmc.fr/RecosDoc/diabete/>.

4.4 Statistiques de fréquentation

La fréquence des accès à la page des RPC textuelles sur le site de la HAS et au système RecosDoc-diabète a été observée. La figure 3 donne la distribution mensuelle des visites sur les 2 sites pour la période allant de février à décembre 2013. Sur cette période, les 10 565 visites du système représentent 19 % des 55 203 visites de la page HAS des RPC textuelles. Les visites du système représentent 37 % des visites des RPC HAS en début de période, et près de 10 % à la fin de l'année 2013. La majeure partie des connexions au système (68 %) provenaient directement du site de la HAS. Les autres provenaient de sites d'information à destination des professionnels de santé relayant la publication des nouvelles RPC de la HAS sur le diabète de type 2, et de manière très minoritaire des moteurs de recherche.

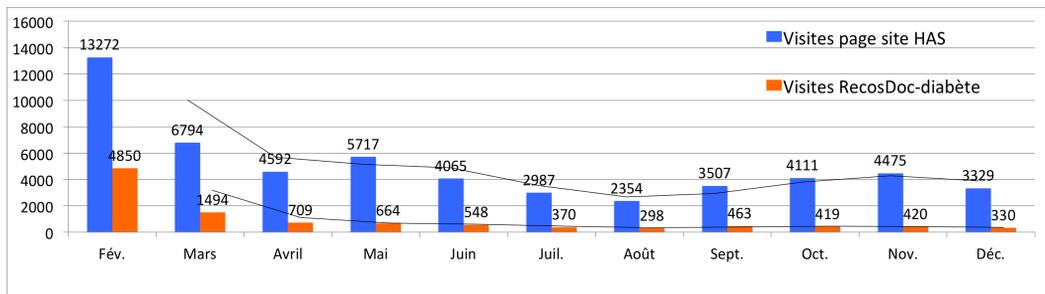


Figure 3 : Distribution des visites mensuelles de la page des RPC textuelles sur le site de la HAS et du système RecosDoc-diabète.

4.5 Évaluation du système RecosDoc-diabète

Au total, 186 visiteurs (2 % des visites du site du système RecosDoc-diabète) ont répondu au questionnaire en ligne. Le tableau 1 donne les caractéristiques des répondants en termes de profession, nombre d'années de pratique, et type d'exercice. On notera que la majorité des répondants étaient des professionnels de santé, parmi lesquels les médecins généralistes étaient majoritaires (64 %). Quelques patients diabétiques ont également utilisé le système. Un tiers des répondants travaille en cabinet individuel et presque la moitié a plus de 20 ans de pratique professionnelle.

Tableau 1 : Profils des répondants au questionnaire en ligne.

Profession (n=184)	Années de pratique (n=179)		Type d'exercice (n=178)		
	(%)	(%)	(%)	(%)	
Généraliste	64 %	< 5 ans	21 %	Établissement de santé	21 %
Diabétologue	6 %	5-10 ans	13 %	Cabinet médical individuel	30 %
Autre spécialiste	11 %	10-20 ans	18 %	Cabinet médical de groupe	26 %
Autre prof. de santé	12 %	> 20 ans	47 %	Centre de santé	4 %
Patient diabétique	2 %	—	—	Maison/pôle de santé	3 %
Autre	5 %	—	—	Autre	15 %

Les réponses aux différentes questions du questionnaire figurent dans le tableau 2. La plupart des répondants (55 %) déclarent avoir utilisé RecosDoc-diabète pour prendre connaissance des recommandations. Le système a été utilisé pour résoudre le cas d'un patient dans 21 % des cas (9 % en synchrone, en consultation, et 12 % en asynchrone, comme un cas clinique). La facilité d'utilisation du système est jugée positivement par 87 % des répondants et 88 % jugent qu'un arbre informatisé comme celui proposé par RecosDoc-diabète serait utile dans la pratique quotidienne. Ils jugent également positivement (86 %) cette structuration du contenu pour mieux comprendre la version textuelle des RPC. Concernant leurs intentions, la plupart des répondants (93 %) déclarent qu'ils pensent utiliser RecosDoc-diabète pour leurs patients parfois, souvent ou toujours. Si l'on considère les réponses des 147 médecins répondants qui ont répondu à l'ensemble des 5 questions, on constate qu'il n'y a aucune association statistique significative⁴ avec la catégorie professionnelle (catégorisée en « généraliste » vs. « spécialiste ») ou avec le nombre d'années de pratique (catégorisée en « moins de 10 ans » vs. « plus de 10 ans »).

Tableau 2 : Réponses au questionnaire en ligne.

Question	(n)	Réponses				
		Connaissance RPC	Formation	Consultation	Cas clinique	Autre
1. Contexte d'utilisation	(185)	55 %	22 %	9 %	12 %	3 %
2. Facilité d'utilisation	(182)	Très difficile 2 %	Plutôt difficile 22 %	Plutôt facile 54 %	Très facile 33 %	—
3. Utile pratique quotidienne	(183)	Pas du tout 3 %	Plutôt non 9 %	Plutôt oui 50 %	Tout à fait 38 %	—
4. Intention d'utiliser	(182)	Jamais 7 %	Parfois 45 %	Souvent 38 %	Toujours 11 %	—
5. Utile compréhension RPC	(180)	Pas du tout 3 %	Plutôt non 11 %	Plutôt oui 50 %	Tout à fait 36 %	—

Parmi les répondants, 67 (36 %) ont saisi des commentaires en texte libre. Dix commentaires, trop succincts, n'ont pas pu être interprétés par les auteurs et 3 faisaient mention de problèmes à caractère technique. Sur les 54 commentaires exploitables, 17, soit près d'un tiers, portaient sur le contenu même des RPC parmi lesquels 15 étaient plutôt critiques à leur égard. Les 37 autres commentaires portaient plus sur la version structurée de RecosDoc-diabète. Neuf commentaires ont été jugés positifs, 12 négatifs et 16 étaient plutôt des suggestions de développement ou d'amélioration. Les commentaires négatifs peuvent être résumés ainsi : « l'arbre décisionnel est trop complexe » (7), « une visualisation de l'arbre dans son entier serait préférable » (3), « la méthode n'est pas didactique » (1) et « la médecine est un art et ne peut être informatisée » (1). Les suggestions d'amélioration ont porté essentiellement sur le développement d'une application pour smartphone et sur l'intégration de l'outil dans les logiciels métiers.

4. Test exact de Fisher.

5 Discussion

5.1 Modélisation de l'ensemble des possibles ou simple reproduction des RPC

La modélisation des connaissances selon les différentes étapes préconisées par Shiffman et al. [12] a permis d'identifier les ambiguïtés, les verbalisations abstraites et les « trous » de connaissances dans la version textuelle des RPC. Cependant, la HAS n'a pas souhaité prendre en compte toutes les observations faites pour améliorer le texte des RPC. Nous pensons néanmoins que la modélisation réalisée conjointement à la rédaction des RPC textuelles a permis d'améliorer la qualité du texte et de son équivalent structuré.

Des difficultés subsistent car lorsque les éléments scientifiques probants manquent pour asseoir les recommandations, il est difficile d'arbitrer entre les différentes opinions qui peuvent s'exprimer, même au sein du GT, et de trouver un consensus. Cela s'applique aussi bien pour la caractérisation des profils patients devant être couverts que pour la spécification des actions recommandées.

Dans la modélisation complète, la majeure partie des profils cliniques identifiés (766) étaient étiquetés « avis d'experts », ce qui a posé un problème pour une diffusion du système RecosDoc-diabète par l'HAS, même si certains membres du GT avaient accepté de revoir tous ces profils. À cause de possibles suites médico-légales, les agences nationales ne sont pas prêtes à diffuser des avis d'experts différents des accords professionnels adoptés par le GT, même en les distinguant, car la méthodologie pour les obtenir n'entre pas dans la méthodologie officielle préconisée pour réaliser les RPC (revue de la littérature et consensus au sein du groupe de travail [2]).

La nouvelle version de la base de connaissances de RecosDoc-diabète « colle » aux RPC selon le souhait de la HAS en ne faisant figurer que les recommandations explicitement mentionnées dans le texte, sans fournir de solution pour les autres profils cliniques. La complétion de la base de connaissances pour fournir une aide à la décision plus large, doit donc être réalisée par des entités externes, sociétés savantes ou privées.

Dans l'état actuel, RecosDoc-diabète peut être utilisé de manière documentaire en naviguant dans l'arbre décisionnel. La production automatisée de recommandations à partir de données patients nécessiterait l'adoption de terminologies d'interface, mais certaines notions qui sont restées abstraites demeurent un obstacle réel.

5.2 Évaluation du système RecosDoc-diabète

De nombreuses études ont rapporté que les SADM basés sur les RPC peuvent avoir un impact sur les pratiques, a priori plus important que la simple diffusion des textes [9]. Le développement de RecosDoc-diabète représente la première expérience de diffusion de RPC sous une forme structurée pour la HAS. Depuis la publication des RPC diabète, RecosDoc-diabète a été consulté moins fréquemment que la version textuelle des RPC. Ceci peut s'expliquer par le fait que la production de référence de la HAS reste les RPC textuelles, et que la version structurée d'accompagnement était secondaire. Cela peut également provenir d'une faible visibilité du système, que ce soit au travers de l'indexation par les moteurs de recherche, ou par le petit paragraphe de présentation de RecosDoc-diabète, peu visible sur la page web des RPC du site de la HAS. Cependant, même après la phase initiale de curiosité suite à l'annonce des RPC en février 2013, les accès au système persistent dans le temps, ce qui est un résultat encourageant.

L'analyse des commentaires montre que presque 30 % des répondants se sont servis de l'enquête en ligne pour critiquer les RPC, peu « fondées sur l'évidence scientifique », constituées essentiellement d'accords professionnels, incomplètes, et formulées à un haut niveau d'abstraction (p. ex. utilisation des classes médicamenteuses alors que les noms des médicaments sont attendus). Les commentaires sur le système sont curieusement contradictoires, certains répondants trouvant le système trop complexe, d'autres le trouvant trop simple. Ceci souligne le fait qu'en termes d'aide à la décision, une solution unique pour le traitement de l'information et l'ergonomie des interfaces ne convient pas car un système statique ne peut pas convenir à la fois aux cas cliniques simples et complexes, ou encore aux utilisateurs experts et novices. D'autres suggestions concernent l'enrichissement de RecosDoc-diabète avec des liens vers des ressources pédagogiques pour les patients, les étudiants voire les médecins. Certains souhaiteraient une intégration du système dans leur logiciel de dossier patient électronique, ou bien une application pour smartphone. Comme les répondants n'ont représenté que 2 % des visiteurs du système, ces commentaires ne reflètent pas nécessairement l'avis de tous les visiteurs, ce qui représente un biais de ce type de sondage reposant sur le volontariat.

6 Conclusion

En France, la HAS a diffusé pour la première fois, en même temps que la version textuelle de ses RPC sur la prise en charge médicamenteuse du diabète de type 2, une version structurée de ces RPC sous la forme du système RecosDoc-diabète. Cette première expérience, certes timide, présente des résultats encourageants. Elle nécessite d'être poursuivie, à la fois pour intégrer les phases de modélisation des connaissances dans le processus de rédaction des RPC et pour favoriser l'utilisation des versions structurées des RPC, que ce soit en ligne ou intégrées dans des applications métier, tout en évaluant l'impact de ces versions structurées sur les pratiques.

Références

- [1] Field MJ, Lohr KN, editors. Clinical Practice Guidelines : Directions for a New Program. Institute of Medicine. Washington DC : National Academy Press ; 1990.
- [2] Haute Autorité de Santé. Méthodes quantitatives pour évaluer les interventions visant à améliorer les pratiques. Guide méthodologique. HAS ; 2007. [Accessible sur www.has-sante.fr].
- [3] Farmer AP, Légaré F, Turcot L, Grimshaw J, Harvey E, McGowan J, et al. Printed educational materials : effects on professional practice and health care outcomes. Cochrane Database of Systematic Reviews. 2008 ;Issue 3. Art. No. : CD004398.
- [4] Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyenne J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes : a systematic review. JAMA. 2005 ;293(10) :1223–1238.
- [5] Shojanian KG, Jennings A, Mayhew A, Ramsay CR, Eccles MP, Grimshaw J. The effects of on-screen, point of care computer reminders on processes and outcomes of care. Cochrane Database Syst Rev. 2009 ;(3). CD001096. DOI : 10.1002/14651858.CD001096.pub2.
- [6] Jaspers MW, Smeulders M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes : a synthesis of high-quality systematic review findings. J Am Med Inform Assoc. 2011 ;18(3) :327–334.

- [7] Bouaud J, Falcoff H, Séroussi B. Simultaneously authoring and modelling clinical practice guidelines : a case study in the therapeutic management of type 2 diabetes in France. In : Data and Knowledge for Medical Decision Support. vol. 186 of Stud Health Technol Inform. IOS Press ; 2013. p. 108–112.
- [8] Haute Autorité de Santé. Recommandation de bonne pratique : Stratégie médicamenteuse du contrôle glycémique du diabète de type 2. HAS ; 2013.
- [9] Flottorp S, Oxman A, Krause J, Musila N, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice : A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. *Implement Sci.* 2013 ;8(1) :35.
- [10] Peleg M, Tu SW, Bury J, Ciccarese P, Fox J, Greenes RA, et al. Comparing computer-interpretable guideline models : a case-study approach. *J Am Med Inform Assoc.* 2003 ;10(1) :52–68.
- [11] Patel VL, Allen VG, Arocha JF, Shortliffe EH. Representing Clinical Guidelines in GLIF : individual and collaborative expertise. *J Am Med Inform Assoc.* 1998 ;5(5) :467–483.
- [12] 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) :418–426.
- [13] Shahar Y, Shalom E, et al AM. A Distributed, Collaborative, Structuring Model for a Clinical-guideline Digital-library. *J Am Med Inform Assoc.* 2003 ;p. 589–93.
- [14] Goud R, Hasman A, Strijbis A, Peek N. A parallel guideline development and formalization strategy to improve the quality of clinical practice guidelines. *Int J Med Inf.* 2009 ;78(8) :513–520.
- [15] Shiffman R, Michel G, Rosenfeld R, Davidson C. Building better guidelines with BRIDGE-Wiz : development and evaluation of a software assistant to promote clarity, transparency, and implementability. *J Am Med Inform Assoc.* 2012 ;19(1) :94–101.
- [16] Séroussi B, Bouaud J, Antoine ÉC. 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.

Adresse de correspondance

J. Bouaud
LIMICS, INSERM U1142
15, rue de l'école de médecine
75006 Paris, France
jacques.bouaud@sap.aphp.fr

Mapping de terminologies diagnostiques en cancérologie par l'intermédiaire du NCI Metathesaurus

Mapping of diagnostic terminologies in oncology using the NCI Metathesaurus

Bérénice Brechat^{1,2} Fleur Mougin¹ Frantz Thiessard^{1,2} Vianney Jouhet^{1,2}

¹ *Equipe de Recherche en Informatique Appliquée à la Santé, ISPED, Université de Bordeaux*

² *Service d'Information Médicale, Pôle de Santé Publique, CHU de Bordeaux*

Résumé

En France, plusieurs terminologies sont utilisées en routine pour coder des diagnostics de cancer (CIM10 pour le PMSI, CIM-O3 au sein des registres, ADICAP pour l'anatomo-pathologie). Cette hétérogénéité est un frein majeur pour l'utilisation secondaire des données diagnostiques recueillies. Il est nécessaire de mettre en correspondance les différentes terminologies diagnostiques de cancérologie au sein d'un système cohérent et hiérarchisé. Le NCI Thesaurus, qui est une terminologie internationale visant à représenter l'ensemble des connaissances en cancérologie, semble être une excellente base pour constituer ce système unifié. Dans ce travail, le mapping existant entre la CIM-O3 et le NCI Thesaurus au sein du NCI Metathesaurus est considéré. Plus de deux tiers des codes CIM-O3 sont associés à un code NCI Thesaurus unique. Une analyse qualitative des mappings multiples et des cas où un code CIM-O3 n'a pas de correspondant NCI Thesaurus a été réalisée. Différents types de causes ont été identifiés et des pistes de solution sont proposées en discussion.

Abstract

Many terminologies are used in France for coding cancer diagnoses (ICD-10 for Diagnosis Related Groups, ICDO3 in cancer registries, ADICAP for pathological anatomy). This heterogeneity largely hinders the use of registries' data. It is thus necessary to integrate the diagnostic terminologies in oncology into a unified and structured system. The NCI Thesaurus is an international terminology, which aims at representing the whole knowledge about oncology. This resource seems to be an excellent basis for constituting this integrated system. In this work, we are investigating the mapping existing between the ICDO3 and the NCI Thesaurus through the NCI Metathesaurus. More than two-third of ICDO3 codes are mapped to a unique NCI Thesaurus code. A qualitative analysis is performed for the multiple mappings and for cases where an ICDO3 code had no corresponding code in the NCI Thesaurus. Some categories of causes were identified and potential solutions are mentioned in the discussion.

Mots-clés : CIM-O3, NCI Thesaurus, mapping de terminologies biomédicales

Keywords: ICDO3, NCI Thesaurus, Terminology mapping

Articles longs des 15^{es} Journées francophones d'informatique médicale, JFIM 2014, pages 34–43

Fès, Maroc, 12–13 juin 2014

1 Introduction

Les activités liées à la santé produisent aujourd'hui une masse importante d'information qu'il faut trier, traiter, recouper pour qu'elle soit exploitable, mais aussi stockée en vue d'une utilisation ultérieure. Face à ces problématiques, les acteurs du domaine médical ont pris conscience de l'importance de structurer l'information qu'ils génèrent [1], mais ces données sont produites par une multitude d'acteurs différents et leur mise en correspondance reste complexe.

Le domaine de la cancérologie illustre bien cette production considérable d'information par des sources multiples. On dispose par exemple de données issues de la clinique, de la recherche ou encore de l'épidémiologie. Pour représenter ces données, un certain nombre de classifications ont été créées et l'enjeu est aujourd'hui de trouver des correspondances entre elles. Les registres des cancers, qui visent à recueillir de façon exhaustive de nombreux éléments relatifs aux cancers incidents dans la population, sont organisés en réseaux avec une volonté de standardiser ce recueil. Le choix s'est porté sur la Classification Internationale des Maladies pour l'Oncologie (CIM-O3) comme nomenclature commune pour la codification des cancers [2]. Pour assurer leur fonction, les registres collectent de l'information issue de sources, et donc de terminologies multiples telles que la classification internationale des maladies ou l'ADICAP (codification des lésions élaborée par l'Association pour le Développement de l'Informatique en Cytologie et en Anatomie Pathologique). Ces registres sont également exploités dans de nombreux travaux et études relatifs au cancer, où leurs données sont croisées avec d'autres bases de données codées grâce à des nomenclatures diverses. Disposer d'un mapping exhaustif entre la CIM-O3 et d'autres terminologies présente donc un intérêt majeur.

Le NCI Thesaurus est une terminologie internationale de référence développée par le National Cancer Institute américain, visant à représenter l'ensemble des connaissances largement utilisée en cancérologie, que ce soit en clinique, en recherche translationnelle et fondamentale, ou encore en épidémiologie. De ce fait, elle semble être un excellent pivot de mapping entre les terminologies du spectre de la cancérologie. On retrouve le NCI Thesaurus, ainsi que la CIM-O3, au sein d'un système unifié de terminologies biomédicales qui contient la plupart des terminologies utilisées dans le domaine de la cancérologie et qui a été développé sur la base de l'UMLS [3] : le NCI Metathesaurus.

L'objectif principal de cette étude est de représenter l'ensemble des concepts de la CIM-O3 au sein du NCI Thesaurus via le NCI Metathesaurus. Pour cela, une analyse de la correspondance entre les concepts CIM-O3/ NCI Thesaurus a été réalisée pour identifier d'éventuelles inexactitudes, erreurs ou absences de mapping.

2 État de l'art

La question de l'interopérabilité entre diverses ontologies a été largement étudiée ces dernières années, avec en particulier la volonté de proposer des méthodes automatiques [4,5]. Dans le domaine biomédical également, cette problématique a généré de très nombreux travaux et l'utilisation de l'UMLS comme pivot a été souvent investiguée [6,7,8].

En revanche, il existe peu de travaux où le mapping est basé sur le NCI Metathesaurus [9]. Disposer d'un mapping de qualité entre la CIM-O3 et une terminologie internationale de référence tel le NCI Thesaurus présente un intérêt majeur pour le partage des données épidémiologiques des registres des cancers. On retrouve des travaux portant sur la recherche d'inconsistances d'un tel mapping, réalisés à partir de versions antérieures du NCI Metathesaurus [9]. Cependant, aucune version corrigée et validée de ce type de mapping n'est actuellement disponible. L'actualisation des concepts portés par ce Metathesaurus justifie également un nouveau travail de mapping.

3 Matériel et méthodes

3.1 Terminologies utilisées pour le mapping

3.1.1 La CIM-O3

La Classification Internationale des Maladies pour l'Oncologie (CIM-O3 ou ICDO3 en version anglaise), est une classification [10], principalement utilisée par les registres de tumeurs cancéreuses pour le codage des caractéristiques de tumeurs. La CIM-O3 possède un axe morphologique pour décrire la morphologie de la tumeur et un axe topographique qui précise une localisation.

Un code CIM-O3 de tumeur est composé de 10 caractères correspondant à l'agrégation des quatre caractères du code topographique, suivis des cinq caractères du code morphologique, éventuellement suivis d'un caractère facultatif (Figure 1). A chaque code CIM-O3 correspond un terme préféré et des inclusions décrivant généralement des notions plus précises. Par exemple, le code 9440/3 a pour terme préféré « Glioblastome, SAI » et pour inclusions « Glioblastome multiforme », « Spongioblastome multiforme » et « Gliome à cellules indifférenciées ».

La CIM-O3 nécessite donc une post-coordination pour le codage des tumeurs car le diagnostic n'est pas codé tel-que dans la terminologie. Il faut, pour coder un diagnostic précis de tumeur, associer un élément morphologique de la terminologie à un élément topographique de la terminologie.

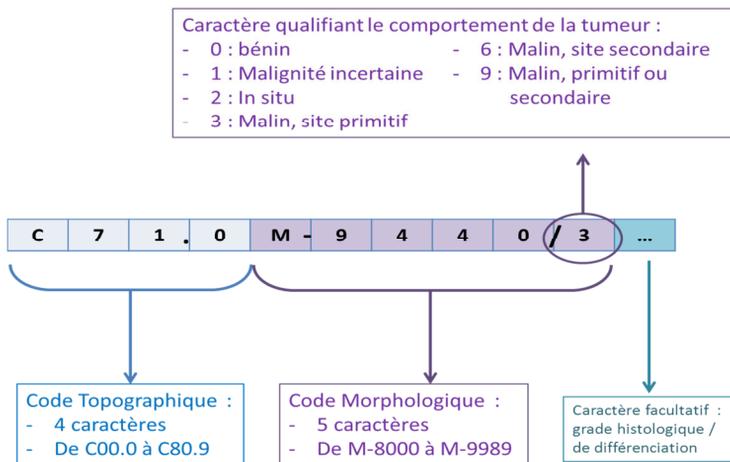


Figure 1 : Eléments permettant la construction d'un code CIM-O3, exemple du glioblastome (code morphologique 9440/3) d'un hémisphère cérébral (code topographique C71.0)

3.1.2 Le NCI thesaurus

Le NCI Thesaurus (NCI) est une terminologie de référence développée par le National Cancer Institute américain, qui couvre le vocabulaire anglais utilisé en cancérologie au sens large [11]. La version utilisée dans ce travail est celle qui est intégrée dans le NCI Metathesaurus (13.06D).

Le NCI comprend plus de 43 000 concepts, organisés selon 20 sous-domaines appelés Kind, tel « Abnormal Cell Kind » qui regroupe les éléments d'anatomo-cyto-pathologie. Associées à ces

concepts, on retrouve des précisions comme le terme préféré (preferred term), des synonymes, des définitions, et d'autres informations complémentaires.

Le NCI possède plus de 90 types de relations définies entre les concepts comme par exemple *disease_has_abnormal_cell* pour qualifier un rapport entre une pathologie et un type de cellule. Aujourd'hui, le NCI possède plus de 200 000 relations entre ses concepts.

Ainsi, le NCI comporte à la fois des entités pré-coordonnées (c'est-à-dire une entité composée d'éléments associés lors de la conception du thésaurus, par exemple « Breast Carcinoma » qui est codé C4872) et des notions pouvant être post-coordonnées grâce aux relations (le terme « Breast Carcinoma » est par exemple relié par *disease_has_primary_anatomic_site* à « Breast »).

Cette prise en charge des termes pré- et post-coordonnés fait du NCI fait du NCI une ressource pour la construction d'un pivot permettant d'intégrer différentes terminologies diagnostiques en cancérologie. Il permet en effet de représenter aussi bien des terminologies utilisant des termes pré-coordonnés (comme par exemple la CIM-10) que des terminologies nécessitant une post-coordination (comme par exemple la CIM-O3), et donc de mettre en relation ces différents éléments.

3.1.3 Le NCI Metathesaurus

Le NCI Metathesaurus est un système utilisé dans le spectre de la cancérologie qui regroupe plus de 75 terminologies biomédicales différentes [12]. Il a été élaboré par le National Cancer Institute américain à partir du Metathesaurus de l'UMLS enrichi par d'autres terminologies biomédicales spécifiques à la cancérologie.

On y retrouve 4 millions de termes représentant 2 millions de concepts biomédicaux.

Ces termes sont reliés entre eux par plus de 22 millions de relations, présentes dans les différentes terminologies. Les différents termes et codes des différentes terminologies représentant un même concept sont codés dans le NCI Metathesaurus sous un code unique : le Concept Unique Identifier (CUI).

A chaque concept est également rattaché l'ensemble des informations complémentaires (définitions, synonymes...) présents initialement dans les terminologies. Enfin, les concepts sont catégorisés par au moins un type sémantique parmi les 133 possibles.

3.2 Mapping entre la CIM-O3 et le NCI par le CUI

Les codes de la CIM-O3 ont été mappés aux codes du NCI par correspondance exacte de leur(s) CUI(s). Ainsi, le mapping a été effectué à partir de la liste des CUIs des codes de la CIM-O3. Pour chacun de ces codes, on a recherché le ou les codes du NCI appartenant au même CUI (Figure 2).

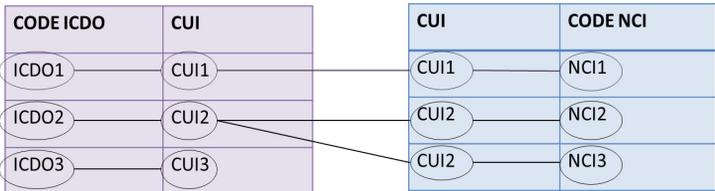


Figure 2 : Principe du mapping des codes de la CIM-O3 avec les codes du NCI par l'intermédiaire des CUI dans le NCI Metathesaurus

Différentes cardinalités de mapping vont émerger :

- Mapping 1-1 : le code CIM-O3 est mappé à un et un seul code du NCI
- Mapping 1-n : le code CIM-O3 est mappé à plusieurs codes du NCI
- Mapping 1-0 : le code CIM-O3 ne retrouve pas de correspondance dans le NCI

L'objectif de l'étude étant de réaliser un mapping de la CIM-O3 vers le NCI, les mappings n-1 n'ont pas été étudiés.

3.3 Analyse du mapping code à code entre la CIM-O3 et le NCI

Une analyse qualitative des mappings 1-n et 1-0 a été effectuée pour identifier des motifs de mapping multiple ou d'absence de mapping et tenter de les grouper par type d'erreur.

Cette analyse a été réalisée manuellement afin de mettre en évidence des inexacitudes, des erreurs et des carences de mapping pouvant être résolues, afin de disposer d'un meilleur mapping de la CIM-O3 vers le NCI.

4 Résultats

4.1 Mapping entre la CIM-O3 et le NCI par le CUI

Le NCI Metathesaurus contient 3634 termes distincts issus de la CIM-O3, codés par 1500 codes CIM-O3 distincts, dont 1091 codes correspondant à des termes de morphologie et 409 codes correspondant à des termes de topographie.

L'utilisation des CUI du NCI Metathesaurus permet de mettre en relation 1261 codes CIM-O3 avec 1670 codes du NCI (Figure 3).

En distinguant le type de code CIM-O3, 963 des 1091 codes morphologiques (soit 88,3%) et 298 des 409 codes topographiques (soit 72,9%) trouvent une correspondance avec un terme du NCI par l'intermédiaire du NCI Metathesaurus. Ainsi, 239 codes de la CIM-O3 (sur les 1500 présents dans le NCI Metathesaurus) ne sont pas mappés à un code NCI (soit 15,9%), correspondant aux mappings de cardinalité 1-0 (Tableau 1). Un exemple de ce type de cardinalité est le code C02.0 codant « Dorsal surface of tongue ».

Tableau 1 : Récapitulatif du mapping des codes CIM-O3 aux codes NCI par cardinalité du mapping obtenu

Type de mapping	Nombre (pourcentage)
Mapping 1-1	1008 (67,2%)
Mapping 1-n	253 (16,9%)
Mapping 1-0	239 (15,9%)

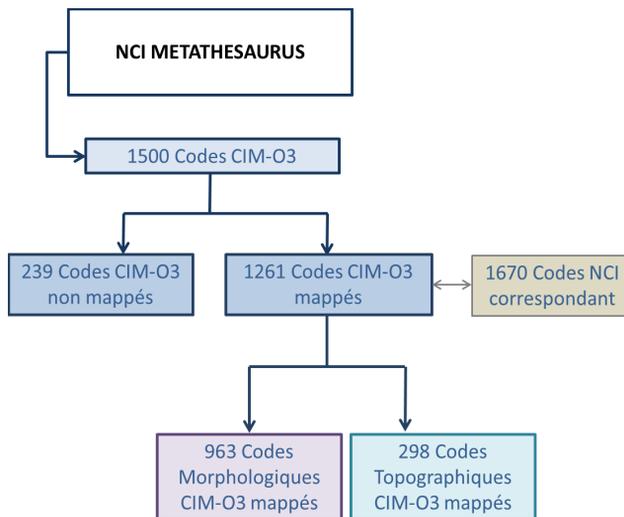


Figure 3 : Répartition et mapping des codes de la CIM-O3 présents dans le NCI Metathesaurus

Par ailleurs, 1008 mappings 1-1 ont été obtenus. Par exemple, le code CIM-O3 8123/3 - « Basaloid carcinoma » est mappé exclusivement au code NCI C4121 - « Basaloid Carcinoma ». Enfin, 253 mappings 1-n ont été identifiés. Par exemple, le code CIM-O3 8000/1 (correspondant au concept de tumeur sans précision) est quant à lui mappé aux codes NCI C65157 et C3262 (Figure 4).

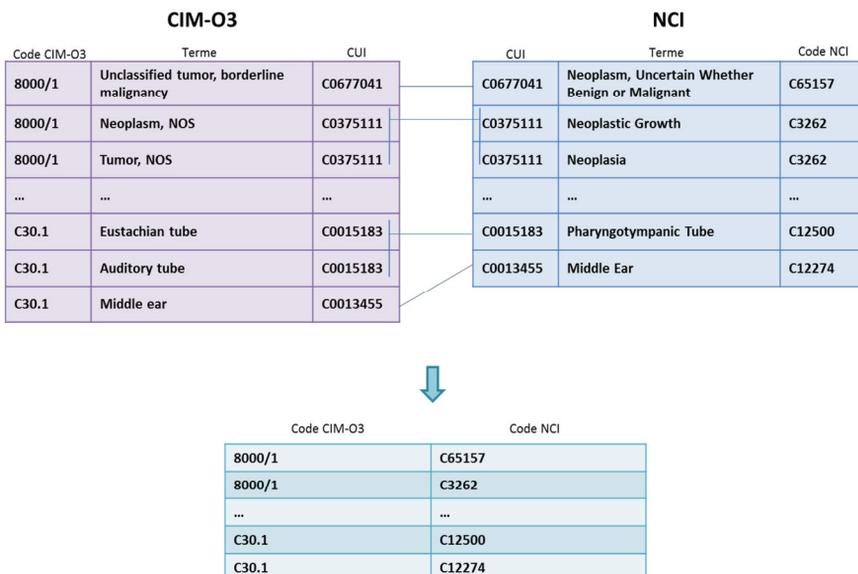


Figure 4 : Exemples de mappings 1-n de codes CIM-O3 avec le NCI

4.2 Analyse du mapping code à code entre la CIM-O3 et le NCI

4.2.1 Différence de granularité

Il existe des mappings multiples dus à la différence de granularité entre les deux terminologies étudiées. Dans ce cas de figure, un même code CIM-O3 est mappé à la fois au code NCI qui lui correspond exactement, mais aussi à ses codes fils plus précis, qui n'ont en fait pas de correspondant dans la CIM-O3. Par exemple, le code CIM-O3 8210/3 - « Adenocarcinoma in adenomatous polyp » est mappé au code NCI C7682 « Carcinoma in a Polyp » mais aussi à son code fils C7676 - « Adenocarcinoma in Adenomatous Polyp » (Figure 5).

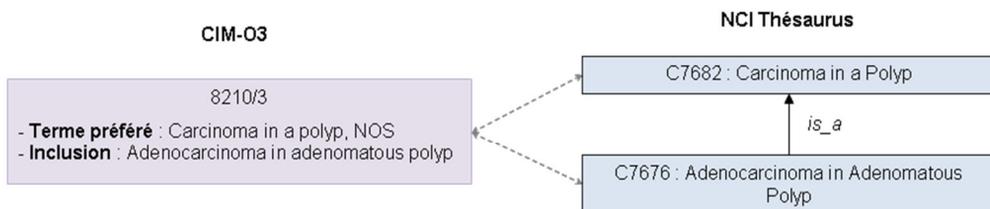


Figure 5 : Exemple de différence de granularité entre la CIM-O3 et le NCI

4.2.2 Pré-coordination au sein du NCI

On retrouve des mappings multiples dus à la présence de concepts pré-coordonnés dans le NCI. Les codes NCI associés à ces concepts sont mappés aux codes morphologiques de la CIM-O3 qui sont eux-mêmes également mappés aux codes NCI décrivant uniquement la morphologie correspondante.

Par exemple, le code CIM-O3 8260/3 - « Papillary adenocarcinoma, NOS » est mappé à la fois au concept morphologique « Papillary Adenocarcinoma » codé C2853 dans le NCI, mais aussi au code C6975 du concept pré-coordonné « Papillary Renal Cell Carcinoma » (Figure 6).

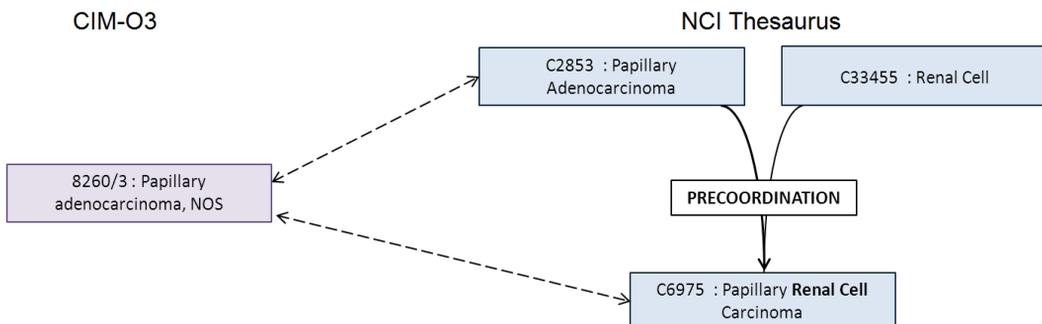


Figure 6 : Exemple de mapping multiple due à la pré-coordination dans le NCI

4.2.3 Les mappings 1-0

L'analyse des mappings 1-0 a montré des lacunes dues à la différence de granularité entre les deux terminologies. Ainsi, on retrouve des concepts fins au sein de la CIM-O3 qui ne retrouvent pas de correspondance aussi spécifique au sein du NCI. Par exemple, le code CIM-O3 C02.0 - « Dorsal surface of tongue, NOS » - codant la surface dorsale de la langue ne retrouve pas le même degré de précision au sein du NCI. Dans ce dernier, le concept de langue (C12422 - « Tongue ») est une feuille, sans concept fils.

Par ailleurs, certains mappings 1-0 ne devraient pas l'être car des concepts apparemment synonymes n'appartiennent pas à un même CUI dans le NCI Metathesaurus (problème de « missed synonymy » fréquent dans ce type de large graphe [13]). Par exemple, le code CIM-O3 C42.0 ayant pour terme préféré « Blood » n'est pas mappé au code NCI C12434 ayant pour terme préféré « Blood » car ils sont dans deux CUIs distincts : respectivement, CL429748 et C0005767.

5 Discussion

L'utilisation du NCI Metathesaurus a permis de retrouver une majorité de correspondances 1-1 entre la CIM-O3 et le NCI. Cependant, plus de 30% des mappings obtenus sont soit 1-n soit 1-0. Le NCI Metathesaurus ne peut donc pas être utilisé directement pour obtenir un mapping code à code exact de la CIM-O3 vers le NCI. Un travail complémentaire doit ainsi être mené pour parvenir à cet objectif. L'analyse qualitative a permis de mettre en évidence les causes de mappings multiples ou d'absence de mapping dans le but de proposer des solutions ultérieurement. En présence de mappings 1-n, des règles de correction automatiques pourront être définies afin de déterminer la correspondance correcte et d'exclure les autres. Dans cette optique, une approche simple sera de comparer le terme préféré du code CIM-O3 avec les termes des codes NCI associés. Dans la figure 6, on voit que le mapping entre le code 8260/3 « Papillary adenocarcinoma, NOS » et le code NCI C2853 « Papillary Adenocarcinoma » est correct puisque les deux termes sont équivalents tandis que celui entre 8260/3 et C6975 « Papillary **Renal Cell** Carcinoma » est faux. En effet, le terme NCI étant pré-coordonné, il décrit l'organe où se situe le carcinome alors que cette information n'est pas présente dans le code CIM-O3 auquel il est mappé. En considérant les termes associés aux codes NCI, on déterminera automatiquement que le mapping correct est vers C2853. Des méthodes de désambiguïsation par l'intermédiaire de la catégorisation sémantique, comme dans [14], pourraient également être appliquées. Comme l'illustre la figure 7, cela permettrait de résoudre des cas comme celui du code CIM-O3 C38.0 (« Heart » catégorisé par le type sémantique « Body Part, Organ, or Organ Component ») qui est mappé aux codes NCI C12727 (« Heart », également catégorisé par « Body Part, Organ, or Organ Component ») et C4570 (« Malignant Endocardial Neoplasm », catégorisé par le type sémantique « Neoplastic Process »). En filtrant les mappings suivant la catégorisation sémantique, on pourrait écarter automatiquement le deuxième mapping.

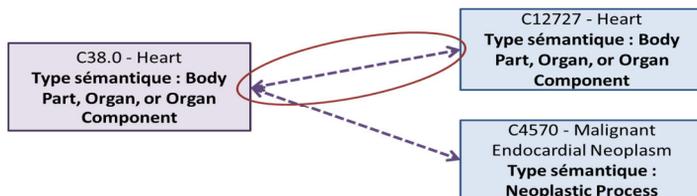


Figure 7 : Exemple de désambiguïsation grâce aux types sémantiques

Pour la plupart des mappings 1-0, la granularité du NCI est insuffisante. Il faudra donc envisager l'ajout de nouveaux codes pour pouvoir représenter les codes CIM-O3 plus fins. Il faudra également traiter les cas dus à la « missed-synonymy ». Par ailleurs, il reste quelques cas isolés de défaut de mapping ne pouvant pas être corrigés automatiquement et pour lesquels une correction manuelle devra être réalisée.

Ce travail de mapping des concepts de la CIM-O3 vers le NCI s'inscrit dans le cadre d'un projet plus large, dont l'objectif est de permettre l'intégration de différentes terminologies diagnostiques en cancérologie. Afin de garantir l'exactitude des mappings obtenus, une évaluation de l'ensemble des résultats sera réalisée par des médecins. Ces mappings validés serviront de jeux de données pour l'évaluation des règles de sélection des mapping multiples.

D'autre part, nous avons identifié un certain nombre de mappings de la forme n-1, à savoir la convergence de plusieurs codes CIM-O3 vers un même code NCI. Par exemple, les codes CIM-O3 8041/3 - « Small cell carcinoma, NOS » et 8042/3 - « Oat cell carcinoma » sont mappés à l'unique code NCI C3915 qui inclut ces deux notions. Ce type de cardinalité n'a pas été pris en compte dans l'analyse qualitative car l'objectif de l'étude était de réaliser un mapping de la CIM-O3 vers le NCI. Cependant, dans un objectif d'intégration des différentes terminologies d'intérêt tout en préservant leur granularité, il est indispensable de traiter ce cas de figure. La prise en compte de ces situations ainsi que le traitement des mappings 1-0 conduiront à l'ajout de termes dans un ensemble basé sur le NCI permettant d'intégrer des terminologies diagnostiques.

Peu de travaux se sont intéressés à l'évaluation des mappings des terminologies diagnostiques à partir du NCI Metathesaurus. Notre travail vise à identifier des mapping uniques et exacts de la CIM-O3 vers le NCI, ce qui diffère de l'objectif principal de l'étude de Burgun et Bodenreider [9] qui cherchait à évaluer la cohérence de la représentation des tumeurs entre ces deux terminologies. Considérant que seuls des termes post-coordonnés existent dans la CIM-O3, nous ne cherchons pas ici à reconstruire les termes pré-coordonnés associant topographie et morphologie pour les mapper vers le NCI. Les résultats du mapping des termes post coordonnés de la CIM-O3 vers le NCI sont cependant similaires à la précédente étude, les différences observées étant dues aux évolutions du NCI Metathesaurus. Le travail de Burgun et Bodenreider [9] s'inscrit dans une perspective d'évaluation de la qualité des terminologies et reste orienté vers l'audit alors que nous cherchons à représenter les connaissances de la CIM-O3 dans une terminologie qui nous permettra, à terme, de mettre en correspondance ses concepts avec d'autres terminologies de cancérologie. Dans cette optique, nous vérifierons que les combinaisons CIM-O3 topographie + morphologie utilisées pour coder des données réelles dans un registre des cancers peuvent être mis en relation avec des termes pré-coordonnés du NCI.

6 Conclusion

Le mapping code à code de la CIM-O3 vers le NCI par l'intermédiaire du NCI Metathesaurus révèle des mappings exacts, des mappings multiples et permet d'identifier les termes absents dans le NCI. L'analyse qualitative a mis en évidence des causes fréquentes aux mappings multiples, ce qui permettra la construction et l'application de règles de correction afin d'obtenir une intégration complète de la CIM-O3 avec le NCI. La confrontation du résultat de cette intégration à des données réelles d'un registre permettra d'identifier l'absence éventuelle de termes pré-coordonnés dans le NCI.

Références

- [1] Kalra D1, Fernando B, Morrison Z, Sheikh A. *A review of the empirical evidence of the value of structuring and coding of clinical information within electronic health records for direct patient care*. Inform Prim Care. 2012;20(3):171-80.
- [2] Jensen OM, Parkin DM, MacLennan R, Muir CS, Skeet RG, ed. *Cancer registration: principles and methods*. Vol. 95. IARC, 1991
- [3] Lindberg DA, Humphries BL, McCray AT. *The Unified Medical Language System. Methods Inf Med*. 1993;32(4):281–291.
- [4] Noy NF. *Tools for mapping and merging ontologies*. In: Staab S, Studer R, editors Handbook on Ontologies: Springer-Verlag. 2004;365–384
- [5] Shvaiko P, Euzenat J. *A survey of schema-based matching approaches*. Journal on Data Semantics. 2005;4.
- [6] Fung KW, Bodenreider O. *Utilizing the UMLS for semantic mapping between terminologies*. AMIA Annu Symp Proc. 2005;266-70
- [7] Mougín F, Burgun A, Bodenreider O. *Comparing drug-class membership in ATC and NDF-RT*. Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, ISBN: 978-1-4503-0781-9, 2012;437-443
- [8] Zhang S, Bodenreider O. *Experience in Aligning Anatomical Ontologies*. Int J Semant Web Inf Syst. 2007;3(2):1-26
- [9] Burgun A, Bodenreider O. *Issues in integrating epidemiology and research information in oncology: experience with ICD-O3 and the NCI Thesaurus*. AMIA Annu Symp Proc. 2007;85-9.
- [10] Fritz A, Percy C, Jack A, Shanmugaratnam K, Sobin L, Parkin M, et al., ed. *International classification of diseases for oncology*. 3rd ed. Geneva : World Health Organization. 2000
- [11] Golbeck J, Fragoso G, Hartel FW, Hendler J, Oberthaler J, Parsia B. *The National Cancer Institute's Thesaurus and Ontology*. J Web Semantics. 2003;75–80
- [12] NCI Metathesaurus. <http://ncimeta.nci.nih.gov/ncimbrowser/> (consulté le 14/04/2014)
- [13] Hole, W, Srinivasan, M. *Discovering Missed Synonymy in a Large Concept-Oriented Metathesaurus*. AMIA Annu Symp Proc. 2003;354- 358.
- [14] Mougín F, Dupuch M, Grabar N. *Improving the mapping between MedDRA and SNOMED CT*. In Peleg M, Lavrač N, and Combi C (editors). Artificial Intelligence in Medicine. Berlin Heidelberg: Springer. 2011; LNAI 6747:220-224

Adresse de correspondance

Equipe de Recherche en Informatique Appliquée à la Santé
ISPED – Université de Bordeaux
146 Rue Léo Saignat
33076 Bordeaux
berenice.brechat@isped.u-bordeaux2.fr

Terminologies et référentiels d'interopérabilité sémantique en Santé

Reference terminologies and templates for semantic interoperability in healthcare

Amina Chniti¹, Nicolas Griffon^{2,3}, Lamine Traore^{1,2}, Sajjad Hussain^{1,2}, Jean-Marie Rodrigues², Stefan Darmoni^{3,2}, Jean Charlet², Eric Sadou², David Ouagne², Eric Lepage^{1,2} et Christel Daniel^{1,2}

¹*AP-HP, F-75006, Paris, France*

²*INSERM, U1142, LIMICS, F-75006, Paris, France, Sorbonne Universités, UPMC Univ Paris 06, UMR_S 1142, LIMICS, F-75006, Paris, France, Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), F-93430, Villetaneuse, France*

³*CISMef, CHU Rouen, France.*

Résumé

Les systèmes d'information de santé (SIS) sont hétérogènes et rarement interopérables, surtout au niveau sémantique. Dans le cadre du projet ANR/TeRSan (Terminologies et Référentiels d'interopérabilité Sémantique en Santé), nous proposons une plateforme d'interopérabilité sémantique ayant pour objectif de faciliter l'échange d'informations cliniques standardisées entre SIS tout en préservant l'usage des terminologies locales. Dans un premier temps, la méthode consiste à modéliser les processus métier, dans le domaine de la biologie, de l'imagerie et de l'anatomie et cytologie pathologiques (ACP), à identifier les modèles standards définis pour l'échange d'informations au cours de ces processus et à formaliser au sein de ces modèles standards le lien préconisé avec des terminologies de référence. Dans un second temps, il s'agit de produire et maintenir des alignements entre terminologies de référence (TR) et terminologies d'interface (TI) locales et de développer les services sémantiques permettant le transcodage entre messages locaux et standards. Les résultats obtenus dans le domaine de l'ACP consistent en la formalisation d'un modèle de demande d'avis par télépathologie en spécialisant le modèle HL7 de demande de sous-traitance, des alignements entre TI et TR des de ce domaine et une preuve de concept du transcodage réalisé par la plateforme d'interopérabilité sémantique lors de l'échange de télépathologie inter-établissements. Les difficultés observées au niveau de l'alignement des ressources terminologiques semblent représenter la principale limite à dépasser.

Abstract

The Healthcare Information Systems (HIS) are heterogeneous and are rarely interoperable,

especially at the semantic level. In the context of the ANR/Tersan project, we propose a semantic interoperability platform aiming to facilitate the exchange of standard clinical information between HIS –based on the use of Reference Terminologies (RT)-while preserving the use of local Interface Terminologies (IT). The first step of the method consists of identifying the standard models used for information exchange in the clinical contexts of the ANR/Tersan project (laboratory, radiology and anatomic and cytology pathology (ACP)) and to formalize the link between these models and Reference Terminologies (RT) used to represent the clinical information. The second step consists of mapping IT and RT in the scope of the selected domains and to design the semantic services supporting the transcoding. As results, in the ACP domain, we formalized a model of the advice request based on an existing HL7 model for outsourcing request, we initiated mappings between IT and RT in this domain and demonstrate our approach by implementing a prototype of a cross-hospital information exchange invoking semantic services. Nevertheless, the difficulties observed in the terminologies mappings seem to represent a significant limit.

Mots-clés : interopérabilité sémantique, terminologie d’interface, terminologie de référence, modèle d’information.

Keywords: *semantic interoperability, local terminology, reference terminology, template.*

1 Introduction

Les systèmes d'information de santé (SIS) sont complexes, hétérogènes et sont rarement interopérables, surtout au niveau sémantique. L'interopérabilité sémantique permet le partage et l'exploitation du sens des données des SIS - faits cliniques, décisions, activités - dès lors qu'elles sont représentées dans un format interprétable par la machine. Dans le domaine de la santé, l'interopérabilité sémantique repose sur des référentiels partagés - modèles d'information («templates») et des terminologies – développés par des organismes de standardisation internationaux. L'émergence opérationnelle de l'interopérabilité sémantique dans le domaine de la santé dépend tout d'abord de la capacité des SIS à intégrer ces référentiels et également de l'adoption, par les professionnels de santé, des solutions de saisie d'information clinique basées sur des interfaces homme-machine adaptées à leurs usages.

En effet, dans la pratique, les principes de structuration et de codage de l'information clinique au sein des SIS sont mis en œuvre de façon spécifique et locale. Même lorsque plusieurs établissements utilisent des SIS du même éditeur, il y a très peu de partage de modèles d'information clinique entre eux. Le niveau de granularité des informations peut aussi varier en fonction de la profession de santé (médecin, infirmiers, kinésithérapeutes, assistantes sociales, etc.), de la spécialité (cardiologie, psychiatrie, imagerie, biologie, etc.) ou encore du mode d'exercice (hospitalisation, consultation, médecine hospitalière, médecine de ville, hospitalisation à domicile, soins ambulatoire, etc.). Les pratiques de documentation locales et les modes de représentation de l'information clinique qui y sont associés représentent des contraintes auxquelles les solutions de partage ou d'échange d'information inter-établissements doivent pouvoir s'adapter. Dès lors que l'information clinique n'est pas d'emblée interopérable lors de sa génération, à la source, des solutions d'interopérabilité sémantique sont nécessaires à la communication et au traitement de cette information au-delà du périmètre où l'information a été générée.

L'objectif de ce travail est le développement de solutions permettant aux SIS de formaliser et de structurer leurs données cliniques selon des référentiels d'interopérabilité sémantique permettant leur partage et leur exploitation standardisée tout en conservant l'usage de leurs terminologies locales. Il s'agit de développer une plateforme de gestion et de distribution centralisée de référentiels d'interopérabilité sémantique et de mettre en œuvre au sein de SIS les outils et services facilitant leur usage.

La solution proposée est fondée sur les technologies du web sémantique afin de faciliter l'échange d'informations cliniques standardisées entre SIS au sein d'un espace de partage. L'originalité de la plateforme est d'offrir des services adaptables à tout SIS déployé dans un établissement de santé permettant de préserver l'usage de terminologies locales, souvent propriétaires mais adaptées aux professionnels de santé. Le caractère innovant de ce travail consiste à gérer au sein de cette plateforme, en plus des terminologies de référence du domaine de la santé (telles que CIM10, ADICAP, LOINC, SNOMED 3.5 VF, CCAM, etc), d'autres types de ressources sémantiques (modèles de structures de données (« templates ») et terminologies d'interface) indispensables à l'émergence d'une utilisation effective et à grande échelle de ces terminologies de référence. Le périmètre métier auquel nous nous sommes limités dans un premier temps est celui du partage inter-établissements dans les domaines de la biologie, de l'imagerie et de l'anatomie et cytologie pathologiques ACP. L'approche proposée est en cours d'évaluation dans le cadre d'une infrastructure pilote d'échange d'information mise en place entre l'AP-HP et le CHU de Rouen.

2 État de l'art

L'initiative internationale *Integrating the Healthcare Enterprise* (IHE) rassemblant des professionnels de la santé et des éditeurs de SIS spécifie l'utilisation coordonnée de standards développés par des organismes tels que HL7¹, CEN TC251² ou DICOM³ pour l'échange d'informations cliniques lors de nombreux cas d'usage. Des projets Européens - tels que epSOS- Smart Open Services for European Patients[2] et Artemis[3] - ont permis de mettre en place des plateformes internationales de partage d'information reposant sur l'utilisation des standards. Malgré ces efforts, l'interopérabilité sémantique des SIS reste encore limitée [1]. L'obstacle majeur reste la difficulté que les éditeurs des SIS rencontrent pour développer et maintenir des solutions permettant l'échange d'informations cliniques selon les standards et surtout d'adapter ces solutions de collecte de données standardisées à l'usage quotidien des professionnels de la santé.

2.1 Référentiels d'interopérabilité sémantique de données de santé

Dans le domaine de la santé, des organismes de standardisation tels qu'HL7, le CEN TC251 ou DICOM définissent des modèles d'information de SIS (messages ou documents) et des organismes tels qu'IHTSDO ou l'OMS définissent des systèmes terminologiques de référence (terminologies, systèmes de codage ou ontologies). Les terminologies de référence sont définies comme des « terminologies conçues pour apporter une représentation complète et

¹ <http://www.hl7.org/>

² <http://www.cen.eu/CEN/Sectors/TechnicalCommitteesWorkshops/CENTechnicalCommittees/Pages/Standards.aspx?param=6232&title=CEN/TC+251>

³ <http://medical.nema.org/>

exacte des concepts d'un domaine donné et de leurs relations et qui sont optimisées pour la classification et la recherche de données cliniques » [4].

Les modèles d'information de santé actuellement proposés par les organismes de standardisation reposent sur des principes communs qui sont i) une modélisation à plusieurs niveaux d'abstraction avec la possibilité de définir des modèles spécifiques de contexte d'usage, ii) une modélisation commune des types de données de santé et iii) des règles définissant la manière d'utiliser les systèmes terminologiques (terminologies, système de codage, ontologies, etc.) lors de l'instanciation de ces modèles – propriété communément désignée par l'expression *terminology binding* [5].

2.1.1 Modélisation à plusieurs niveaux d'abstraction

Ainsi en ce qui concerne les modèles d'information de santé définis dans le cadre de la version 3 d'HL7, un modèle de référence générique (*Reference Information Model* d'HL7 version 3) est spécialisé afin de définir des modèles spécifiques de contextes d'usage tel que le modèle d'architecture de document clinique (*Clinical Document Architecture (CDA)*). Le CDA est le modèle clinique détaillé d'HL7 version 3 le plus utilisé dans les établissements de santé.

2.1.2 Types de données standard

Les modèles standard d'information en santé – qu'il s'agisse d'HL7 ou des modèles CEN TC251- intègrent les modèles communs de types de données définis par la norme ISO 21090:2011 « Types de données harmonisées pour une interchangeabilité d'informations ».

Cette norme standardise la sémantique de types de données de santé (*e.g.* quantité physique ou donnée codée associée à un jeu de valeurs codées et éventuellement ordonnées).

2.1.3 « Terminology binding »

L'association entre modèle d'information et terminologies est spécifiée au niveau d'« éléments de donnée » qui constituent le plus petit élément d'information au sein des modèles standard. Le standard ISO/IEC 11179-3:2013 « Registres de métadonnées » est de plus en plus utilisé dans le domaine de la santé afin de partager des définitions non ambiguës d'« éléments de données » réutilisables faisant référence à des concepts de systèmes terminologiques. S'il est communément admis que l'ensemble des éléments de données cliniques ne peut être défini in extenso une fois pour toute et pour tous les usages par les organismes de standardisation, un certain nombre d'initiatives ont constitué des ensembles d'« éléments de données » cliniques standardisés et codés selon des Terminologies de Référence (TR) partagées.

Ainsi, les modèles d'information de santé standard actuels - tels que les modèles HL7 version 3 – permettent de formaliser la façon dont l'information est structurée et codée. Cependant, ce sont les modèles HL7 version 2 – moins formels – qui restent les standards d'échange d'information clinique les plus implémentés au monde et le plus souvent préconisés par IHE. De plus, si les établissements de santé adoptent de plus en plus largement ces modèles HL7 v2 pour les échanges, la collecte de l'information au sein de SIS repose sur des modèles et des terminologies locaux.

2.2 Modèles locaux et terminologies d'interface

Les établissements de santé développent le plus souvent des modèles d'information clinique locaux prenant en compte de façon évolutive les caractéristiques organisationnelles et cliniques locales. L'information clinique au sein de ces modèles est codée selon des Terminologies d'Interface (TI). Les terminologies d'interface sont définies par Rosenbloom al. comme « une collection systématique de phrases (termes) du domaine de la santé définies afin de faciliter la saisie d'information par les utilisateurs au sein des SIS » [6]. Les terminologies d'interface sont construites pour des utilisateurs et des usages spécifiques et représentent une solution de flexibilité par rapport aux problèmes d'incomplétude et de lenteur de mise à jour des terminologies de référence. Ces terminologies d'interface doivent être alignées à des terminologies de référence afin de permettre le partage et le traitement de l'information clinique [6-7-8].

2.3 Plateforme d'interopérabilité sémantique

Le partage ou l'échange de données sémantiquement annotées s'inscrit dans la problématique plus générale de la mise en correspondance de schémas [9]. La mise en correspondance de schémas consiste à prendre deux schémas en entrée et à construire en sortie un ensemble de correspondances entre les éléments des deux schémas. Notre travail s'inscrit dans le cadre des approches de médiation [10]. Nous nous sommes particulièrement intéressés aux travaux d'intégration de données guidés par une ontologie [9-10-11-12] et notamment à l'approche de type *global as view* dans laquelle une ontologie globale est utilisée comme source de médiation, chaque source de données alignant ses données à cette représentation pivot.

3 Matériel et méthodes

La méthode utilisée afin d'atteindre notre objectif comporte trois étapes : i) modélisation des processus, analyse des besoins et identification des référentiels, ii) spécifications fonctionnelle et technique de la plateforme d'interopérabilité sémantique et iii) formalisation des ressources sémantiques (modèles d'information standard, TIs, TRs et alignements terminologiques) et expérimentation de la plateforme.

3.1 Modélisation des processus et analyse des besoins

La modélisation des processus et d'analyse des besoins est réalisée en deux étapes. La première étape est la modélisation des processus d'échange d'information inter-établissements lors de la gestion des examens. Cette étape consiste à analyser et formaliser les activités réalisées lors de la prescription et du retour des résultats dans les domaines de la biologie, l'imagerie et l'ACP. La deuxième étape est l'identification des référentiels, c'est à dire des modèles d'informations et terminologies de référence sur lesquels se basent les échanges d'information. A chaque fois que les cas d'usage concernés entrent dans le champ du cadre d'interopérabilité de l'ASIP Santé, les ressources sémantiques de référence ont été définies en conformité avec ce cadre d'interopérabilité. Enfin, les besoins d'alignement entre TIs et TRs ont été analysés. Il a été convenu que les SIS des établissements partenaires devaient être en capacité d'extraire les informations à échanger et de les structurer conformément au modèle d'information de

référence partagé mais que l'information clinique au sein de messages structurés de façon standard pouvait être codée en utilisant des TIs locales.

3.2 Architecture de la plateforme d'interopérabilité sémantique et intégration à l'infrastructure d'échange

L'étape de spécification fonctionnelle et technique de la plateforme d'interopérabilité sémantique a consisté à i) spécifier les fonctionnalités de la plateforme en ce qui concerne la gestion des différentes ressources sémantiques (ressources locales et de référence) et des alignements entre ces ressources et ii) définir les services sémantiques requis lors de l'exécution des processus inter-établissements. La plateforme d'interopérabilité sémantique est fondée sur un serveur central, des serveurs locaux situés au niveau de chaque hôpital partenaire et un ensemble de services sémantiques. Ainsi, le serveur central gère les différentes versions de référentiels d'interopérabilité partagés (*i.e.* modèles d'information et TRs) et assure la distribution des TRs au niveau des différents serveurs locaux. Les serveurs locaux gèrent les TIs locales et leurs alignements avec les TRs partagées.

La plateforme doit fournir les services sémantiques permettant de réaliser les transformations des informations échangées lors de l'exécution des processus inter-établissements. Il s'agit de mettre en œuvre des services de transcodage entre TIs et TRs permettant de transcoder l'information clinique au sein des messages ou documents échangés selon le modèle de référence partagé servant de modèle pivot.

La Figure 1 présente l'architecture de la plateforme pour la gestion des ressources sémantiques au sein d'un serveur central et des serveurs locaux et illustre l'utilisation de services de transcodage lors d'un scénario d'échange d'information entre deux établissements (CHU de Rouen et l'AP-HP) dans le cadre d'une demande d'avis expert en ACP⁴. Le serveur central (TeRSan) gère les différentes versions de référentiels d'interopérabilité partagés (3) et les serveurs locaux (A (2) et B (4)) les TIs locales et leurs alignements avec les TRs partagées. Dans l'exemple de la demande d'avis en télépathologie, les services de transcodage permettent aux établissements A et B d'échanger des informations cliniques standardisées (hypothèse diagnostique codée en SNOMED-CT (3)) tout en continuant à utiliser leurs TIs locales (par exemple ADICAP (1) et CIM-O (4)).

⁴ Dans le cadre du profil d'intégration IHE Inter Laboratory Workflow (ILW) de sous-traitance d'examen de laboratoire. http://www.ihe.net/uploadedFiles/Documents/Laboratory/IHE_LAB_Suppl_ILW.pdf

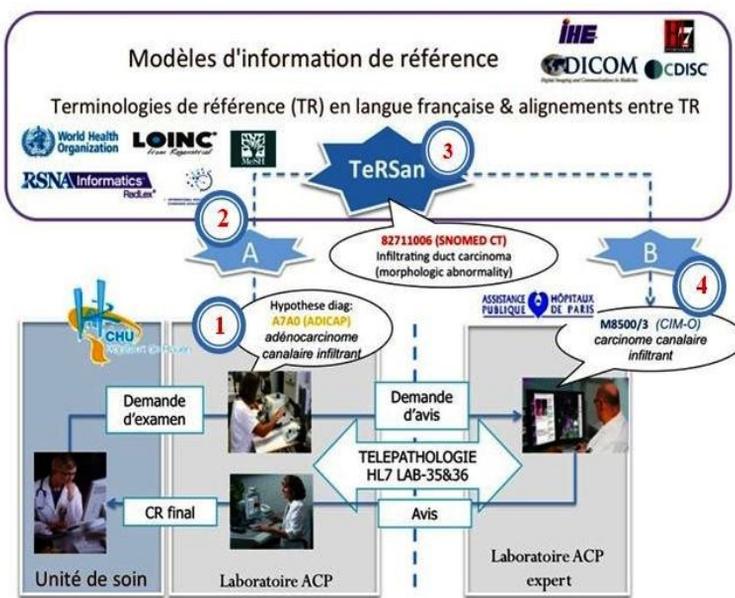


Figure 1 : Architecture de la plateforme de gestion des ressources sémantiques

3.3 Formalisation des ressources d'interopérabilité sémantique

3.3.1 Modèles d'information et terminologies

Cette étape consiste à formaliser les référentiels d'interopérabilité sémantique standard utilisés dans les différents échanges inter-établissements. En pratique, il s'agit de formaliser les modèles de message HL7 version 2 et notamment les modalités de codage des informations au sein de ces modèles avec des TRs dans des contextes d'échange d'information spécifiques. Cette étape a été réalisée en utilisant l'application termApp, espace collaboratif et intuitif de conception de modèles d'information en santé⁵.

3.3.2 Alignement entre terminologies d'interface et terminologies de référence

Cette étape consiste à mettre en correspondance les TIs locales avec les TRs et s'appuie sur des outils d'alignement semi-automatique soit proposés par la plateforme centrale tels qu'ITM (Intelligent Topic Manager) Match⁶ – soit déployés au sein des serveurs locaux des hôpitaux. A l'AP-HP, ITM Match permet de réaliser les alignements selon 3 algorithmes linguistiques de mise en correspondance. Au CHU de Rouen, un outil d'alignement automatique [13] a été utilisé et les résultats validés manuellement [6]. Lors de cette étape, les limitations des TRs par rapport à l'usage ont conduit à définir des besoins d'extension et/ou d'actualisation et /ou de traduction de ces TRs.

⁵ <http://termapp.davidouagne.com/>

⁶ <http://www.mondeca.com/fr/Produits/Intelligent-Topic-Manager>

3.4 Intégration des ressources et services d'interopérabilité sémantique

La plateforme logicielle (outils et des services) centrale de gestion des référentiels est utilisée pour permettre la création, la mise à jour et le partage des ressources sémantiques utilisées au sein de la communauté d'échange (figure 1). Les ressources sémantiques – centrales et locales - produites lors des étapes précédentes ont été intégrées au sein des serveurs locaux des établissements de santé partenaires. Dans ce travail, l'outil de gestion de référentiels ITM, est utilisé au niveau du serveur central et au niveau du serveur local à l'AP-HP. Pour le CHU de Rouen, la gestion des terminologies est confiée au Portail Terminologique de Santé (PTS)⁷.

4 Résultat

4.1 Modélisation des processus et analyse des besoins

L'analyse des processus a permis de constater que, quel que soit le domaine considéré, le cycle prescription-exécution-retour de résultat repose sur le processus générique synthétique suivant: création du message ou document de demande par le composant de prescription du SIS de l'établissement de santé source de la demande (acteur IHE Order Placer), échange de message standard (message HL7 version 2), exécution de la demande par le composant d'exécution du SIS de l'établissement de santé destinataire de la demande (acteur IHE Order Filler), facturation (hors domaine d'intérêt), création du message ou document de résultat par le composant d'édition de résultat ou compte rendu SIS de l'établissement de santé source du résultat ou compte rendu (acteur IHE Order Filler/Report Creator), échange de message standard (message HL7 version 2).

Pour chaque domaine, un responsable a recensé l'ensemble des modèles standards de messages ou de documents et des TRs utilisées ou utilisables lors des échanges, et a proposé, en collaboration avec les experts métiers des deux sites pilotes (AP-HP et CHU de Rouen), des solutions d'interopérabilité entre les deux sites. Chaque site utilise en général des TIs spécifiques, le partage ou l'échange d'information (message ou document) nécessite que la TI utilisée par les composants soit mise en correspondance avec la TR partagée spécifiée pour le message ou document.

4.2 Architecture d'échange et de déploiement

L'architecture d'échange présentée dans la Figure 2 se base sur les plateformes d'Intégration d'Applications d'Entreprise (IAE) et le Système de Messagerie Sécurisé (SMS) implémentés dans les établissements cliniques.

⁷ <http://www.hetop.fr>

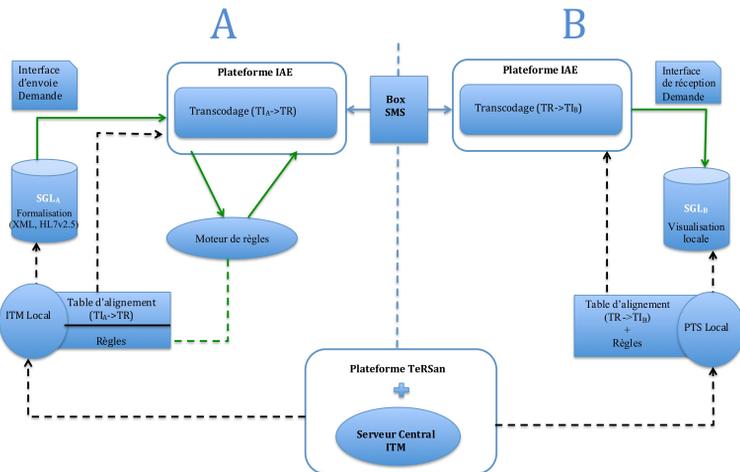


Figure 2 : Architecture du prototype d'échange d'information entre établissements de santé et utilisation des services sémantiques des serveurs locaux.

Les IAE prennent en charge la réception des messages, le routage applicatif et la transformation des messages à destination d'hôpitaux destinataires. Le SMS assure la transmission inter-établissements des messages et prend en charge la sécurité des échanges.

Les services de transcodage de la plateforme d'interopérabilité sémantique sont utilisés par la plateforme IAE lors de l'étape de transformation des messages. Les IAE fournissent un environnement de développement pour le traitement des messages HL7 v2.5. Ce traitement consiste à transcoder le contenu des messages selon des règles de transcodage en utilisant des ressources d'alignements entre TIs et TRs gérées au sein des serveurs de terminologies locaux. Les règles permettent d'informer l'IAE sur la/les ressources d'alignement à utiliser en fonction du contexte (i.e. biologie, imagerie ou ACP) et des informations véhiculées dans le message (i.e. examen demandé, région anatomique, informations cliniques, etc).

Dans l'exemple de la [figure 1](#), la valeur de l'information clinique « Hypothèse diagnostique » est codée localement en ADICAP au niveau du Système de Gestion de Laboratoire (SGL) du laboratoire ACP demandeur (1), transcodée au niveau du serveur local (A) avant envoi (2) en SNOMED CT - TR définie pour ce type d'information dans ce contexte d'échange (3) – puis, après réception par le laboratoire ACP expert, transcodée en CIM-O – TI utilisée localement au niveau du SGL du laboratoire ACP destinataire - au niveau du serveur local (B) (4) avant d'être intégrée au SGL du laboratoire.

4.3 Formalisation des ressources d'interopérabilité sémantique

4.3.1 Templates et terminologies

La construction de référentiels partagés (modèles d'information de message HL7 version 2 et TRs) a été initiée dans les trois d'application. En ce qui concerne l'ACP, domaine choisi pour la mise en œuvre du premier prototype, nous avons formalisé un modèle de demande d'avis par

télépathologie. Le modèle de demande d'avis a été formellement modélisé en fonction du modèle de message HL7 version 2 adapté au contexte des demandes de sous-traitance inter-établissements⁸.

L'organisation hiérarchique du message HL7 utilisé dans le cadre de la transaction de sous-traitance entre laboratoires a été modélisée (voir Figure 3). Ce message permet de véhiculer l'information au sein de champs organisés en segments. Les champs des différents segments contiennent les informations relatives au message lui-même, au patient, aux éléments d'assurance intervenant dans la facturation et à la demande de sous-traitance elle-même. Les informations de demande de sous-traitance sont constituées d'informations générales de la demande, d'observations cliniques pertinentes dans le contexte de la demande, d'informations relatives aux prélèvements associés à la demande et d'informations relatives à des demandes d'examens ou observations antérieures pertinentes dans le contexte de la demande de sous-traitance.

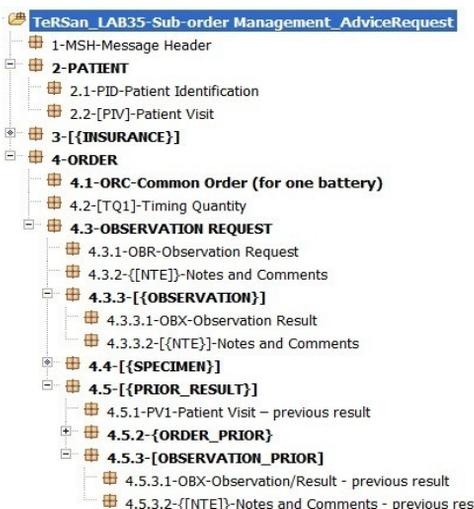


Figure 3 : Organisation hiérarchique du message HL7 v2.5.1 OML^O21

Une deuxième étape de modélisation nous a permis de spécialiser le modèle HL7 de demande de sous-traitance au contexte de la demande d'avis en télépathologie. Lors de cette étape, des observations spécifiques à ce contexte ont été modélisées, au sein des segments OBX, sous forme d'« éléments de données » en se basant sur le standard ISO/IEC11179-3:2013. Chaque « élément de donnée » *Attribute Code* (champ 3 du segment OBX (OBX-3)) des observations modélisées a été associé un concept médical issu d'une TR du domaine (LOINC ou SNOMED CT) et son domaine de valeurs (*Attribute Value* (OBX-5)) a été formalisé en se fondant sur le standard ISO 21090:2011. Lorsque la valeur de l'observation est codée (type de données de l'*Attribute Value* (OBX-5) est *Coded Element* (CE) ou *Coded With Exception* (CWE)) chacune des valeurs possibles du domaine de valeur a été explicitement associée à un concept médical issu d'une TR du domaine (SNOMED CT) (voir Figure 4).

⁸ Message HL7 v2.5.1 OML021 http://www.hl7.org/implement/standards/product_brief.cfm?product_id=144, conformément à la transaction LAB-35 du profil IHE ILW

Au sein de chaque établissement de santé, les champs du modèle de message instanciés par des informations codées en utilisant des TIs ont été identifiés. Pour chaque champ identifié, une règle de transcodage a été explicitée.

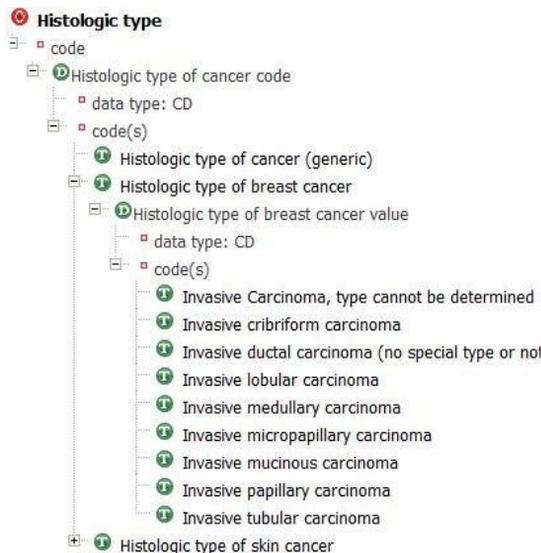


Figure 4 : Modèle de l'observation Hypothèse diagnostique (type histologique).

4.3.2 Alignement entre terminologies d'interface et terminologies de référence

L'alignement entre TIs et TRs a été initié dans les domaines de la biologie (TI de demande et de résultat) et de l'imagerie (TI de demande). Ces travaux ont mis en évidence que la granularité des TRs n'était pas toujours adaptée aux besoins nécessaires en interface. Par exemple, en ce qui concerne les TIs de demandes d'analyses de biologie, les termes LOINC sont souvent trop précis par rapport à ce qui peut être demandé. Il est ainsi nécessaire de représenter une analyse demandée par un ensemble de codes LOINC susceptibles de correspondre à la réalisation de cette demande. A contrario, en ce qui concerne la TI de demandes d'actes d'imagerie, certains termes de la CCAM ont un grain plus gros que ce qui peut être attendu au niveau des demandes d'actes d'imagerie et différents termes de demandes peuvent être alignés au même terme de la TR. Dans le domaine de l'ACP, nous nous sommes intéressés dans un premier temps aux alignements des TIs et TRs utilisées pour représenter les observations ACP. Les termes de TI concernant les observations ont été alignés à LOINC ou PathLex. Les termes de TI concernant les valeurs possibles des observations codées ont été alignés à SNOMED CT ou PathLex. En ce qui concerne la valeur de l'information clinique « Hypothèse diagnostique », les TI locales utilisées sont ADICAP ou CIM-O et la TR SNOMED CT. Les ressources d'alignement ADICAP-SNOMEDCT et CIM-O-SNOMEDCT ont été intégrés respectivement dans les serveurs de l'AP-HP et du CHU de Rouen.

4.4 Intégration des ressources et services d'interopérabilité sémantique

La synchronisation entre le serveur central – gérant les modèles standard et les TRs – et les serveurs locaux - gérant les TIs et leurs alignements aux TRs - est réalisée via des services web. A chaque nouvelle version de TR, le différentiel entre versions est transmis aux serveurs locaux. Une réflexion est menée pour mettre en œuvre une solution facilitant la mise à jour des alignements entre TIs et TRs lors de l'intégration d'une nouvelle version de la TR.

Les services d'interopérabilité sémantique permettant le transcodage dynamique des informations à échanger avant l'envoi de messages ou documents au(x) destinataire(s) ont été spécifiés et sont en cours de développement. Le prototype d'échange d'information inter-établissements démontrant l'apport des services sémantiques lors d'échanges de télépathologie est en cours de mise en œuvre.

5 Discussion

La plateforme et les services d'interopérabilité sémantique proposés dans le cadre de ce travail permettront aux établissements de santé d'échanger des informations cliniques selon les standards tout en conservant au sein du SIS des solutions de collectes de données cliniques utilisant des TIs locales et respectant ainsi l'usage quotidien des professionnels de la santé. L'objectif de la plateforme TeRSan va donc au-delà des plateformes développée dans le cadre des projets Européens - tels que epSOS[2] ou Artemis[3] – se focalisant essentiellement sur la gestion de transformation de modèles entre modèles standard et sur le traitement l'information multi-lingue au sein de TRs. Ces projets n'abordent pas la gestion des TIs, ni les problématiques d'alignement entre TIs et TRs. Par ailleurs, en ce qui concerne le projet epSOS, les solutions d'interopérabilité sémantiques mises en place concernent un périmètre restreint au partage de documents cliniques correspondant à trois modèles précis de CDA et n'abordent la le passage à l'échelle des solutions proposées grâce à des règles de transcodage génériques.

En ce qui concerne le travail réalisé, les difficultés essentielles ont été rencontrées lors de l'étape d'alignement des terminologies. Plusieurs pistes d'optimisation de cette étape sont envisageables. Dans un premier temps, en ce qui concerne les TRs multi-axiales telles que LOINC ou la CCAM, il est possible d'exploiter les informations contenues au sein des axes afin d'améliorer la performance des algorithmes d'alignement semi-automatique. Même si l'utilisation axiale de ces terminologies n'est pas prévue par leurs créateurs, les limites des termes pré-coordonnés imposent l'utilisation du vocabulaire axial, plus riche. Une deuxième solution consisterait, dans les situations où il n'existe pas de TR permettant un double transcodage satisfaisant – c'est-à-dire sans perte d'information lors des échanges - à considérer d'autres terminologies en tant que TR. Si LOINC reste actuellement la TR indiscutable dans le domaine de la biologie, dans le domaine de l'imagerie en revanche, RadLex⁹ représente une alternative internationale à la CCAM, tout à fait crédible. Enfin, si aucune terminologie existante n'est susceptible de répondre aux besoins, il serait nécessaire d'enrichir les TRs existantes (à un rythme toutefois plus soutenu que celui proposé actuellement par les institutions les gérant). Enfin, afin de faciliter la mise en œuvre de processus de prescription connectée au sein des établissements, la création de TIs consensuelles partagées alignées aux TRs doit être considérée. L'usage de la TR pivot aurait lieu d'être alors essentiellement dans les

⁹ <http://www.rsna.org>

échanges au niveau international.

La gestion du transcodage par les plateformes IAE au sein des établissements de santé est envisageable tant que les traitements à mettre en œuvre sont relativement simples. En effet, l'IAE est en charge de la transmission de tous les messages inter-applicatifs des établissements et il n'est donc pas possible de la saturer avec des messages nécessitant des traitements complexes tels que l'exécution de certaines règles de transcodage. Ainsi, la mise en œuvre d'un moteur de règle interfacé à l'IAE et au serveur de terminologie est nécessaire et Drools¹⁰ s'avère être une solution intéressante.

6 Conclusion

Dans ce travail, nous avons présenté une plateforme d'interopérabilité sémantique permettant la standardisation de l'information clinique échangée tout en respectant les usages des professionnels de santé qui continuent à utiliser les interfaces de saisie adaptées à leur pratique quotidienne.

Il s'agit de contribuer ainsi, dans le respect de la gouvernance de l'ASIP Santé¹¹, à la mise en œuvre d'une stratégie de production, de maintenance et de distribution durables de référentiels nationaux d'interopérabilité sémantique. Sur le plan opérationnel, ce travail démontre dans le domaine de la prescription d'actes que la synchronisation entre SIS de référentiels incluant des TIs permettra l'exécution facile de processus inter-établissements. Concernant les travaux en cours, ce travail démontrera aussi que la saisie de données de santé conformément aux référentiels permettra, dans chacun de ces établissements, leur exploitation standardisée pour l'accès personnalisé à des connaissances médicales, à des recommandations ou à des essais cliniques en cours. De plus des TIs partagées consensuelles pour les trois domaines considérés seront proposées et la prise en compte des spécificités locales au sein de ces TIs partagées sera gérée par des synonymes.

Des règles d'alignement entre TIs et TRs seront définies en s'appuyant sur le guide d'implémentation de la SNOMED CT et les travaux du groupe de travail HL7 TermInfo. Ces règles d'alignement sont des règles de formalisation d'expressions complexes des TIs (pré-coordonnées) avec les concepts d'une TR (*e.g* SNOMED).

D'un autre côté, la gestion des mises à jour des TIs et TRs ainsi que l'impact de ces mises à jour sur les alignements est une problématique très importante qui sera gérée au sein de la plateforme. Le serveur central gèrera la synchronisation des mises à jours des TRs au sein des serveurs locaux et l'impact sur les alignements sera géré au niveau des serveurs locaux.

¹⁰ <https://www.jboss.org/drools/>

¹¹ <http://esante.gouv.fr/asip-sante>

Remerciements

Ces travaux ont été financés par l'Agence National de la Recherche, programme Technologie pour la Santé, dans le cadre du projet TeRSan (Terminologie et Référentiels d'interopérabilité en Santé) ANR-11-TECS-019. <http://www.chu-rouen.fr/tersan>. Nous remercions Sylvie Cormont, Christophe André, Vincent Galais, Déa Giardella, Naémé Nekooguyan, Lydia Bascarane, Florence Amardeilh et Julien Grosjean pour leur contribution dans le cadre du projet TeRSan et de ce travail.

Références

- [1] European Commission Information Society and Media, "Semantic interoperability for better health and safer healthcare," Semantic Health Report, Research and Deployment Roadmap for Europe, Jan. 2009.
- [2] "epSOS: Home." [Online]. Available: <http://www.epsos.eu/home.html>. [Accessed: 26-May-2014].
- [3] A. Dogac, G. B. Laleci, S. Kirbas, Y. Kabak, S. S. Sinir, A. Yildiz, and Y. Gurcan, "Artemis: Deploying Semantically Enriched Web Services in the Healthcare Domain," *Inf Syst*, vol. 31, no. 4–5, pp. 321–339, juin 2006.
- [4] S. T. Rosenbloom, S. H. Brown, D. Froehling, B. A. Bauer, D. L. Wahner-Roedler, W. M. Gregg, and P. L. Elkin, "Using SNOMED CT to Represent Two Interface Terminologies," *J. Am. Med. Informatics Assoc. Jamia*, vol. 16, no. 1, pp. 81–88, 2009.
- [5] A. L. Rector, R. Qamar, and T. Marley, "Binding ontologies and coding systems to electronic health records and messages," *Appl. Ontol.*, vol. 4, no. 1, pp. 51–69, Jan. 2009.
- [6] S. T. Rosenbloom, R. A. Miller, K. B. Johnson, P. L. Elkin, and S. H. Brown, "Interface Terminologies: Facilitating Direct Entry of Clinical Data into Electronic Health Record Systems," *J. Am. Med. Informatics Assoc. Jamia*, vol. 13, no. 3, pp. 277–288, 2006.
- [7] F. Bakhshi-Raiez, L. Ahmadian, R. Cornet, E. de Jonge, and N. F. de Keizer, "Construction of an interface terminology on SNOMED CT. Generic approach and its application in intensive care," *Methods Inf. Med.*, vol. 49, no. 4, pp. 349–359, 2010.
- [8] N. Griffon, C. Savoye-Collet, P. Massari, C. Daniel, and S. J. Darmoni, "An interface terminology for medical imaging ordering purposes," *Amia Annu. Symp. Proc. Amia Symp. Amia Symp.*, vol. 2012, pp. 1237–1243, 2012.
- [9] A. Doan and A. Y. Halevy, "Semantic integration research in the database community: A brief survey," *Ai Mag.*, vol. 26, pp. 83–94, 2005.
- [10] G. Wiederhold, "Mediators in the architecture of future information systems," *Computer*, vol. 25, no. 3, pp. 38–49, Mar. 1992.
- [11] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hübner, "Ontology-Based Integration of Information - A Survey of Existing Approaches," presented at the In IJCAI-01 Workshop: Ontologies and Information Sharing, 2001, pp. 108–117.
- [12] Y. Kalfoglou and M. Schorlemmer, "Ontology Mapping: The State of the Art," *Knowl Eng Rev*, vol. 18, no. 1, pp. 1–31, Jan. 2003.
- [13] N. F. Noy, "Semantic Integration: A Survey Of Ontology-Based Approaches," *Sigmod Rec.*, vol. 33, p. 2004, 2004.

- [14] J. Euzenat and P. Shvaiko, *Ontology Matching*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2007.

Adresse de correspondance

Hôpital Rothschild, pavillon 3

5, rue de santerre, 75012,

Paris, France

amina.chniti@sap.aphp.fr

Intégration des données de prescription dans un entrepôt de données biomédicales

Integration of prescription data in a clinical data warehouse

Kevin Dalleau¹ Denis Delamarre^{1,2}
Thibault Ledieu^{1,2} Marc Cuggia^{1,2}

¹Laboratoire Traitement du Signal et de l'Image (LTSI) INSERM : U1099 – Université de Rennes 1

²CHU Rennes, Pôle Pharmacologie et Investigation cliniques, Unité « Fouille de données », F35033 Rennes, France

Résumé

L'intégration de données relatives au médicament au sein d'un entrepôt de données biomédicales et leur réutilisation est d'une importance majeure, les prescriptions étant au centre de la prise en charge du patient au cours de séjour. En utilisant une base de données médicamenteuse indépendante, Thériaque, et une définition standardisée des prescriptions et plus généralement du circuit de médicament, la CIOsp (Classification InterOpérable des spécialités), nous avons implémenté une nouvelle facette à notre moteur de recherche, Roogle. Nous avons permis à l'utilisateur de rechercher la prescription d'un médicament, soit via la sélection de code ATC, soit directement via l'utilisation d'un formulaire de recherche avec autocomplétion et gestion des doublons. Bien que beaucoup de travail reste à faire, ces développements ouvrent la voie à une toute nouvelle manière de constituer des cohortes en vue de réaliser des essais cliniques, ou encore de trouver de potentiels effets indésirables liés à une molécule.

Abstract

Integration of drugs-related data into a clinical data warehouse and its good use is a key issue, prescriptions being at the center of patient care in a hospital. Using an independent drug-database, Thériaque, and a standardized definition of prescriptions and medications - the CIO (InterOperable Classification), we were able to implement a brand new facet to Roogle, our search engine. We enabled the user to search for a drug prescription either by selecting corresponding ATC codes, or directly with a search form including autocompletion and duplicates management. Although a lot of work is still to be done, these developments pave the way to whole new ways to create cohorts of patients for clinical studies, or to find potential adverse effects related to a specific drug.

Mots-clés : Informatique médicale ; Conférence ; Entrepôt de données ; Médicaments ;

Prescription ; Big data

Keywords: Medical Informatics ; Conference ; Data Warehouse ; Drugs, Prescriptions ; Big

Articles longs des 15^{es} Journées francophones d'informatique médicale, JFIM 2014, pages 59–70

Fès, Maroc, 12–13 juin 2014

1 Introduction

Le domaine médical connaît à l'heure actuelle une massification des données en santé (notion de Big Data)[1]. Ce gisement de données peut être réutilisé pour soutenir la recherche. Dans ce but, les établissements, en particulier les CHU, sont en train de se doter d'entrepôts de données biomédicales, dont l'objectif est d'intégrer et d'exploiter les données issues du système d'information hospitalier dans un objectif d'interrogation populationnelle. Parmi les données importantes exploitées dans ces entrepôts, celles concernant le médicament représentent plusieurs défis tant sur les plans de leur intégration que de leur exploitation.

Premier défi: gérer l'hétérogénéité des sources contenant des informations sur le traitement des patients. En effet, dans un SIH, ces informations peuvent être contenues dans des documents textuels (compte rendus d'hospitalisation, courriers), mais aussi dans des documents structurés, depuis les modules de prescriptions des DPI où cette information est le plus souvent codée selon une norme sémantique, ou encore dans des formulaires structurés de dossiers de spécialité. Il s'agit donc de pouvoir intégrer ces données hétérogènes dans un même entrepôt de données, et de permettre leur exploitation.

Deuxième défi: valoriser l'exploitation sémantique de ces données hétérogènes. Pour cela, l'interrogation des données sur le médicament doit s'appuyer sur un modèle de représentation sémantiquement riche, qui permette par exemple de rechercher des patients ayant pour traitement une classe médicamenteuse spécifique (par exemple, les patients sous fluoroquinolone), ou ayant un traitement indiqué dans une pathologie particulière (les patients ayant un médicament indiqué dans la chimiothérapie du sein, par exemple).

Pour cela, l'utilisation d'une base de connaissance sur le médicament devient incontournable.

Le travail présenté ici vise à exposer l'approche d'intégration et de recherche d'information concernant le médicament au sein de l'entrepôt de données biomédicales du CHU de Rennes (système Roogle).

2 État de l'art

À l'heure actuelle, le projet le plus avancé dans l'interrogation et l'exploitation d'entrepôts de données biomédicales est celui porté par le National Institute of Health (NIH): I2B2 (Informatics for Integrating Biology and the Bedside). Il s'agit d'une plateforme open source dont l'objet est l'exploitation de données stockées au sein d'entrepôts de données biomédicales, initialement sur des données de génétique [2, p. 2].

I2B2 permet d'effectuer des recherches sur le médicament, et ce sous de nombreux angles (voie, spécialité commerciale, dose, etc.).

Cependant :

- I2B2 se base sur des données, référentiels et thésaurus propres aux Etats-Unis (RxNorm, nom de spécialités spécifiques), le rendant peu utilisable en France
- Les flux PN13 [PréNorme 13, majoritaires en France] ne sont pas pris en charge par la solution

Nous proposons, par l'ajout de cette fonctionnalité à Roogle, une exploitation des données liées

au médicament se basant sur un standard d'interopérabilité vivant et largement adopté en France, le standard PN13 – SIPH2, et s'appuyant sur une base de connaissances médicamenteuse puissante et exhaustive, Thériaque.

3 Matériel et méthodes

3.1 Matériel

3.1.1 *Roogle*

Projet débuté au sein de CHU de Rennes en 2007, Roogle était à ses débuts un système de recherche d'information au sein du dossier d'imagerie médicale, gérant à la fois la recherche sur des données structurées et sur des données plein-texte[3].

L'outil s'est au fil des années enrichi de nombreux autres flux, tels que les comptes-rendus hospitaliers, les actes réalisés, les comptes-rendus d'anatomopathologie, etc. Les champs d'applications sont nombreux, allant de la constitution de cohorte pour des fins de recherche à l'évaluation des pratiques professionnelles, en passant par la recherche de cas spécifiques.

À ce jour, le système exploité à Rennes contient plus d'un million de patients et près de 15 millions de documents.

3.1.2 *Base Thériaque*

Editée par le Centre National Hospitalier d'Information sur le médicament (CNHIM), Thériaque constitue l'une des plus grandes bases de données médicamenteuses en France. Elle regroupe de manière exhaustive toutes les informations sur les médicaments commercialisés en France, en puisant l'information de sources officielles (agence du médicament, journaux officiels, avis de la HAS, etc.), complétées d'informations issues d'ouvrages de référence. [4]

Ces données sont très structurées, permettant leur exploitation et leur intégration dans des applications tierces. Celles-ci sont par ailleurs mises à jour quotidiennement.

La base nous est fournie par le CNHIM sous de nombreux formats. L'utilisation des informations qu'elle contient permet de répondre à de nombreuses questions concernant les médicaments, allant de leur code ATC à leurs indications.

L'interopérabilité est assurée via l'intégration des fichiers UCD. Fournis par le CIP (Club Inter Pharmaceutique), le code UCD est un code numérique à 7 chiffres, correspondant à la plus petite unité de dispensation (gélule, comprimé, etc.).

3.1.3 *Le standard PN13 – SIPH2*

Il s'agit d'un modèle du circuit du médicament, partagé par les différentes applications du système d'information du circuit, qui définit les flux d'informations à émettre à chaque transmission. Le standard est constitué de 3 éléments indissociables :

- Le référentiel d'architecture technique, dont l'objectif est de décrire les messages qui transitent entre les systèmes : messages de prescriptions, de dispensation nominative, etc.

- Le référentiel des nomenclatures, contenant l'ensemble des nomenclatures permettant de coder les attributs des éléments définis dans le référentiel d'architecture technique.
- La Classification InterOpérable des spécialités (CIOsp), composée d'un noyau contenant les informations de base (code UCD, nom commercial, classe ATC, etc.), et de composants additionnels.

Le standard garantit l'interopérabilité, permettant aux applications l'utilisant de communiquer et de se comprendre.

La distribution consiste en une archive contenant toutes les données de la CIOsp sous forme de fichiers texte, chaque fichier correspondant à une table du modèle physique de données.

3.1.4 Données de prescription issues de DxCare

Le CHU de Rennes a choisi DxCare ® (Medasys) comme solution au sein de son système d'information. Elle permet, via l'existence de plusieurs modules, de gérer de nombreux aspects du fonctionnement hospitalier : mouvement, séjour, dossier médical, planification des rendez-vous, prescriptions, etc.

DxCare ® est "CIO-inside", tout comme de nombreuses autres solutions du marché (Pharma ®, Géniois ®, etc.). Ainsi, l'ensemble des prescriptions réalisées au sein de l'établissement transite au format PN13. L'architecture d'un élément de prescription dans ce format est présenté figure 1.

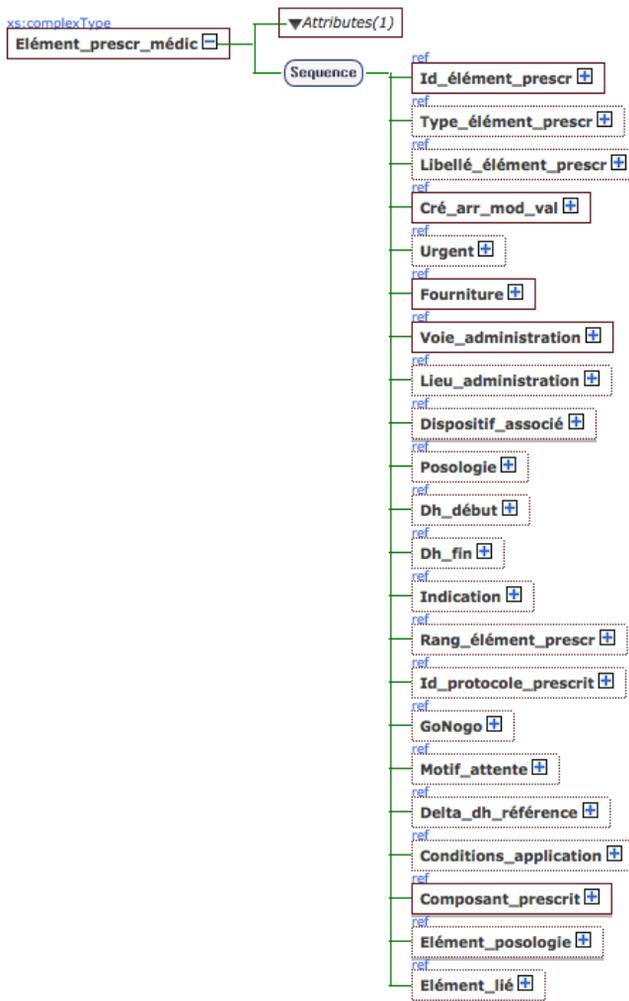


Figure 1 : Architecture d'un élément de prescription au format PN13

3.2 Méthodes

3.2.1 Cas d'usages adressés

L'intégration des prescriptions, et, plus largement, des médicaments à l'entrepôt de données répond à une demande croissante de la part des professionnels, dans plusieurs domaines (recherche clinique, épidémiologie, professionnels de la pharmacovigilance, etc.).

Afin d'identifier les besoins et la méthode d'intégration et de recherche d'information sur ce domaine, nous définissons les cas d'usages de Roogle suivants.

- Constitution de cohorte :

La recherche de patients ayant reçu un ou plusieurs traitement(s) spécifique(s) est l'un des éléments clés de la constitution de cohorte. L'intégration de la temporalité, rendue possible grâce à la haute structuration des flux de prescriptions, est par ailleurs un besoin très fort afin de mener de la manière la plus fine possible cette tâche, et de se rapprocher des contraintes du screening.

- Etude de faisabilité

Sur le même principe que la constitution d'une cohorte de patients, il est possible qu'un promoteur, industriel ou institutionnel, utilise Roogole pour déterminer les capacités d'inclusion de son site et d'un site investigateur. Contrairement à la phase de screening, les résultats de la requête ne porteront pas sur les informations individuelles de chaque patient mais sur les statistiques du site.

- Suivi des protocoles

Les protocoles ayant souvent des critères stricts devant être respectés, rendre possible la visualisation des différents évènements ayant eu lieu au cours de leur déroulement, notamment au niveau de la prise de médicaments, est aussi l'un des cas d'usage retenus.

- Recherche de potentiels effets indésirables

La prise en charge du médicament au sein de l'entrepôt de données permettra, en les croisant avec les données déjà présentes en son sein (courriers de sorties, résultats de laboratoire, etc.), de rechercher d'éventuels évènements consécutifs à la prise d'un médicament, et ce sur de grandes quantités de données.

3.2.2 Intégration

L'un des principes clés de Roogole est de conserver une structure de document la plus proche possible du document d'origine. Il s'agit de permettre à l'utilisateur qui fait une requête sur une information précise (la prescription d'un médicament par exemple) de retrouver lors de l'affichage des résultats le contexte dans lequel a été prescrit le médicament (en l'occurrence l'ordonnance). La conservation du document le plus proche du format natif dans l'entrepôt de données est aussi un moyen d'assurer la meilleure intégrité de l'information, au cas où celle-ci deviendrait opposable. C'est pour cette raison que les données de prescription sont intégrées au format natif PN13 dans l'une des tables, contenant déjà les autres documents précédemment intégrés à l'entrepôt (compte-rendus, etc.).

Afin de tirer parti au maximum du fait que les flux de prescriptions soient structurés, les données sont par ailleurs stockées dans une autre table dédiée, où cette fois ci chaque ligne correspond à une ligne de la prescription. Cette méthode permet de conserver la plus fine granularité, chaque ligne étant indépendante, et pouvant être retrouvée via ses différents éléments. De plus, cette deuxième intégration permet de faciliter une éventuelle agrégation future des résultats. Plusieurs lignes de cette table peuvent donc correspondre à une seule et même prescription, stockée entièrement dans la première table.

L'ensemble de la base de données médicamenteuse Thériaque est par ailleurs intégrée, afin qu'elle puisse être utilisée comme base de connaissance pour la construction de requêtes sémantiquement riches.

Pour illustrer l'utilisation de la base Thériaque au sein de Roogle, nous avons développé un script permettant de charger la classification ATC, à partir de la base Thériaque, dans une table dédiée aux Thésaurus. Ce modèle de représentation sous forme de chemin parcourant un arbre hiérarchique a déjà été utilisé pour l'exploitation d'autres données structurées qui utilisent des terminologies dédiées (comme la CIM-10 pour les diagnostics, et la CCAM pour les actes médicaux et chirurgicaux).

La dernière distribution en date de la CIOsp a enfin été intégrée, afin de pouvoir assurer le la bonne marche de fonctions spécifiques de décodage des documents au format PN13.

3.2.3 Recherche d'information et visualisation

La problématique de visualisation des données lors d'une recherche d'information est double :

- l'utilisateur doit pouvoir construire sa requête de façon intuitive et exhaustive,
- la visualisation des résultats doit permettre à l'utilisateur d'extraire l'information de façon simple et non ambiguë.

Le constructeur de requêtes de Roogle propose actuellement à l'utilisateur 2 types de recherche :

- Les recherches sur les informations structurées comme les données issues du PMSI,
- Les recherches plein texte (full-text) portant sur toutes les informations du Dossier Patient Informatisé, que cette information soit structurée ou non.

Dans le cas particulier de la recherche sur des prescriptions, l'utilisateur s'orientera plus généralement vers une recherche sur des données structurées afin de profiter de la structuration des données des flux de prescriptions. Le formulaire de recherche dans les données structurées préexistant sera donc augmenté de la classification ATC, afin de pouvoir y effectuer des recherches simplement. La recherche de prescription peut également se faire en plein texte sur les courriers médicaux comme sur les flux de prescriptions.

Le pivot de toute recherche est le code UCD. Ce code, assurant l'interopérabilité au sein de tous les systèmes partageants les informations sur le médicament, permet à notre système de rechercher directement au sein des documents de prescriptions, au sein desquels chaque ligne est identifiée par son code UCD et son code ATC.

La CIOsp est enfin très largement utilisée afin de traduire en clair les informations présentes au sein des documents. En effet, de nombreux éléments sont codifiés (type de substance, voie, forme, etc.), et il est nécessaire de les convertir en langage humain afin de simplifier la lecture. Des fonctions se chargeant du décodage sont donc utilisées, s'appuyant directement sur la distribution la plus à jour.

4 Résultats

4.1 Recherche et visualisation

4.1.1 Recherche via code ATC

L'intégration de la classification ATC au thésaurus permet de naviguer dans les branches de la classification, notamment dans le cadre de la recherche et de la sélection de substances, comme vu figure 2.

Devant chaque nœud de l'arborescence figure le nombre de patients ayant une prescription d'un médicament appartenant à ce nœud. L'utilisateur peut choisir un nœud comme critère de sélection (critère d'inclusion ou d'exclusion) et les combiner à d'autres critères.

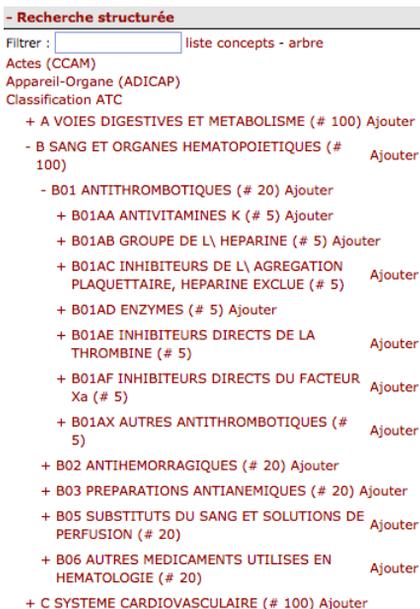


Figure 2 : Recherche et inclusion de médicaments par navigation dans la classification ATC

Le formulaire de recherche est composé :

- D'un arbre de concepts où chaque embranchement représente une sous-catégorie de l'ATC. Cet arbre peut être exploré manuellement par l'intermédiaire d'un menu dépliant.
- D'un champ texte permettant de filtrer les différents concepts présents dans les thésaurus intégrés, et de retourner uniquement ceux qui contiennent des patients, comme le montre la figure 3.

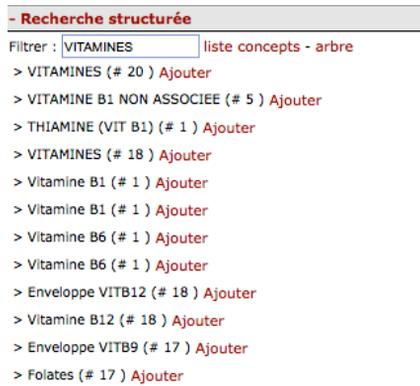


Figure 3 : Recherche de code par un moteur de recherche terminologique dans tous les thésaurus indexés.

4.1.2 Recherche via UCD

Roogle permet également de rechercher un médicament directement par son code UCD, comme le montre la figure suivante.

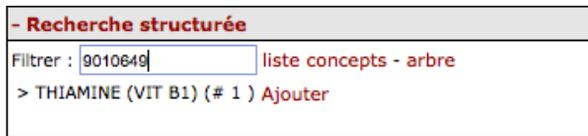


Figure 4 : Recherche via un code UCD. Ici, il est retourné la Thiamine, l'UCD recherché correspondant à celle d'une unité de dispensation de Bevitine®

4.1.3 Recherche par substance active

Une autre voie de recherche est explorée, permettant une recherche plus spécifique au médicament (figure 5).

Recherche par substance active

Commercialisation :

- Ne pas inclure les spécialités n'étant plus commercialisées
- Inclure les spécialités n'étant plus commercialisées

Gestion des principes actifs multiples :

- N'inclure que les spécialités contenant uniquement les principes actifs sélectionnés
 - Inclure toutes les spécialités, y compris celles contenant plusieurs principes actifs (dont ceux sélectionnés)
- Rechercher seulement une spécialité

Figure 5 : Recherche de spécialités par entrée de substances actives

Ce mode de recherche gère l'autocomplétion, et permet une gestion des doublons. Des fonctionnalités supplémentaires telles que la recherche ou non de spécialités n'étant plus commercialisées sont par ailleurs implémentées.

L'affichage des spécialités correspondantes se fait à la volée dans un bloc de droite, permettant à l'utilisateur de vérifier les résultats correspondant à sa sélection, et éventuellement à exclure certaines spécialités renvoyées (figure 6).



The screenshot shows a search bar with 'SARTAN' entered. A dropdown menu is open, listing several specialities: ENALAPRIL MALEATE, LOSARTAN POTASSIQUE (highlighted), VALSARTAN, IRBESARTAN, CANDESARTAN CILEXETIL, TELMISARTAN, and EPROSARTAN MESILATE. To the right, a list of specialities is displayed with checkboxes: RENITEC 20MG CPR, RENITEC 5MG CPR, ENALAPRIL BGA 5MG CPR, ENALAPRIL BGA 20MG CPR, and ENALAPRIL MYL 5MG CPR. All checkboxes are checked.

Figure 6 : Ajout de substance active à la recherche. Si une même substance est sélectionnée plusieurs fois, les blocs fusionnent afin d'éviter les doublons. Une fois le terme ajouté, les spécialités correspondantes s'affichent dans un bloc dédié.

Les médicaments inclus dans la recherche sont *in fine* convertis en code UCD pour recherche.

La liste de l'ensemble des patients possédant un ou plusieurs documents contenant les termes recherchés sont retournés. L'affichage pourra se faire soit par prescription unique, soit par prescription agrégée sur l'ensemble du séjour du patient, par UF (unité fonctionnelle). Un exemple de résultats d'une recherche, avec affichage par prescription unique est présenté figure 7.

Requête		Statistiques		Map		Résultats		
Ouvrir les concepts		Mettre tous les patients dans le panier						
ipp	nom	prenom	Date naissance	S	CP	Actions	Panier	Doc
10026	GY_ZZ	Fabienne_	10/12/1926	F	-			Voir 121212 1212 23/09/2050 Prescription KM0007

Fabienne_ GY_ZZ , Née le 10/12/1926

23/09/2050 - Prescription
 IPP : 10026 , IEP : 121212
 UF : 1212-

Mots clés

Typ	Avis	Voie	Libellés	Début	Fin	Classe
ID	IV		BEVITINE 100 MG/2 ML, SOL INJ, AMP	23/09/2050 17:00:00	23/09/2050 17:20:00	THIAMINE (VIT B1)
ID	IV		SODIUM CHLORURE 0.9%, SOL PR PERF, POCHE 50 ML VIAFLO	23/09/2050 17:00:00	23/09/2050 17:20:00	ELECTROLYTES

Figure 7 : Affichage des résultats d'une recherche (ici, sur « Thiamine »). Notons qu'il s'agit ici d'une prescription anonymisée.

5 Discussion

L'intégration et l'exploitation des données de prescription restent complexes. Ces données sont présentes dans le système d'information de façon hétérogène. L'exploitation des données structurées suppose l'utilisation d'un référentiel à la fois stable et mis à jour très régulièrement. Le choix de la CIOsp, lié intrinsèquement au format de flux PN13, nous permet d'avoir un système continuellement à jour et en parfaite adéquation avec le contenu des documents intégrés dans Roogle.

Actuellement, ce référentiel est proposé par PHAST sous la forme de fichier texte. Dans notre expérimentation, ce fichier a été parsé pour être intégré dans Roogle. L'évolution vers une approche webservice, comme le développe actuellement thériaque pour sa base, permettrait une mise à jour de ce référentiel plus efficiente.

L'intégration de données de prescription au format PN13 a été aisée. Ce format n'étant pas facilement lisible par un humain, nous avons développé un parseur permettant de visualiser les prescriptions de façon similaire à la façon dont elles étaient représentées dans le DPI DxCare.

Nous prévoyons d'intégrer également les données d'administration des médicaments, toujours au format PN13. Dans ce cas, la visualisation de ces informations sera plus complexe

puisqu'elle nécessitera de reconstituer l'équivalent d'une pancarte dans l'interface de Roogle.

L'intégration de la base Thériaque nous permet d'envisager l'exploitation des relations contenues dans la base. Dans notre système, Thériaque n'est pas utilisé comme composant d'intégration des données sources (pour cela, nous utilisons la CIOsp), mais comme une base de connaissance, utilisée pour la construction des requêtes et l'annotation de résultats.

Thériaque met à présent à disposition sa base sous forme de webservice. Nous étudions actuellement cette voie comme modèle de couplage entre Roogle et Thériaque, dont l'intérêt serait d'avoir un entrepôt de données exploitable à l'aide de connaissance sur le médicament à jour.

6 Conclusion

L'intégration des flux liés aux médicaments dans l'entrepôt de données ouvre le champ des possibles quant à ses utilisations, et apporte un outil supplémentaire aux différents métiers, notamment aux professionnels de la recherche clinique pour les études de faisabilité ainsi que pour la constitution de cohorte.

De nombreuses améliorations restent cependant à apporter, notamment au niveau performance et interface utilisateur. Il est enfin possible d'imaginer d'autres utilisations encore plus évoluées des données de prescriptions : recherche des coordonnées de patients en cas de rappels de lots, autocomplétion des e-CRF via un plugin connecté à Roogle, etc. Une évaluation de l'outil est prévue.

Remerciements

Nous remercions

- PHAST, pour nous avoir transmis la dernière distribution à jour,
- La CNHIM, pour nous avoir permis d'accéder à l'ensemble de la base Thériaque

Références

- [1] F. F. Costa, « Big data in biomedicine », *Drug Discov. Today*.
- [2] S. N. Murphy, M. E. Mendis, D. A. Berkowitz, I. Kohane, et H. C. Chueh, « Integration of Clinical and Genetic Data in the i2b2 Architecture », *AMIA. Annu. Symp. Proc.*, vol. 2006, p. 1040, 2006.
- [3] M. Cuggia, N. Garcelon, B. Campillo-Gimenez, T. Bernicot, J.-F. Laurent, E. Garin, A. Happe, et R. Duvauferrier, « Roogle: an information retrieval engine for clinical data warehouse », *Stud Health Technol Inf.*, vol. 169, p. 584–588, 2011.
- [4] M.-C. Husson, « Thériaque® : base de données indépendante sur le médicament, outil de bon usage pour les professionnels de santé », *Ann. Pharm. Fr.*, vol. 66, n° 5-6, p. 268-277, nov. 2008.

Adresse de correspondance

Dalleau Kevin, Laboratoire Traitement du Signal et de l'Image (LTSI) INSERM : U1099 – Université de Rennes 1, 2 avenue du Professeur Léon Bernard. 35043 Rennes

kevin.dalleau@gmail.com

Utilisation des techniques de modélisation des connaissances pour la comparaison de recommandations de pratique clinique: application à la prise en charge de l'hypertension

Using knowledge modeling to compare clinical practice guidelines: a case study with the management of hypertension

Alexandre Galopin^{1,2,3,4}, Jacques Bouaud^{5,2,3,4},
Suzanne Pereira¹, Brigitte Sérroussi^{3,6,2,4}

¹ Vidal, Issy-les-Moulineaux, France

² INSERM, U1142, LIMICS, Paris, France

³ Sorbonne Universités, UPMC Université Paris 06, UMR_S 1142, LIMICS, Paris, France

⁴ Université Paris 13, Sorbonne Paris Cité, LIMICS, (UMR_S 1142), Bobigny, France

⁵ AP-HP, DRCD, Paris, France

⁶ AP-HP, Hôpital Tenon, DSP, Paris, France; APREC, Paris, France.

Résumé

Les guides de bonne pratique clinique (GBPC) sont des documents publiés pour orienter les pratiques des professionnels de santé dans un objectif d'amélioration de la qualité des soins. La profusion de GBPC sur un même thème et leurs disparités limitent leur effet sur les pratiques et posent le problème du choix des GBPC à suivre. Nous proposons d'utiliser la modélisation formelle pour comparer le contenu des GBPC publiées sur un même sujet. Dans une première étape, nous identifions et comparons les couvertures conceptuelles des GBPC via des techniques d'extraction terminologique. Dans une seconde étape, nous modélisons le contenu des GBPC sous la forme de règles décisionnelles que nous comparons par la mise en œuvre d'un raisonnement ontologique. L'application à trois GBPC contemporains sur la prise en charge de l'hypertension artérielle a montré des différences mais une cohérence globale des recommandations proposées. Les trois GBPC considérés sont ainsi complémentaires. Leur prise en compte simultanée pour élaborer la base de connaissances d'un système d'aide à la décision médicale permettrait d'améliorer la qualité des propositions du système.

Abstract

Clinical Practice Guidelines (CPGs) are textual documents published for healthcare professionals to improve the quality of care. The large number of CPGs for a given topic and their disparities reduce their impact on practices and raise the problem of the choice of the "right" CPGs. We propose to use formal modelling techniques to compare CPG contents. In a first step, we analyse the conceptual coverage of CPGs obtained through terminological extraction. In a second step, we model the content of CPGs as a set of decision rules that we compare with an ontological reasoning. The method has been applied to the comparison of three contemporary CPGs about the management of arterial hypertension. We observed some

differences but the overall consistency of the proposed recommendations was preserved. These three CPGs would be complementary and thus good candidates for building the knowledge base of a clinical decision support system.

Mots-clés : Guides de bonne pratique clinique comme sujet, Hypertension artérielle, Gestion des pharmacothérapies, Traitement du langage naturel, Extraction terminologique, Systèmes d'aide à la décision clinique

Keywords: *Practice guidelines as topic, Hypertension, Medication therapy management, Natural language processing, Term extraction, Clinical decision support*

1 Introduction

Les Guides de Bonne Pratique Clinique (GBPC) sont des documents qui, pour une pathologie donnée, fournissent une synthèse de l'état de l'art scientifique actualisé. L'objectif de ces publications est de fournir aux professionnels de santé des recommandations de prise en charge afin de réduire les pratiques inappropriées, et d'améliorer la qualité des soins. Ces documents textuels sont généralement rédigés à l'initiative d'agences nationales de santé telles la Haute Autorité de Santé (HAS) pour la France ou le National Institute for Health and Care Excellence (NICE) pour le Royaume-Uni. Des sociétés savantes comme le Programme d'Education Canadien sur l'Hypertension (PECH) au Canada, ou la Société Française d'Hypertension Artérielle (SFHTA) en France, éditent également leurs propres GBPC. Aussi, pour une même pathologie, il existe de nombreux GBPC, d'origines diverses et publiés à des dates différentes.

En dépit de leur large diffusion, il existe des obstacles à la mise en œuvre sur le terrain des GBPC [1]. Ainsi la seule diffusion sous forme de textes des GBPC a peu d'impact sur les pratiques alors que leur implémentation dans des systèmes informatisés d'aide à la décision permet d'améliorer leur suivi par les praticiens [2]. En effet, les systèmes d'aide à la décision médicale (SADM) s'appuient sur une modélisation des connaissances d'un GBPC permettant de produire, pour un patient donné caractérisé par un ensemble de critères, les recommandations personnalisées adaptées à ce patient.

Si la construction d'une base de connaissances pour un SADM à partir d'un GBPC est une étape non triviale, la question du choix du ou des GBPC à modéliser reste entière. Face à la profusion des GBPC sur un thème donné, quel GBPC choisir ? Établis en théorie sur une même base scientifique, celle de la littérature disponible à une date donnée, des GBPC contemporains devraient délivrer les mêmes recommandations par delà des variations éditoriales et de format. Pourtant, une étude de 2008 comparant plusieurs GBPC sur l'hypertension artérielle mettait en évidence des différences sur le contenu des documents [3].

Dans l'objectif de développer un SADM pour la prise en charge de l'hypertension artérielle, le but de cette étude est d'évaluer dans quelle mesure des GBPC contemporains sont équivalents en terme de contenu, c'est-à-dire de mesurer si les connaissances proposées au niveau des recommandations de pratique clinique sont comparables. Le travail a été réalisé en utilisant les méthodes de l'ingénierie des connaissances afin de modéliser le contenu des textes des GBPC. Dans un premier temps, nous nous sommes intéressés à la couverture conceptuelle des GBPC en utilisant des techniques d'extraction terminologique [4]. Puis, dans un second temps, nous avons formalisé les GBPC sous forme de règles décisionnelles correspondant aux recommandations proposées. La comparaison des GBPC a ainsi porté sur les concepts mobilisés et les règles décisionnelles formalisant les profils cliniques et les recommandations de prise en charge associées.

L'article est organisé de la manière suivante. La section suivante relate les travaux existant sur la comparaison de GBPC, l'extraction de concepts à partir de textes et la modélisation des GBPC. La section 3 présente les ressources utilisées dans cette étude, les 3 GBPC publiés en français en 2013 que nous avons choisis, et l'ontologie médicale que nous avons utilisée, puis les méthodes employées pour analyser les textes des GBPC. Dans la section 4, nous présentons les résultats de la comparaison des 3 GBPC. La dernière section discute ces résultats et envisage les perspectives de ce travail.

2 État de l'art

2.1 Comparaison de GBPC

Chaque année, plusieurs dizaines de GBPC sont publiés sur un même sujet par des agences nationales ou des sociétés savantes. Si l'évolution continue des connaissances médicales [5] justifie l'actualisation des GBPC, cette actualisation n'est pas réalisée au même moment par les différents producteurs de GBPC. En faisant abstraction des différences liées au pays de production, ou à la date de publication, il existe des différences de contenu en termes de recommandations qui constituent des obstacles à la mise en œuvre des GBPC. Ainsi, Flottorp et al. [1] citent parmi les facteurs s'opposant au suivi des GBPC l'existence de conflits entre recommandations. En pratique, en dehors d'un cadre réglementaire, la question du choix des GBPC est posée [6]. Ainsi, une étude prospective récente comparant 3 SADM s'appuyant sur 3 GBPC pour l'indication d'un scanner cérébral montre un accord inter-GBPC allant de 60 % à 80 %, illustrant par là même l'existence de 20 % à 40 % de divergences [7]. Georg et al. [3] ont réalisé en 2008 une étude comparative de quatre GBPC sur la prise en charge de l'hypertension artérielle. Alors que par principe, les GBPC s'appuient sur l'« état de l'art », les auteurs ont établi que le nombre de références communes était lié aux différences entre les recommandations. Chaque GBPC s'appuie en effet sur une base bibliographique non exhaustive constituée d'articles scientifiques issus de la recherche documentaire conduite par les auteurs, qui lui sont donc propres. Les contenus des GBPC, construits à partir des références sélectionnées, peuvent ainsi présenter des différences au niveau des recommandations. De plus, en l'absence de preuve (*evidence*), l'élaboration d'une recommandation passe souvent par une interprétation des sources par un groupe de travail. Ces interprétations s'écartent du principe de l'*evidence-based medicine* et représentent un second biais pour la construction de recommandations, ce qui est également critiqué par les professionnels de santé [1]. Actuellement, il n'existe pas de méthode permettant de comparer des GBPC portant sur un même sujet. La comparaison des textes d'origine n'est pas une solution du fait de la faiblesse des textes, qui sont ambigus, abstraits, incomplets, et nécessitent d'être interprétés. En revanche, des représentations formalisées de leur contenu, comme on le réalise pour la construction de SADM, devraient plus facilement permettre de comparer plusieurs GBPC. C'est cette approche qui est testée dans ce travail.

2.2 Analyse du texte des RPC

Les techniques de traitement automatique du langage naturel ne permettent pas de « comprendre » les textes des GBPC de façon satisfaisante. De manière moins ambitieuse, mais plus opérationnelle, l'extraction terminologique s'attache à la récupération d'information pour un domaine particulier. Les approches permettant une extraction terminologique automatique utilisent généralement des opérations linguistiques (*part-of-speech tagging*, *phrase chunking*) pour extraire les candidats-termes, c.-à-d. les groupes de mots potentiellement pertinents. Les

entrées terminologiques sont ensuite filtrées au sein de la liste des candidats en utilisant des méthodes statistiques. Après ce filtre, il ne reste plus que les termes avec une faible ambiguïté et une forte spécificité qui seront utiles pour la conceptualisation d'un domaine de connaissance ou pour la création d'une ontologie (voir [8,9]).

De nombreux logiciels dédiés à l'extraction de termes ont été développés. Parmi ces logiciels, on retrouve TreeTagger [10], un *part-of-speech tagger* performant qui se charge d'annoter un texte en identifiant les fonctions de chaque mot d'une phrase. S'il s'avère efficace pour des textes narratifs bien construits, les résultats sont moins convaincants avec des textes composés de nombreuses listes et tableaux comme c'est le cas des GBPC.

L'extraction des termes des GBPC nécessite ainsi des outils plus performants. Développé par l'équipe du professeur Drouin, TermoStat [11] est un extracteur terminologique disponible en ligne capable d'analyser des termes simples et complexes à partir de phrases construites ou de mots isolés. Comparé à d'autres outils, TermoStat ajoute une autre étape de comparaison statistique avec un corpus de référence composé d'une grande collection hétérogène de documents textuels. Un score est alors donné à chaque candidat-terme, calculé à partir de sa fréquence relative. TermoStat prend en charge plusieurs langues dont le français. Dans le contexte d'identification des concepts référencés dans un texte, la sélection des candidats-termes pertinents doit être réalisée manuellement.

2.3 Modélisation des connaissances

Si il existe de nombreux formalismes de représentation des connaissances spécifiques des GBPC [12], le passage du texte non structuré comme c'est le cas des GBPC à une base de connaissances formalisée est un problème difficile. Il n'existe en effet pas de méthode systématique et de nombreux travaux proposent des méthodologies, dont la mise en œuvre est facilitée par des outils. Ainsi, avec l'outil DeGeL [13], les GBPC sont structurés manuellement, d'abord convertis dans un texte sémantiquement semi-structuré et dans un langage semi-formel par l'expert médical. D'autres approches impliquent l'utilisation d'information heuristique [14].

GEM est un modèle XML permettant de repérer et d'extraire les structures décisionnelles canoniques dans un texte de GBPC [15]. Le traitement par GEM-Cutter aboutit à un fichier XML proposant une première structuration des GBPC. Dans ce contexte, Shiffman et al. [15] ont développé une méthode permettant de traduire un GBPC narratif en un ensemble de recommandations formalisées. Cette méthode repose essentiellement sur quatre étapes. La première étape d'atomisation est l'identification des notions pertinentes qui seront les critères décisionnels et les actions recommandées. La suivante est l'étape de déabstraction qui permet de reformuler les notions identifiées dans une interprétation opérationnalisable. L'étape de désambiguïsation vise à éliminer toute ambiguïté des recommandations textuelles pour faciliter leur interprétation et leur modélisation formelle. Enfin, une dernière étape de vérification de la complétude permet d'envisager des situations cliniques théoriquement valides mais non explicitement mentionnées dans les GBPC et qui correspondent aux « trous de connaissances » des GBPC. Cette dernière étape a pour objectif de proposer une modélisation complète du domaine couvert, mais elle possède l'inconvénient de s'éloigner du contenu strict des GBPC. Toutefois, quel que soit le formalisme de représentation visé, et que l'on décide ou non de réaliser la dernière étape, les étapes précédentes sont des préalables à la formalisation.

3 Matériel et méthodes

L'objectif de ce travail est de comparer le contenu des GBPC sur la base de leur formalisation en utilisant des techniques de l'ingénierie des connaissances. Nous décrivons ci-dessous les GBPC qui ont été choisis pour l'étude ainsi que l'ontologie qui servira de point de départ pour la représentation des GBPC sous forme de règles de production. Puis, nous décrivons les méthodes employées pour d'abord caractériser la couverture conceptuelle des textes de chaque GBPC et pour ensuite construire et comparer les bases de connaissances correspondant aux GBPC.

3.1 Les Guides de Bonne Pratique

Lors de cette étude, en l'absence de recommandations nationales officielles récentes sur la prise en charge de l'hypertension artérielle (les dernières recommandations de l'HAS datent de 2005), nous avons choisi de considérer trois recommandations contemporaines, datant de 2013, écrites en français et d'origines diverses. « VR-HTA » (Vidal-Recos HTA) [16] est un GBPC réalisé par Vidal, une société française spécialisée dans la mise à disposition d'informations sur les médicaments, récemment engagée dans la publication sous une forme synthétique de recommandations établies à partir d'autres GBPC déjà publiés. VR-HTA compte 19 pages et inclut 5 références bibliographiques. « SFHTA » est un GBPC rédigé par la Société Française d'HyperTension Artérielle et se présente sous la forme d'un mémento synthétique [17]. Il s'agit d'un guide relativement court (4 pages, 40 références bibliographiques) qui explore uniquement les cas les plus généraux de l'hypertension. Enfin, le GBPC « PECH » [18] est publié par le Programme d'Education Canadien sur l'Hypertension et décrit tous les aspects de la prise en charge de l'hypertension. Il est mis à jour chaque année et compte 35 pages.

3.2 OntoUrgences

Afin de permettre la modélisation des GBPC sous forme de règles et de comparer ces règles, nous avons cherché à réutiliser une ontologie médicale existante comprenant les principaux concepts liés à l'HTA. Nous avons utilisé OntoUrgences qui est une ontologie développée par notre laboratoire dans le cadre du projet LERUDI [19]. Elle organise près de 10,000 concepts relatifs aux signes, symptômes et pathologies des patients admis aux urgences. Ce projet visait à donner une visualisation rapide d'un dossier médical et d'orienter le patient vers le service approprié. Il est possible d'enrichir cette ontologie avec de nouveaux concepts adaptés à nos besoins. Enfin, sa représentation en OWL permet d'utiliser un raisonneur pour classifier l'ensemble des nouveaux concepts et comparer les règles exprimées à différents niveaux d'abstraction.

3.3 Comparaison des couvertures conceptuelles des GBPC

L'analyse que nous avons développée dans le but d'identifier les concepts médicaux pertinents et produire les couvertures conceptuelles de chaque GBPC s'appuie sur l'extraction automatique de candidats termes à partir des textes des GBPC. L'hypothèse que sous-tend cette analyse est que si les GBPC ne possèdent pas la même couverture conceptuelle, leur contenu sera différent. En revanche, même s'ils font référence aux mêmes concepts, cela ne garantit pas que leur contenu sera semblable. L'analyse conceptuelle réalisée repose sur les 5 étapes suivantes qui nécessitent des traitements « manuels » :

- a) *Prétraitement des GBPC.* Les GBPC textuels ont besoin d'être prétraités afin de permettre une analyse informatique. Chaque élément structuré (tableaux, images, graphiques) a été

converti en texte brut afin d'éviter les mises-à-plats incorrectes. De plus, le travail étant centré sur la partie thérapeutique de la prise en charge, nous avons éliminé les autres sections, c.-à-d. la liste des auteurs, l'évaluation, le diagnostic, et les références. Les trois GBPC ont été concaténés sous la forme d'un seul document pour les étapes suivantes de l'analyse.

- b) *Extraction terminologique automatique.* Le document complet a été soumis à TermoStat qui a retourné une liste de candidats-termes, ordonnés selon leur score de spécificité et associés à leur contexte syntaxique, c.-à-d. les expressions constituant le voisinage de chacun des candidats termes.
- c) *Suppression manuelle des termes non pertinents.* Les termes vides de sens tels que les prépositions, p. ex. 'de', ou les unités, p. ex. 'mmHg', ont été supprimés. Les candidats-termes restants ont ensuite été considérés selon leur contexte syntaxique. Lorsqu'un candidat terme est jugé hors-sujet (p. ex. 'pluridisciplinaire', 'épidémiologie',...), il est aussi supprimé. Ce traitement conduit à un sous-ensemble de termes liés à la prise en charge thérapeutique de l'hypertension.
- d) *Regroupement des candidats-termes et identification des concepts.* Cette étape consiste à rassembler les candidats-termes qui sont liés à un même concept. Nous avons ici aussi utilisé les contextes syntaxiques pour guider les regroupements. Par exemple, le terme 'occlusion' est utilisé dans le seul contexte de 'AVC ischémique' et a donc été groupé avec ce dernier. Trois personnes (AG, JB et BS), avec des profils différents, ont réalisé cette étape de manière indépendante. Les regroupements correspondants ont été comparés et les conflits ont été résolus de façon consensuelle. Nous avons finalement obtenu des regroupements de termes liés à un même concept. Par exemple, le regroupement R7 associé au concept d'AVC ischémique est constitué des termes « AVC ischémique, occlusion de l'artère cérébrale, artère carotide, carotide extracrânienne, ischémique, occlusion ».
- e) *Construction de la couverture conceptuelle.* Pour chaque GBPC, les termes regroupés ont été recherchés dans les textes et mis en relation avec des concepts médicaux. L'ensemble des concepts ainsi identifiés au sein d'un GBPC constitue sa couverture conceptuelle.

Les couvertures conceptuelles des trois GBPC ont ensuite été comparées 2 à 2, et toutes les 3 au moyen des opérations ensemblistes classiques (intersection). Un sous-objectif était caractériser les concepts communs aux 3 GBPC et ceux qui étaient spécifiques d'un GBPC.

3.4 Comparaison des modèles décisionnels des GBPC

La modélisation des connaissances décisionnelles sous forme de règles dans la perspective de la réalisation d'un SADM permet de considérer le contenu des GBPC uniquement sous l'angle des connaissances opérationnalisables. L'hypothèse sous-jacente à la comparaison des modélisations est que si des textes différents conduisent à des modélisations identiques alors leur contenu est identique. Si les modélisations diffèrent, une analyse plus fine doit être réalisée car ces modélisations peuvent se compléter, être plus précises, ou encore être en conflit en proposant des actions contradictoires pour une situation identique.

L'approche adoptée ici repose sur une modélisation sous la forme de règles de production puis sur une comparaison des parties conditions de ces règles utilisant la classification ontologique. Elle a été réalisée en suivant les 4 étapes suivantes :

- a) *Extraction des règles.* Une règle est composée d'une partie condition et d'une partie action. Les trois GBPC ont été manuellement traités par trois personnes (AG, JB, BS) qui ont travaillé de manière indépendante à l'identification des règles décisionnelles relatives à la prise en charge thérapeutique de l'HTA. Pour chaque GBPC, les trois bases de règles ont été construites et consolidées de manière consensuelle sous la forme d'une seule base de règles.

Les critères des parties condition ont été normalisés afin d'éviter les variations syntaxiques en se référant à OntoUrgences. Lorsqu'un critère d'une règle a une correspondance conceptuelle dans OntoUrgences, nous utilisons le label du concept. La partie condition est ainsi une conjonction de concepts qui décrit un profil patient. La partie action, qui comporte les prises en charge associées, a été standardisée mais non formalisée. Il est à noter qu'un profil patient peut être utilisé dans les recommandations de plusieurs GBPC.

- b) *Enrichissement ontologique.* Comme certains critères utilisés dans la définition des profils patients n'avaient pas de correspondance dans OntoUrgences, ces concepts ont été ajoutés et classés au sein de l'ontologie en prenant en compte les liens hiérarchiques existants. Par exemple, le concept HTA n'avait pas de sous-classe, nous avons donc créé le concept de HTA Systolique et l'avons classé en tant que sous-classe de HTA. Pour réaliser ces opérations, nous avons utilisé le logiciel Protégé.
- c) *Classification des profils.* Les profils patients étant une conjonction de concepts, ils peuvent implicitement être classés selon des liens de subsumption. Par exemple, un profil P_1 : "C_a", subsume P_2 : "C_a et C_b" et P_1 subsume P_3 : "C_c" si "C_a" subsume "C_c" dans l'ontologie. Cette propriété implique que toutes les actions s'appliquant à un profil donné s'appliquent nécessairement à tous les profils qu'il subsume. Tous les profils extraits des GBPC sont donc définis dans l'ontologie et classifiés grâce à un raisonneur OWL produisant, après élimination des concepts qui ne sont pas des profils, un graphe de subsumption de profils patients.
- d) *Evaluation des recommandations.* Toutes les actions associées à un profil patient donné et aux profils le subsumant sont ensuite comparées. Pour la comparaison, nous avons identifié trois situations :
 - a. *Pas d'incohérence:* toutes les actions sont identiques ou équivalentes (par exemple, "changer le traitement" et "modifier le traitement").
 - b. *Incohérence potentielle:* les actions peuvent conduire à un conflit décisionnel dépendant de la nature et de l'origine de l'incohérence :
 - i. Complétion: une action est complétée par une nouvelle action. Par exemple, une situation plus spécifique nécessite une action complémentaire (p. ex. "ajouter de l'aspirine au traitement courant").
 - ii. Restriction: Une action est restreinte par une nouvelle action. La restriction peut s'expliquer par une précision du profil patient. Par exemple, prescrire des IEC alors que la recommandation précédente conseillait des IEC, des ARA2 ou des bêtabloquants.
 - c. *Contradiction:* des actions sont contradictoires (p. ex. "Prescrire des Bêtabloquants" et "Ne pas prescrire de Bêtabloquants"). Ces conflits peuvent provenir de recommandations issues de différents GBPC ou au sein d'un même GBPC dans le cadre de profils spécifiques.

4 Résultats

4.1 Comparaison des couvertures conceptuelles des GBPC

La compilation des trois GBPC est un document contenant 11,313 mots. TermoStat a identifié 723 candidats termes parmi lesquels 246 ont été considérés comme non pertinents et supprimés. Les 477 termes restants ont été manuellement groupés en 88 clusters, chaque cluster étant lié à un seul concept médical. Concernant les couvertures conceptuelles, VR-HTA comprend 81

concepts, SFHTA 53 concepts et PECH 73 concepts. La Figure 1 illustre quantitativement la couverture conceptuelle de chaque GBPC. Elle montre aussi les concepts spécifiques à chaque GBPC.

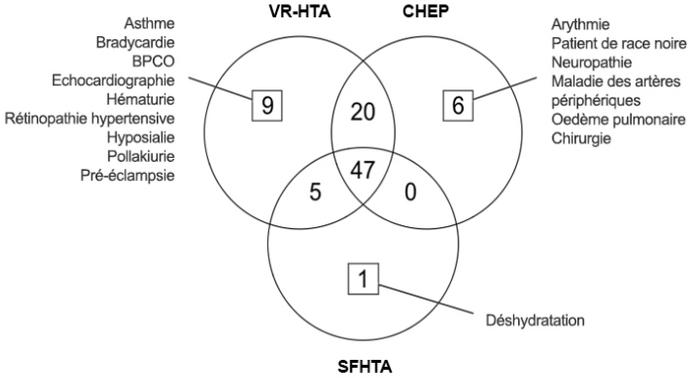


Figure 1 : Couverture conceptuelle des 3 GBPC

Quarante-sept concepts (53% du total) sont communs aux 3 GBPC. Ils représentent les concepts centraux de la prise en charge de l'hypertension artérielle. Seize concepts (18% du total) apparaissent uniquement dans une seule RPC: 9 dans VR-HTA (56%), 1 dans SFHTA (6%), et 6 dans PECH (38%). L'intersection de VR-HTA et CHEP contient 67 concepts et 98% des concepts de SFHTA sont aussi présents dans l'union de VR-HTA et CHEP.

4.2 Comparaison des règles décisionnelles

La modélisation des trois GBPC a conduit à l'identification de 205 règles décisionnelles parmi lesquelles 165 traitaient de l'aspect thérapeutique (28 pour SFHTA, 62 pour VR-HTA, 115 pour PECH). Le tableau 1 donne des exemples des règles extraites. Parmi ces 165 règles, nous avons identifié 27 doublons (même partie condition) conduisant ainsi à 138 profils patients différents.

Lors de la comparaison des profils patients, seulement 2 profils sont communs aux trois GBPC (P1 : "HTA" et P11 : "HTA + Objectif Tensionnel non Atteint + Bithérapie courante"), un profil est commun à VR-HTA et PECH, un est commun à VR-HTA et SFHTA et 21, 35 et 78 profils sont spécifiques à SFHTA, VR-HTA et PECH, respectivement.

GBPC	Id	Profils Patients	Actions
SFHTA	R1	P1:HTA	IEC, ARA2, ICA, Diurétiques Thiazidiques, Betabloquants recommandés, IEC +ARA2 contre-indiqué ...
SFHTA	R11	P11:HTA \wedge ObjectifNonAtteint \wedge BithérapieCourante	Modifier dosage ou modifier bithérapie ou prescrire trithérapie
VR-HTA	R49	P11:HTA \wedge ObjectifNonAtteint \wedge BithérapieCourante	Modifier dosage ou prescrire trithérapie

VR-HTA	R78	P78:HTA \wedge Bradycardie	Betabloquants contre-indiqués
CHEP	R113	P11:HTA \wedge ObjectifNonAtteint \wedge BithérapieCourante	Prescrire trithérapie
CHEP	R162	P162:HTA \wedge InsuffisanceCardiaque \wedge ObjectifNonAtteint	Prescrire IEC+ARA2

Tableau 1 : Exemples de règles décisionnelles extraites des GBPC

Au total, 121 concepts ont été utilisés pour représenter l'ensemble des profils. Seulement 16 d'entre eux (13%) ont une correspondance dans OntoUrgences (e.g. HTA, Bradycardie...). Les 105 concepts restants ont été ajoutés manuellement (e.g. Objectif non atteint, Traitement courant...).

Les 138 profils ont ensuite été classés par un raisonneur OWL conduisant à un graphe de subsomption. La Figure 1 montre un extrait de ce graphe avec P1 : "HTA" qui subsume tous les autres profils. Nous avons observé 36 profils isolés (seulement liés à P1) qui correspondent à des conditions particulières (e.g. "HTA et Bradycardie"). Nous avons aussi observé 10 groupes de profils liés par des liens de subsomption. (e.g. HTA et Age>50 subsume HTA et Age>80).

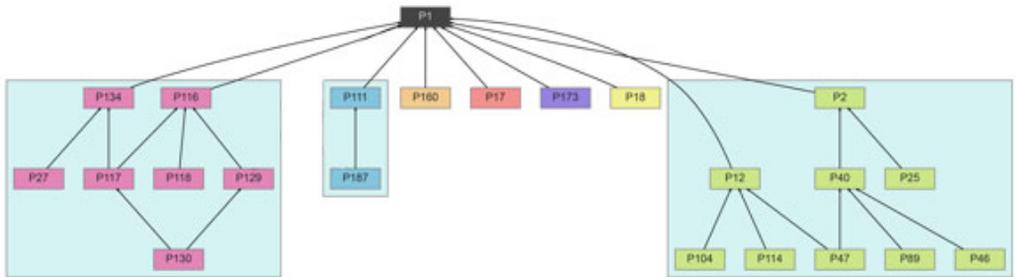


Figure 2 : Graphe de subsomption des profils patients décrits dans les GBPC

Lorsque les liens de subsomption sont pris en compte, nous trouvons que chaque profil est subsumé par au moins une règle issue de chaque GBPC. Ceci est dû au fait que le profil racine P1 est utilisé par les trois GBPC. Toutefois, si l'on ne considère pas P1, trop général, 10 profils sont couverts par les 3 GBPC, 44 par 2 GBPC et 83 par seulement 1 GBPC.

Concernant la cohérence des conclusions associées à un profil patient et aux profils le subsumant, nous avons observé la distribution suivante: 38 cas de non incohérence, 95 cas d'incohérence potentielle (15 cas d'extension et 80 cas de restriction) et 6 cas de contradiction (par exemple, P162 : "HTA + Insuffisance Cardiaque + Objectif non atteint" recommande une association IEC+ARA2 qui est contre-indiqué par P1: "HTA", un profil plus général).

5 Discussion

Les méthodes que nous avons détaillées permettent de comparer le contenu de GBPC selon deux niveaux de formalisation. Ces méthodes ont été appliquées à trois GBPC écrits en français traitant de la prise en charge de l'hypertension artérielle. Il est ainsi possible de quantifier les différences potentielles qu'il peut exister entre les GBPC et évaluer la pertinence d'une base de connaissance hybride constituée à partir de règles décisionnelles extraites des trois GBPC.

La première analyse, conceptuelle, a mis en évidence des disparités entre les différents GBPC. Ces différences se retrouvent tant sur la forme que sur le contenu. Le style d'écriture, le volume, le niveau de précision des recommandations sont autant de choix éditoriaux qui influent de manière certaine sur les couvertures conceptuelles des GBPC. Ainsi, l'ensemble des concepts présents dans SFHTA qui est un GBPC synthétique et général est presque totalement inclus dans l'ensemble des concepts de VR-HTA et PECH. Seulement la moitié des concepts est effectivement partagée par les trois GBPC et près de 20% des concepts ne se retrouvent que dans un seul GBPC. On remarque que les concepts spécifiques à VR-HTA sont généralement liés aux contre-indications médicamenteuses tandis que pour PECH, ce sont essentiellement des cas particuliers additionnels. Les résultats donnent des informations intéressantes quant aux thèmes abordés mais en aucun cas si les concepts considérés sont impliqués dans des règles décisionnelles de prise en charge. L'utilisation de TermoStat pour l'extraction terminologique apporte aussi un biais. Il s'agit d'un procédé automatique

Lors de la seconde analyse qui s'attachait à comparer les règles décisionnelles, nous avons véritablement comparé les recommandations proposées par les GBPC. Cette partie de l'étude s'intéresse directement aux difficultés que peuvent avoir les médecins lors de la recherche d'une recommandation adaptée à un cas patient. L'obtention des résultats a mis en évidence le faible nombre de profils partagés par les trois GBPC (seulement 2). Il se trouve que les profils traités par les GBPC sont à des niveaux de précision différents. C'est ainsi que des recommandations qui semblaient contradictoires, concernaient en pratique des profils patients différents. Les conflits décisionnels sont essentiellement dûs au gain de précision du profil. Cependant, il est à noter qu'aucune contradiction critique n'a été relevée pour un profil patient considéré au même niveau d'abstraction. La mise en commun des bases de règles ne soulèverait a priori aucune difficulté décisionnelle et permettrait de réduire les silences causés par un profil non traité par la base de connaissance.

Quelle soit la méthode considérée, une des limites réside dans la subjectivité introduite par l'intervention des étapes manuelles. En effet, que ce soit pour le regroupement des termes ou encore l'identification des règles décisionnelles, il n'existe aucun outil efficace capable de réaliser automatiquement ces actions. Afin de limiter le biais, ces tâches ont été réalisées par 3 des auteurs (AG, JB, BS) de manière indépendante. Les résultats obtenus ont ensuite été comparés, discutés et fusionnés.

Comme toute ontologie, OntolUrgences a été développée dans un but précis: proposer un aperçu rapide d'un dossier patient pour accélérer l'orientation du patient vers un service médical. Ainsi, OntolUrgences contient essentiellement des concepts et relations relatifs aux pathologies, signes et symptômes ainsi que de nombreux concepts utiles à la caractérisation du patient. On constate par ailleurs des manques en termes de traitement et de logique de prise en charge.

Cette étude a permis de proposer des méthodes de comparaison de GBPC (réutilisables et applicables à d'autres thèmes) mais aussi d'obtenir une information quantitative et qualitative sur les différences qui peuvent exister entre trois GBPC sur la prise en charge de l'hypertension de l'adulte. Afin de poursuivre sur la voie du développement d'un système d'aide à la décision qui saura s'adapter aux attentes des cliniciens et fournir efficacement des recommandations, une consolidation de la base de connaissance s'avère indispensable. Cette consolidation devra prendre en compte des aspects d'évolutivité et de performances en réutilisant des outils existants et en s'appuyant sur les standards du domaine.

Références

- [1] Flottorp S, Oxman A, Krause J, Musila N, Wensing M, Godycki-Cwirko M, et al., A checklist for identifying determinants of practice: A systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice, *Implement Sci* 8/1 (2013), 35.
- [2] Jaspers MW, Smeulers M, Vermeulen H, Peute LW. Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings. *J Am Med Inform Assoc.* 2011;18(3):327–34.
- [3] Georg G, Colombet I, Durieux P, Ménard J, Meneton P. A comparative analysis of four CPGs for hypertension management. *J Hum Hypertens.* 2008 Dec;22(12):829-37.
- [4] Galopin A, Bouaud J, Pereira S, Séroussi B, Comparison of clinical practice guidelines from a knowledge modeling perspective: a case study with the management of hypertension. EFMI Special Topic Conference (STC), (2014), Budapest, submitted.
- [5] Wyatt J. Use and sources of medical knowledge. *Lancet* 1991; 338: 1368-1373
- [6] Ferket BS, Colkesen EB, Visser JJ, Spronk S, Kraaijenhagen RA, Steyerberg EW, et al., Systematic review of guidelines on cardiovascular risk assessment: Which recommendations should clinicians follow for a cardiovascular health check?, *Arch Intern Med*, 170/1 (2010), 27-40.
- [7] Korley FK, Morton MJ, Hill PM, Mundangeppufu T, Zhou T, Mohareb AM, Rothman RE. Agreement between routine emergency department care and clinical decision support recommended care in patients evaluated for mild traumatic brain injury, *Acad Emerg Med.* 2013 May;20(5):463-9.
- [8] Liu K, Hogan WR, Crowley RS. Natural Language Processing methods and systems for biomedical ontology learning. *J Biomed Inform.* 2011 Feb;44(1):163-79.
- [9] Dramé K, Diallo G, Delva F, Dartigues JF, Mouillet E, Salamon R, Mouglin F. Reuse of termino-ontological resources and text corpora for building a multilingual domain ontology: An application to Alzheimer's disease. *J Biomed Inform.* 2013 Dec 29. pii: S1532-0464(13)00204-9.
- [10] Schmid H. Improvements in Part-of-Speech Tagging with an Application to German. *Text, Speech and Language Technology Volume 11, 1999*, pp 13-25.
- [11] Drouin P. Term extraction using non-technical corpora as a point of leverage. *Terminology.* 2003;9(1),99-115.
- [12] Peleg, M., Tu, S., Bury, J., Ciccarese, P., Jones, N. and Miksch, S., Comparing computer-interpretable guideline models. A case-study approach. *J Am Med Inform Assoc.* v10. 52-68.
- [13] Hatsek A, Young O, Shalom E, Shahar Y. DeGeL: a clinical-guidelines library and automated guideline-support tools. *Stud Health Technol Inform.* 2008;139:203-12.
- [14] Kaiser K, Akkaya C, Miksch S. How can information extraction ease formalizing treatment processes in clinical practice guidelines? A method and its evaluation. *Artif Intell Med.* 2007 Feb;39(2):151-63. Epub 2006 Sep 8.
- [15] 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) :418–426.
- [16] Vidal. Vidal Recos HTA. 2013. [Commercial content available at

http://www.vidal.fr/recommandatio-ns/1640/hta_hypertension_arterielle/. Last accessed January 6, 2014].

- [17] SFHTA. Recommandation sur « La prise en charge de l'hypertension artérielle de l'adulte». 2013. [Available at <http://www.sfhta.eu/wp-content/uploads/2012/12/Recommandation-SFHTA-2013-Prise-en-charge-HTA-de-l'Adulte.pdf>. Last accessed January 6, 2014].
- [18] Hypertension Canada. Recommandations 2013 du Programme Éducatif Canadien sur l'Hypertension. 2013. [Available at http://www.hypertension.ca/images/CHEP_2013/2013_CHEPRecsFullVersion_FR_HCP1000.pdf. Last accessed January 6, 2014].
- [19] Charlet J, Declerck G, Dhombres F, Miroux P, Gayet P, Vandebussche PY, Building a medical ontology to support information retrieval: terminological and metamodelization issues. Proc. of the 10th International Conference on Terminology and Artificial Intelligence (2013), Villetaneuse, France.

Adresse de correspondance

Pour toute demande d'informations, merci de contacter Alexandre Galopin à l'adresse: alexandre.galopin@vidal.fr

Enrichissement sémantique associé à la détection de la négation et des antécédents familiaux dans un entrepôt de données hospitalier

Semantic enrichment based on detection of negation and history family in a French hospital data warehouse

Nicolas Garcelon^{1,2}, Rémi Salomon³, Anita Burgun²

¹*Institut Imagine, Paris Descartes Université Paris Descartes-Sorbonne Paris Cité, Paris, France*

²*INSERM, Centre de Recherche des Cordeliers, UMR 1138 Equipe 22, Paris*

³*Service de Néphrologie Pédiatrique, Hôpital Necker-Enfants Malades, Assistance Publique - Hôpitaux de Paris (AP-HP), Université Paris Descartes, Sorbonne Paris Cité -*

Résumé

(Background) Nous avons développé dans des travaux précédents un entrepôt de données biomédicales permettant de fouiller les comptes rendus par un moteur de recherche plein texte. Mais l'absence de prise en compte de la négation et des antécédents familiaux engendraient du bruit. (Methods) Nous avons développé un algorithme permettant de détecter en français la notion de négation dans les comptes rendus médicaux, ainsi que le contexte d'antécédents familiaux. L'enrichissement sémantique tient compte de ces nouvelles informations sémantiques. Nous intégrons dans le schéma de l'entrepôt de données ainsi que dans l'algorithme de fouille de données ces nouvelles informations afin d'améliorer les performances du moteur de recherche. (Results) L'intégration de la détection de la négation et des antécédents familiaux améliore nettement la précision du moteur de recherche malgré une légère diminution du rappel. (Conclusion) Cet enrichissement syntaxique et sémantique a été intégré dans l'entrepôt de données avec succès.

Abstract

(Background) We developed in previous work a biomedical data warehouse for mining clinical reports by full text search engine. But the lack of consideration of negation and family history generate noise. (Methods) We developed an algorithm to detect the French notion of negation in medical reports and the context of family history. The semantic enrichment reflects this new semantic information. We integrate in the data warehouse schema and in data mining algorithm this new information to improve the performance of the search engine. (Results) the integration of the detection of negation and family history significantly improves the precision of the search engine despite a slight decrease in recall. (Conclusion) The syntactic and semantic enrichment has been successfully integrated into the data warehouse.

Mots-clés : Entrepôt de données ; fouille de données ; Traitement automatique du langage

Keywords: Data warehouse; data mining; Natural Language Processing

1 Introduction

Lors de travaux précédents, nous avons développé un entrepôt de données essentiellement orienté sur des comptes rendus plein texte [1]. Il s'agissait de proposer un outil simple permettant de rapidement fouiller les comptes rendus médicaux des patients. Dans l'onglet résultat du moteur de recherche, l'utilisateur peut facilement vérifier si le patient est vrai positif par l'affichage de la section du document contenant les termes recherchés. Nous avons ensuite développé un module d'enrichissement sémantique [2] utilisant le Metathesaurus de l'UMLS [3]. L'utilisateur choisit s'il veut utiliser l'enrichissement sémantique en fonction de la sensibilité et spécificité souhaitée. Le moteur de recherche exploite les relations de synonymie et de subsumption du Metathesaurus pour retrouver le plus grand nombre de patients dont les documents contiennent les termes recherchés, ce qui a pour effet d'augmenter la sensibilité aux dépens d'une moins bonne spécificité. Le bruit obtenu est lié non seulement aux phrases négatives (« absence de », « pas de » etc.) mais aussi au contexte d'antécédents familiaux dans lequel l'expression recherchée est située (« la sœur du patient est atteinte de »). Ce constat réalisé nous proposons une méthode pour détecter le contexte d'antécédents familiaux, ainsi que les sous phrases négatives afin de les exclure de la recherche plein texte. Ces méthodes sont particulièrement nécessaires pour traiter les comptes rendus produits dans les services prenant en charge des patients atteints de maladies génétiques héréditaires, comme l'Hôpital Necker à Paris et l'Institut Imagine (www.institutimagine.org/). Il s'agit aussi de compléter l'enrichissement sémantique par l'ajout de ce niveau de certitude et du contexte dans lequel le concept est retrouvé.

2 État de l'art

La plupart des travaux de Traitement Automatique du Langage concernent la langue anglaise. En 2001, Chapman et al. développent NegEx [4]. Il s'agit d'un algorithme qui se veut simple et ouvert. L'auteur associe une extraction sémantique à partir des concepts UMLS et une détection par expression régulière d'une forme négative avant ou après le concept UMLS extrait. Si la phrase ne contient pas de concept UMLS, la phrase n'est pas traitée. Goryachev et al. [5] ont évalué en 2006 4 méthodes de détection de la négation. Il a notamment comparé la méthode NegEx et la méthode NegExpand développée par Aronow et al. [6]. Tandis que NegEx utilise l'intégralité de la phrase, NegExpand se base sur l'analyse de syntagmes nominaux en découpant le texte suivant certains critères prédéfinis (ponctuation, « and », « or »). La détection de la négation est réalisée indépendamment de l'extraction de concepts UMLS. Les 2 autres méthodes décrites dans cet article se basent sur un système d'apprentissage automatique (Weka) sur un corpus de plus de 1700 résumés de sortie pour lesquels les 8 000 termes extraits ont été manuellement annotés par un expert comme un fait, une négation ou une éventualité. Ces 2 méthodes divergent ensuite par l'algorithme de classification : la classification bayésienne naïve et le Vector Space Model. Il conclut que la méthode NegEx reste la plus efficace. L. Deléger & Grouin ont proposé d'adapter NegEx à la langue française [7]. L'adaptation française de NegEx donne un rappel et une précision corrects. Toutefois, l'auteur fait remarquer que leur méthode ne fonctionne pas lorsque la phrase est trop longue et contient plusieurs informations négatives et affirmatives.

La méthode que nous développerons ici s'approche davantage de NegExpand adapté au français. De plus, notre méthode doit fonctionner indépendamment d'un enrichissement sémantique. C'est à dire que l'utilisateur pourra rechercher un terme dans le plein texte sans nécessairement qu'il y ait un concept UMLS retrouvé.

En ce qui concerne l'extraction automatique des antécédents familiaux dans les comptes rendus

médicaux, Friedlin & McDonald [8] ont utilisé l'outil REX (REgenstrief data eXtraction tool) afin d'associer 12 maladies au contexte d'antécédents familiaux avec une précision de 97% et un rappel de 96%. Goryachev et al. [9] ont proposé une méthode afin de déterminer le contexte d'antécédents familiaux des concepts UMLS extraits. Cette méthode consiste à découper le compte rendu en sections (« antécédents », « antécédents familiaux », etc.), puis chaque section est découpée en phrase puis en groupe nominal. Enfin il détermine par des règles de décision si la maladie détectée dans le groupe nominale est associée à un parent du patient ou au patient lui-même avec une sensibilité de 97,2% et une spécificité de 99,7%. Lewis et al. [10] quant à lui prend l'intégralité du texte sans passer par un découpage en section. Il détecte les chaînes de dépendances entre la relation de parenté et une maladie grâce au Stanford NLP Parser. Cela lui permet d'extraire des extraits de phrase contenant le lien entre un parent du patient et une maladie avec une précision de 61% et un rappel de 51%.

Notre méthode de détection des antécédents familiaux s'inspire davantage de la méthode NegEx utilisée pour la détection de la négation, car elle est beaucoup plus simple à mettre en œuvre et très facilement évolutive.

3 Matériel et méthodes

3.1 Prétraitement des textes

Il est indispensable d'effectuer un « nettoyage » des documents avant l'extraction des contextes et de la négation. En effet, les textes sont issus d'une conversion automatique de documents Word en format txt, ou sont issus de documents HTML. Une première étape consiste donc à supprimer les balises HTML, et surtout à réajuster les sauts de ligne. Suivant la version de Word, la conversion au format texte induit des sauts de ligne au milieu des phrases, ce qui peut modifier l'interprétation de l'algorithme. Il s'agit donc de déterminer si un saut de ligne doit être remplacé par un espace ou par un point de ponctuation. Et inversement, certains documents contiennent des pseudos tableaux ou des listes sur deux colonnes. La distinction entre deux propositions ou syntagmes nominaux n'est donc plus représentée par un saut de ligne ou par une ponctuation mais par plusieurs espaces. Par ailleurs, une phrase peut contenir des doubles espaces liés à une erreur de saisie. Nous avons donc développé quelques règles sous forme d'expression régulière qui prend en compte ces variabilités afin de recréer la phrase sans saut de ligne ou de forcer l'ajout d'un point séparant deux phrases distinctes. Cette étape est essentielle pour la suite des traitements.

3.2 Détection du contexte familial

Nous avons établi une liste d'expressions décrivant un lien de parenté avec le patient : sœur, frère, cousin, tante, père, mère, grand parent, oncle, cousin, antécédent familial, neveu, nièce, antécédents familiaux, papa, maman. Cette liste initiale peut facilement être complétée par la suite pour la rendre plus exhaustive. Les expressions régulières tiennent compte des variations orthographiques des formes nominales : féminin, masculin, singulier, pluriel, présence ou erreur sur l'accentuation : `[^a-z] fr [éeèê] res?`

L'algorithme découpe le texte en phrases en utilisant les éléments de ponctuation et les sauts de ligne du texte. Puis pour chaque phrase il vérifie si elle contient une des expressions exprimant la notion de famille ou un lien de parenté. La phrase est alors extraite de manière à obtenir à la fin de cette étape deux textes correspondants aux deux contextes 'patient' et 'antécédents familiaux'.

Nous avons choisi d'attribuer l'antécédent familial à l'intégralité de la phrase plutôt qu'à une sous-partie de la phrase. En effet, nous avons testé les deux méthodes, et il apparaît plus que l'auteur du

compte rendu décrit les antécédents familiaux d'un seul bloc sans évoquer les données cliniques du patient. En considérant les propositions, nous attribuons trop souvent des antécédents familiaux au patient lui-même car le lien de parenté n'est citée qu'une fois dans la phrase. Ainsi afin de réduire le nombre de faux négatifs dans la détection des antécédents familiaux et le nombre de faux positifs dans la recherche de patients, nous avons décidé d'utiliser le découpage par phrase.

3.3 Détection de la négation

La détection de la négation ressemble davantage à la méthode NegExpand. Il existe une grande hétérogénéité stylistique entre les comptes rendus médicaux, suivant les auteurs et les services. En raison de cette complexité syntaxique des comptes rendus dans lesquels une phrase peut exprimer plusieurs informations (affirmatives et négatives), nous avons créé une liste de règles permettant de découper ces phrases en se basant sur les expressions suivantes : mais ; pour ; qui ; entre ; car ; , [^d] ; sans ; lequel : laquelle ; hormis ; parce qu ; bien qu ; en dehors : malgré ; en raison de.

Nous découpons après une virgule uniquement si elle n'est pas suivie de la lettre « d », cela afin de ne pas découper les listes qui suivent une négation : « le patient n'a pas de diabète, d'insuffisance cardiaque, de cholestérol ». Nous avons créé une liste d'expressions régulières permettant de détecter une tournure négative. Les règles de détection de la négation prennent en compte des formes variées de la négation : pas [a-z]* de, absence de, oriente pas vers, exclure, non, aucun, absence, absent, negati, elimine, infirme etc. A cela nous avons ajouté des règles pour prendre en compte les doubles négations (« pas d'exclure », « pas d'éliminer ») ou les expressions exprimant une forte affirmation par l'utilisation de la négation : « pas de doute », « pas de soucis », « pas de problème ». Nous avons aussi tenu compte de l'expression de la « normalité » surtout utilisée en génétique : « gène OPA1 normal » qui signifie « non muté ». Nous avons donc considéré le terme « normal » comme l'expression de la négation d'un symptôme, d'une mutation ou d'un dysfonctionnement.

Nous considérons le triplet : texte-contexte-certitude afin de décrire le contenu d'un document. Suivant la complexité du document il sera décrit par plusieurs triplets, avec un maximum de 6 triplets :

[texte] - texte intégral -1 : le texte intégral

[texte] - texte intégral -0 : les sous phrases négatives du texte intégral

[texte] - texte patient -1 : le texte patient

[texte] - texte patient -0 : les sous phrases négatives du texte patient

[texte] - antécédents familiaux -1 : les antécédents familiaux

[texte] - antécédents familiaux -0 : les sous phrases négatives des antécédents familiaux

3.4 Enrichissement sémantique

Nous avons extrait le sous-ensemble français du Metathesaurus UMLS. Tous les concepts, leurs synonymes et leurs déclinaisons sont listés dans une table puis triés par ordre décroissant de longueur de chaîne de caractères. La présence de chaque terme ou expression est testée dans le texte pour chaque triplet texte-contexte-certitude.

3.5 Intégration à l'entrepôt de données

L'entrepôt de données a été développé sous Oracle, en langage PHP, HTML, jQuery. Le schéma relationnel d'origine de la base de données a été modifié afin de pouvoir stocker les parties de texte reliées au contexte et au niveau de certitude et aux concepts éventuellement extraits du texte (**Erreur ! Source du renvoi introuvable.**). Nous avons donc (i) une table contenant le document original à afficher, (ii) une table contenant le texte (indexé par Oracle text, et composé des phrases et sous phrases extraites par notre méthode TAL) en lien avec le contexte (patient, antécédents familiaux) et le niveau de certitude (0 ou 1) soit le triplet « texte-contexte-certitude », (iii) une table contenant les concepts extraits de chaque triplet. Oracle text indexe aussi la liste des synonymes de chaque concept UMLS afin de faciliter l'expansion sémantique d'une requête.

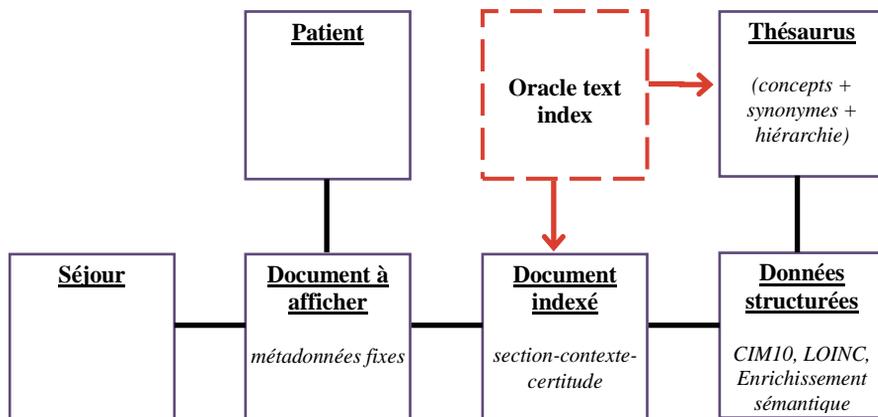


Figure 1 : schéma relation de l'entrepôt de données

3.6 Intégration au moteur de recherche

Le moteur de recherche multicritère (**Figure 2**) permet de créer plusieurs sous requêtes atomiques. Chaque sous requête peut être une recherche plein texte, une recherche sur donnée structurée (CIM10, LOINC) ou une recherche sur donnée démographique (sexe du patient, âge). L'intersection entre ces sous requêtes se fait soit sur un séjour du patient, soit sur l'ensemble du dossier patient. L'utilisateur peut préciser si une sous requête correspond à un critère d'exclusion ou d'inclusion, et éventuellement sur un contexte particulier (s'il a été détecté : motif, diagnostic etc.). La recherche plein texte exploite le module Oracle Text fourni par défaut dans Oracle 11g. Il permet d'utiliser des opérateurs booléens (or, and, not) ainsi que des opérateurs spécifiques (near, fuzzy etc.). Dorénavant l'utilisateur peut préciser si la recherche peut se faire sur les antécédents familiaux, sur le patient ou sur l'intégralité du compte rendu. Par défaut, le moteur prend en compte la détection de la négation en soustrayant du résultat les patients dont le terme recherché est présent dans le texte dont la certitude est à 0, c'est à dire considéré comme une négation. L'utilisateur peut désactiver ce critère s'il veut obtenir tous les documents qui parlent du concept même avec une négation.

L'enrichissement sémantique est aussi optionnel pour l'utilisateur, il permet d'étendre la recherche non seulement aux concepts synonymes mais aussi aux concepts fils : par exemple la requête « insuffisance cardiaque » permettra de retrouver les valvulopathies. Vu que l'enrichissement sémantique est relié au triplet texte-contexte-certitude, si le texte contient « le patient n'a pas de

valvulopathie», ce ne sera pas un critère d'exclusion du résultat, parce que cela ne sous entend pas qu'il n'a pas d'autres formes d'insuffisance cardiaque dans une autre partie du texte. La relation de subsomption n'est alors pas utilisée dans ce cas.

The screenshot displays a search engine interface with the following components:

- Top Bar:** "Recherche rapide : Intégralité de l'entrepôt (18997 patients)" with a "Rechercher rapide" button.
- Search Form:** "Formulaire de recherche : Intégralité de l'entrepôt (18997 patients)". It includes sections for "Requêtes enregistrées", "Critères généraux", and "Recherche en Full-text".
- Search Criteria 1:** A search box containing "atrophie optique" with a patient count of 235. It includes options for "Inclure", "Type de doc", "Patient", "Titre", and "UF".
- Search Criteria 2:** A search box containing "diabete" with a patient count of 11314. It includes similar options to the first criterion.
- Structured Search:** A section titled "Recherche structurée" with a filter box and a list of concepts:
 - Actes
 - Diagnostics
 - + I-certaines maladies infectieuses et parasitaires (# 3259) Ajouter
 - + XX-causes externes de morbidite et de mortalite (# 586) Ajouter
 - + XXI-facteurs influant sur l etat de sante et motifs de recours aux services de sante (# 9405) Ajouter
 - + III-maladies du sang et des organes hematopoiétiques et certains troubles du systeme immunitaire (# 2275) Ajouter
 - + IV-maladies endocriniennes, nutritionnelles et metaboliques (# 3210) Ajouter
 - + XXII-Codes d utilisation particuliere (# 584) Ajouter
 - + V-troubles mentaux et du comportement (# 1445) Ajouter
 - + VI-maladies du systeme nerveux (# 2351) Ajouter
- Buttons:** "Inclure", "Exclure", and "Annuler".
- Right Panel:** "Requête" and "Statistiques" tabs. It contains a list of search criteria with patient counts and checkboxes:
 - Age du patient à l hospit : de 0 à 20
 - Les documents (avec enrichissement sémantique) contenant l'expression "atrophie optique " spécifiquement dans la partie texte_patient (235 patients) [X]
 - Les documents (avec enrichissement sémantique) contenant l'expression "diabete" spécifiquement dans la partie texte_patient (11314 patients) [X]
 - Absence de VI-maladies du systeme nerveux dans Diagnostics (3011 patients) [X]

Figure 2 : Moteur de recherche

3.7 Evaluation

Nous avons tout d'abord réalisé une évaluation intrinsèque de la détection de la négation et des antécédents familiaux associée à l'enrichissement sémantique. Pour cela nous avons exécuté notre algorithme sur 130 comptes rendus médicaux. Nous obtenons un tableur qui donne pour chaque phrase les concepts extraits, la notion de négation ou d'affirmation et le contexte (patient ou famille). Les 1700 lignes du tableur ont été évaluées manuellement afin de calculer la précision et le rappel pour la détection de la négation d'une part, et la détection du contexte d'autre part.

Pour évaluer l'apport dans le moteur de recherche, nous avons exécuté notre algorithme sur l'intégralité de l'entrepôt de données de l'hôpital Necker Enfants Malades : 189 997 Patients et 653 047 documents. Les documents sont issus de plusieurs sources différentes (CR hospitalisation, consultation, radiologie, foetopathologie, les résumés PMSI sont volontairement exclus) et plusieurs formats (compte rendu Word, compte rendu HTML, compte rendu texte). Les textes sont dé-identifiés. Nous avons testé le moteur de recherche sur 3 requêtes plein textes : « crohn and diabete », « NPHP1 » et « lupus and insuffisance renale ». Une technicienne médicale a évalué le

nombre de patients vrais positifs, faux positifs pour les 3 requêtes avec et sans la détection de la négation, puis avec la détection des antécédents familiaux. Pour évaluer les résultats, la technicienne médicale utilise l'interface web de l'entrepôt de données permettant de vérifier rapidement le contexte des termes retrouvés au survol de la souris (**Figure 3**). En cliquant sur le document elle peut aussi accéder à l'intégralité du compte rendu afin de vérifier en cas de doute s'il s'agit de vrai positif ou non. Pour le calcul du rappel, nous considérons que le moteur de recherche plein texte sans prise en compte de la négation ni des antécédents familiaux renvoie l'intégralité des vrais positifs. La technicienne médicale génère ainsi le gold standard pour une requête en séparant les vrais positifs des faux positifs. A partir de ces données nous déterminons la précision et le rappel pour la détection de la négation puis pour la détection des antécédents familiaux.

Requête								Statistiques		Résultats			
Ouvrir les concepts				Mettre tous les patients dans le panier									
ipp	nom	prenom	Age aujourd'hui	S	CP	Actions	Panier	Doc					
NIP	NOM_PATIENT1	PRENOM	65 ans	F	75018			Voir					
NIP	NOM_PATIENT2	PRENOM	16 ans	F	75006			Voir					
NIP	NOM_PATIENT3	PRENOM	16 ans	M	93210			Voir					

Figure 3 : Aperçu du résultat

4 Résultats

L'évaluation intrinsèque de la détection de la négation et du contexte associée à l'enrichissement sémantique est représentée dans les tableaux 1 et 2. Le tableau 1 présente la détection du contexte d'un concept.

Le calcul de la précision = Vrais Positifs / (Vrais Positifs + Faux Positifs).

Le calcul du rappel ou sensibilité = Vrais Positifs / (Vrais Positifs + Faux Négatifs).

La spécificité = Vrais Négatifs / (Vrais Négatifs + Faux Positifs).

En considérant l'Histoire familiale comme l'élément positif, nous obtenons une précision de 95,2%, un rappel de 94,4% et une spécificité de 98,7%.

Tableau 1 : Evaluation de la détection du contexte

Contexte	Vrai	Faux
Histoire familiale	338 (VP)	17 (FP)
Patient	1325 (VN)	20 (FN)
Total	1663	37

Le tableau 2 représente la détection de la négation et de l'affirmation associée au concept. En

considérant la négation comme l'élément positif, nous obtenons une précision de 95,6%, un rappel de 92,8% et une spécificité de 99,1%

Tableau 2 : Evaluation de la détection de la négation

Contexte	Vrai	Faux
Négation	283 (VP)	13 (FP)
Affirmation	1382 (VN)	22 (FN)
Total	1665	35

Sur les 653 047 documents contenus dans l'entrepôt de données, nous avons pu extraire des antécédents familiaux sur 110 774 documents. Le tableau 3 montre la répartition du nombre de textes sur le couple contexte-certitude. Le tableau 4 montre le nombre de concepts UMLS extraits des textes par contexte et certitude. Le niveau de certitude à 1 correspond à l'intégralité du texte, le niveau de certitude à 0 correspond uniquement aux parties de phrases exprimant la négation.

Tableau 3 : Nombre de documents par contexte et certitude

Contexte	Certitude	Nombre de textes
Texte Intégral	1	653 047
Texte Intégral	0	540 147
Patient	1	653 047
Patient	0	541 217
Antécédents familiaux	1	110 774
Antécédents familiaux	0	5 780

Tableau 4 : Nombre de concepts UMLS par contexte et certitude

Contexte	Certitude	Nombre de concepts
Texte Intégral	1	9 027 310
Texte Intégral	0	1 663 988
Patient	1	8 947 279
Patient	0	1 354 236
Antécédents familiaux	1	80 031
Antécédents familiaux	0	4 032

Les tableaux 5, 6, 7 indiquent le nombre de patients Vrai Positifs (VP), Faux Positifs (FP), Vrai Négatifs (VN), Faux Négatifs (FN), la précision et le rappel pour chacune des requêtes dans 3 cas de figure : la recherche plein texte sans prise en compte de la négation ou des antécédents familiaux (texte intégral), la recherche plein texte avec exclusion de la négation (texte intégral - négation), la recherche plein texte avec exclusion de la négation et exclusion des antécédents familiaux (texte intégral - négation - antécédents familiaux).

L'ajout du module gérant la négation apporte un gain de précision (de 0,75 à 0,83 pour « lupus et insuffisance rénale »), même si pour la requête « crohn and diabete » il y a une perte du rappel (de 1 à 0,83). Ceci s'explique par une tournure négative n'exprimant pas de négation et que nous n'avions pas prévu dans les exceptions : « je ne reviens pas sur le crohn ». La perte du rappel lors de la prise en compte des antécédents familiaux pour la recherche sur le lupus et l'insuffisance rénale (de 1 à 0,95) s'explique par l'utilisation au sein d'une même phrase du patient et de son frère, l'un ayant la maladie l'autre non. La phrase a alors été considérée comme relatif aux antécédents familiaux et non au patient.

La détection des antécédents familiaux augmente aussi la précision sans modifier le rappel dans les 3 tests effectués, excepté pour la recherche du gène « NPHP1 » qui n'a pas été testé dans les familles des patients. On voit par ailleurs que la précision de la requête « NPHP1 » reste faible malgré une amélioration de la précision de 0,42 à 0,52. Dans notre expérience à l'Hôpital Necker, les utilisateurs qui requêtent avec un nom de gène recherchent les patients avec une mutation sur ce gène. Nous avons donc considéré dans notre évaluation que le doute et la prescription du test génétique sont des faux positifs. Si nous considérons ces éléments comme vrai positif nous améliorons la précision à 0,81 pour « NPHP1 » comme indiqué dans le tableau 8.

Tableau 5 : Précision et rappel pour « lupus and insuffisance renale »

Lupus and insuffisance renale	VP	FP	VN	FN	Précision	Rappel
Texte Intégral	77	26	0	0	0,75	1,00
Texte Intégral - négation	77	16	10	0	0,83	1,00
Texte Intégral - négation - antécédents familiaux	73	7	19	4	0,91	0,95

Tableau 6 : Précision et rappel pour « Crohn and diabete »

Crohn and diabete	VP	FP	VN	FN	Précision	Rappel
Texte Intégral	12	82	0	0	0,13	1,00
Texte Intégral - négation	10	57	25	2	0,15	0,83
Texte Intégral - négation - antécédents familiaux	10	10	72	2	0,50	0,83

Tableau 7 : Précision et rappel pour « NPHP1 »

NPHP1	VP	FP	VN	FN	Précision	Rappel
Texte Intégral	11	15	0	0	0,42	1,00
Texte Intégral - négation	11	10	5	0	0,52	1,00
Texte Intégral - négation - antécédents familiaux	11	10	5	0	0,52	1,00

Tableau 8 : Précision et rappel pour « NPHP1 » avec doute et prescription

NPHP1 (doute et prescription)	VP	FP	VN	FN	Précision	Rappel
Texte Intégral	17	9	0	0	0,65	1,00
Texte Intégral - négation	17	4	5	0	0,81	1,00
Texte Intégral - négation - antécédents familiaux	17	4	5	0	0,81	1,00

5 Discussion

L'évaluation intrinsèque de la détection du contexte familiale montre une précision (95,2%), un rappel (94,4%) et une spécificité (98,7%) relativement élevés. Nos scores sont sensiblement inférieurs aux résultats de Goryachev & al. [9] avec une sensibilité de 97,2% et une spécificité de 99,7%. Notre rappel est un peu supérieur au rappel de Friedlin et al. [8] de 93%, mais notre précision est légèrement inférieure (97%). Notre évaluation diffère de celle de Goryachev et al parce qu'ils effectuent l'analyse sur une extraction automatique de la section du compte rendu qui concerne l'histoire familiale. Friedlin évalue lui aussi son algorithme uniquement sur les sections concernant l'histoire familiale, mais avec une détection exacte des sections car précisément labélisées comme histoire familiale. Notre évaluation porte sur l'intégralité de chaque compte rendu. De plus, leur algorithme repose sur la langue anglaise alors que la nôtre est spécifique au français.

L'évaluation intrinsèque de la détection de la négation montre une précision (95,6%), un rappel (92,8%) et une spécificité (99,1%) relativement élevés. Les résultats obtenus par Deléger et Grouin [7] montrent une précision et un rappel respectivement de 84,4% et 83,3% pour NegEx, et 88,8% et 84,6% pour leur adaptation au français de NegEx.

L'apport de notre méthode pour améliorer la qualité du moteur de recherche est variable suivant les requêtes. On voit toutefois qu'il y a à chaque fois un gain de la précision avec peu ou pas de perte de rappel. Les erreurs sont généralement liées à un problème de découpage des phrases lors du nettoyage des textes. Si une phrase se termine sans ponctuation par un saut de ligne et qu'à la ligne suivant la 1ere lettre du mot est en majuscule, l'algorithme considère que ce sont 2 phrases distinctes. Mais cette méthode a ses limites pour les noms propres qui commencent par une majuscule. En effet, pour Crohn, certaines phrases sont découpées de telle manière que le terme Crohn est précédé d'un saut de ligne :

```
« La sœur du patient a une maladie de  
Crohn »
```

L'algorithme considère alors ces deux phrases distinctes : « La sœur du patient a une maladie de » et « Crohn ». L'antécédent familial ne prend alors pas en compte « Crohn » qui reste associé au patient.

L'évaluation nous montre aussi que certains résultats Faux Positifs sont dus à une expression du doute ou d'un examen à réaliser. Par exemple, les médecins font mention du gène NPHP1 non par pour exprimer le fait qu'il est muté ou non muté mais pour émettre une suspicion de mutation ou une prescription d'examen génétique sur ce gène.

Nous pensons pouvoir encore améliorer la qualité de la détection de la négation en ajoutant de nouvelles règles et exceptions découvertes au cours de l'évaluation. Nous souhaitons ajouter un niveau intermédiaire de certitude permettant d'exprimer la suspicion ou le doute avec un score à 0,5. Nous prévoyons de développer un algorithme spécifique aux informations génétiques afin de discriminer les gènes effectivement mutés des tests génétiques prescrits et des suspicions de mutations (doute).

6 Conclusion

Nous avons pu intégrer dans l'entrepôt de données de l'hôpital Necker Enfants Malades un algorithme permettant de diminuer le bruit de la recherche plein texte dans les comptes rendus médicaux en extrayant automatiquement les antécédents familiaux des patients et les sous phrases exprimant une négation. L'interface utilisateur permet de choisir le niveau de spécificité souhaité

en fonction de la rareté ou de la complexité des requêtes effectuées. L'extraction des concepts associée à la détection de la négation et des antécédents familiaux nous permettra de travailler sur l'enrichissement des descriptions phénotypiques [11] des patients porteurs d'une anomalie génétique.

Remerciements

Nous remercions chaleureusement Malika-Anais Boujemaoui pour avoir évalué l'amélioration de la précision du moteur de recherche par la détection de la négation et des antécédents familiaux.

Références

- [1] Cuggia M, Bayat S, Garcelon N, Sanders L, Rouget F, Coursin A, Pladys P. A full-text information retrieval system for an epidemiological registry. *Stud Health Technol Inform.* 2010;160(Pt 1):491-5.
- [2] Cuggia M, Garcelon N, Campillo-Gimenez B, Bernicot T, Laurent JF, Garin E, Happe A, Duvauferrier R. Roogle: an information retrieval engine for clinical data warehouse. *Stud Health Technol Inform.* 2011;169:584-8.
- [3] Humphreys, B.L., et al., The Unified Medical Language System: an informatics research collaboration. *J Am Med Inform Assoc*, 1998. 5(1): p. 1-11.
- [4] W. Chapman, W. Bridewell, P. Hanbury, G. Cooper, and B. Buchanan. A simple algorithm for identifying negated findings and diseases in discharge summaries", 2001; *J Biomed Inform*, vol. 34, no. 5, pp. 301-310.
- [5] S. Goryachev, M. Sordo, Q.T. Zeng, and L. Ngo. Implementation and evaluation of four different methods of negation detection, 2006; Technical report, DSG
- [6] Aronow DB, Fangfang F, Croft WB. Ad hoc classification of radiology reports. *J Am Med Inform Assoc.* 1999 Sep-Oct; 6(5):393-411.
- [7] L. Deléger & C. Grouin. Detecting Negation of Medical Problems in French Clinical Notes. *IHI'12 Proceedings of 2nd ACM SIGHIT International Health Informatics Symposium.* 2012.
- [8] Friedlin J, McDonald CJ. Using a natural language processing system to extract and code family history data from admission reports. *AMIA Annu Symp Proc.* 2006:925.
- [9] Goryachev S, Kim H, Zeng-Treitler Q. Identification and extraction of family history information from clinical reports. *AMIA Annu Symp Proc.* 2008 Nov 6:247-51.
- [10] N. Lewis, D. Gruhl, and H. Yang, "Extracting family history diagnosis from clinical texts," in *International Conference on Bioinformatics and Computational Biology*, March 2011.
- [11] Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, Wang D, Masys DR, Roden DM, Crawford DC. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics.* 2010 May 1;26(9):1205-10. doi: 10.1093/bioinformatics/btq126. Epub 2010 Mar 24.

Adresse de correspondance

Institut Imagine, 24 boulevard Montparnasse, 75015 Paris
nicolas.garcelon@institutimagine.org

Avantages, effets et préoccupations reliés au dossier de santé personnel électronique (DSPé) : Une mise à jour de la littérature

Benefits, Effects and Concerns Related to the Electronic Personal Health Record: Update of the Literature

El Kebir Ghandour^{1,2}; Marie-Pierre Gagnon^{1,2,3}; Jean-Paul Fortin^{2,4}

¹*Axe Santé Publique et Pratiques Optimales en Santé CRCHUQ, Hôpital Saint-François d'Assise, Québec, Canada*

²*Centre de Recherche sur les Soins et Services de Première Ligne au CSSS-VC, Québec, Canada*

³*Faculté des Sciences Infirmières, Université Laval, Québec, Canada*

⁴*Faculté de Médecine-Département de Médecine Sociale et Préventive. Université Laval, Québec, Canada*

Résumé

Les dossiers de santé personnels électroniques (DSPé) permettent aux individus de gérer leur santé et de communiquer avec l'équipe de soins. Ces outils pourraient avoir un impact important sur l'efficacité et l'efficience des soins, suscitant ainsi de plus en plus d'intérêt envers leur implantation et leur utilisation. Cependant, les preuves scientifiques sur leurs bénéfices et effets potentiels demeurent insuffisantes. Cet article présente les résultats d'une mise à jour de la littérature scientifique sur certains aspects d'intérêt pour les décideurs qui s'impliqueraient dans l'implantation de ces systèmes au Canada. Parmi les 1070 articles identifiés dans les sept bases de données consultées, 96 ont fait l'objet d'analyse du texte intégral et 52 ont été retenus pour une analyse fine. Les résultats des études consultées demeurent encore divergents. Cependant, les preuves scientifiques sont de plus en plus évidentes pour certains aspects utiles suggérant des bénéfices pour les patients, les professionnels, les organisations et le système de santé (ex. prestation de soins préventifs, l'amélioration de la communication avec les professionnels et soutien aux soins centrés sur le patient). Particulièrement, les DSPé s'avèrent bénéfiques pour le suivi et la gestion de maladie et de l'état de santé de clientèles particulières, ou de catégories de patients considérées comme grandes consommatrices des ressources de soins dans le système de santé (ex. maladies chroniques). Toutefois, certaines études soulèvent l'importance de se préoccuper de l'alignement de ces systèmes avec le processus de soins et le travail des cliniciens lors de l'implantation, en lien surtout avec la circulation de l'information pour l'amélioration des résultats cliniques des patients.

Abstract

Electronic personal health records (ePHR) allow patients to manage their health and to communicate with the health care team. These tools could have an impact on the effectiveness and efficiency of care, giving rise to more interest in their implementation and use. However, the scientific evidence on their benefits and potential effects remain insufficient. In this paper, we

present the results of the scientific literature review update to bring out evidence concerning some interesting aspects to guide ePHR implementation in Canada. We have conducted searches in seven electronic databases and identified 1070 studies, of which 52 were included in the final analysis. Empirical studies accessed showed that evidence about several issues and concerns around ePHRs implementation remains divergent. However, it was noted that the scientific evidence is stronger for some useful aspects of implementation, suggesting benefits for patients, professionals, organizations and the health care system (preventive care delivery, patient-providers communication improvement, and patient-centered care support). In particular, ePHR show benefits for disease management and health status monitoring by specific groups (e.g. chronic diseases). However, some studies raise the importance of these systems impact on the care process and the clinical practice, related especially to the flow of information for the improvement of clinical outcomes.

Mots-clés : Dossier de santé personnel électronique, revue de la littérature, effets, avantages, impacts, adoption, utilisation, implantation, politiques publiques, Canada.

Keywords: *Electronic personal health record, literature review, effects, benefits, impacts, adoption, use, implementation, public policy, Canada.*

1 Introduction

La majorité des Canadiens veulent avoir accès à de l'information sur la santé et à des services en ligne avec une volonté de fournir plus d'informations aux prestataires de soins et services afin de recevoir des recommandations individualisées sur la façon d'optimiser leur état de santé [1].

On note dans la littérature que plusieurs résultats de santé positifs découlent de la mise en place d'approches qui combinent et intègrent des actions axées sur le prestataire de soins, le patient et ses proches, et sur l'organisation même des soins et services [2, 3]. En effet, les patients peuvent avoir un rôle central dans leurs propres soins et dans la gestion efficace de leurs maladies [4, 5] et peuvent être des partenaires avec des droits et des responsabilités [6]. Un tel rôle du patient dépend de sa participation active, qui est une exigence évidente pour l'autogestion efficace de conditions telles que les maladies chroniques (MC) [7-9] et pour l'amélioration de l'interaction avec les professionnels [10]. De plus, la participation du patient est reconnue comme un élément clé dans la réorganisation des processus de soins de santé pouvant être appliquée pour certains aspects des soins aux patients, notamment la gestion des MC et à la prise de décision clinique [11]. En effet, les patients sont reconnus comme des décideurs clés dans le processus de soins [10, 12]. la participation des patients leur permet de tirer le plus grand bénéfice des services de santé auxquels ils ont accès [13], principalement au niveau de la première ligne de soins.

D'autre part, le soutien à la participation du patient à la gestion de son état de santé peut constituer une mesure permettant d'entretenir deux approches organisationnelles, « l'approche centrée sur le patient » et « l'autonomisation ». Ainsi, d'un point de vue organisationnel, le véritable enjeu réside dans la façon de soutenir la personne dans son développement, dans son autonomisation et dans le maintien de sa capacité à gérer sa maladie et à prévenir les complications [14]. À cet égard, les technologies de l'information et de communication (TIC) sont considérées comme des outils essentiels pour faciliter l'implantation de tels modèles [15], d'autant plus que les patients au XXIe siècle sont de plus en plus connectés aux nouvelles technologies et à internet [16]. Il devient ainsi possible de soutenir la prestation de soins, essentiellement en première ligne, en initiant une nouvelle approche qui met à contribution les patients, les cliniciens et les TIC. Notamment, le recours à divers outils de communication interactifs entre le patient et

son équipe de soins permettrait de mieux supporter la coordination des soins en assurant des soins plus complets, mieux intégrés et continus [17]. À cet égard, le dossier de santé personnel électronique (DSPé) et l'e-santé (traduction du terme *e-health*, signifiant « santé électronique ») ont gagné en popularité et attirent l'intérêt des chercheurs, professionnels et celui de l'industrie dans le monde entier [18].

2 Contexte et problématique

Le dossier de santé personnel électronique (DSPé) représente une importante innovation des TIC en santé qui est de plus en plus reconnue comme nécessaire au soutien des soins centrés sur le patient, à l'autogestion et à l'utilisation efficace des ressources des systèmes de santé [19, 20]. Le DSPé est défini comme étant : « *Une application électronique qui permet aux individus d'accéder, gérer et partager les informations sur leur santé, et celle d'autres individus pour lesquels ils sont autorisés, dans un environnement privé, sécurisé et confidentiel* » [21] (p.14). Grâce au DSPé, les patients peuvent entrer leurs données de santé, communiquer avec leurs médecins et accéder aux informations dans leurs dossiers médicaux et celles d'autres sources [15, 22, 23]. Dans certains cas, le DSPé peut inclure des informations issues du dossier médical électronique (DME), du dossier clinique informatisé (DCI) de l'hôpital ou des systèmes de dossiers électroniques régionaux. Un portail patient est également un exemple de DSPé qui fournit l'accès aux dossiers de santé du patient. La distinction principale avec le DME est que le DSPé est opéré et contrôlé par le patient alors que le DME est opéré par le médecin et les autres membres de l'équipe de soins.

Plusieurs auteurs avancent des bénéfices potentiels liés à l'utilisation du DSPé aussi bien pour les patients, les professionnels, les organisations que pour le système de soins [24, 25]. En effet, l'utilisation du DSPé permettrait aux citoyens d'accumuler de l'information sur leur santé et sur la prestation de services provenant de plusieurs sources, de suivre plus régulièrement leur état de santé et d'ainsi mieux gérer leurs soins. Le DSPé peut aussi soutenir l'autogestion, la planification des soins [26-33], l'accès aux soins [26] et ultimement contribuer à l'amélioration de l'état de santé de la personne [34-37]. En facilitant la communication patient-équipe de soins, le DSPé peut contribuer considérablement à l'amélioration de l'interaction avec les professionnels [38] et de la relation patient-professionnels [24]. De plus, le DSPé peut fournir de l'aide à la décision [39] et peut agir en tant qu'outil de promotion à la prise de décision partagée [37, 40]. Ainsi, et grâce à l'intégration d'informations de santé plus individualisées aux soins, le DSPé permettrait au système de soins de santé d'évoluer vers un modèle de médecine plus personnalisée [24]. Également, l'accès facilité à l'information de santé ciblée et éventuellement adaptée aux besoins des personnes peut constituer un moyen novateur pour améliorer la littératie en santé pour les patients (deux Québécois sur trois présentent des niveaux de compétences inférieurs au niveau souhaitable pour ce qui est de la littératie en santé [41]). De plus, le DSPé peut contribuer à l'habilitation des patients à prendre plus de responsabilités et à acquérir plus de contrôle sur leur santé et leur vie [42]. Aussi, le DSPé peut contribuer à l'amélioration de la sécurité des soins et à la réduction des erreurs [26-32, 37]. Ainsi, le DSPé pourrait avoir un impact important sur l'efficacité et l'efficacités des soins en contribuant à la réduction de la complexité de la prestation de soins [43-46], à la diminution des soins redondants [37] et à une utilisation plus efficace des ressources, réduisant éventuellement les coûts du système [37, 43-46].

En raison de ses avantages potentiels, les chercheurs, décideurs et gestionnaires du système de santé ainsi que l'industrie manifestent de plus en plus d'intérêt pour l'implantation et l'utilisation des DSPé. De plus, on voit émerger au Canada un certain nombre d'initiatives qui introduisent cette solution dans leur plan d'action, soulevant ainsi la nécessité de mieux cerner l'ensemble des questions qui l'entourent. Ceci est particulièrement difficile étant donné que les

preuves scientifiques sur les avantages du DSPé et ses effets potentiels sur les pratiques professionnelles, les organisations de soins et les systèmes de santé demeurent insuffisantes.

Dans ce contexte s'inscrit notre projet de recherche, dont l'objectif est de formuler des options de politiques viables en relation à des enjeux qui sont d'intérêt pour les décideurs qui auront à se pencher sur l'implantation du DSPé par les organisations de soins au Canada. Afin d'atteindre cet objectif, nous avons réalisé l'actualisation d'une revue systématique de la littérature sur les DSPé.

Dans les sections suivantes, nous présenterons les résultats en lien avec la mise à jour de la littérature scientifique sur le DSPé de notre projet ainsi qu'une discussion des résultats saillants, utiles à la formulation de notes de politique autour de l'implantation d'un DSPé dans le système de soins au Canada.

3 Objectifs

Cette étude est une mise à jour sélective d'une revue de la littérature réalisée récemment par Archer et collaborateurs [47] sur les DSPé qui cible certains aspects dans le but de guider la prise de décisions liée à la mise en œuvre du DSPé. Ainsi, nous nous sommes concentrés sur quelques aspects jugés les plus utiles pour les décideurs, en nous attardant principalement à des éléments de preuves scientifiques en lien avec: **1)** les effets et avantages du DSPé ; **2)** les principales préoccupations envers l'utilisation ; **3)** l'acceptation, l'adoption et l'utilisation du DSPé par les utilisateurs (patients et professionnels); et **4)** des questions liées à la mise en œuvre du DSPé. Aussi, nous avons examiné les évidences scientifiques en lien avec les perceptions, les attitudes des patients et professionnels qui peuvent constituer des barrières ou des facilitateurs à l'utilisation du DSPé.

4 Méthodologie

Dans cette mise à jour de la littérature, nous nous sommes concentrés sur les articles scientifiques revus par les pairs décrivant un DSPé, publiés entre avril 2010 et septembre 2013. Nous avons réalisé la recherche d'études dans sept bases de données électroniques : Cochrane Library; Embase; EBSCO; PsYINFO; PubMed; Scopus; Web of Science. Les mots-clés utilisés sont: electronic Personal Health Records; ePHR; Patient Internet Portal, Patient Portal; Patient Shared/ Held/ Record; Patient Accessible Records; Personal Medical Record; ePHR Effects / Benefits; ePHR Purposes; ePHR Acceptance; ePHR Adoption/ Use (healthcare providers, patients); ePHR Barriers to Adoption and Use; ePHR Satisfaction.

Nous avons sélectionné les études empiriques, qualitatives et quantitatives, portant sur des systèmes de DSPé, publiées en anglais, français ou espagnol, et réalisées dans des pays de niveau socioéconomique comparable au Canada. Nous avons limité notre sélection aux DSPé attachés (tethered) ou intégrés permettant des communications asynchrones entre patient et professionnels, ainsi qu'un accès du patient à ses données médicales électroniques, par opposition aux DSPé autonomes (stand-alone). Les critères d'inclusion sont présentés à la figure 1 (voir Annexes). Une grille d'extraction a été élaborée en se basant sur la littérature pertinente [25, 47] et a servi à extraire les données manuellement dans une base de données Excel. Nous avons procédé principalement à une analyse qualitative des données.

5 Résultats de la revue de littérature

Parmi les 1070 articles identifiés dans les bases de données électroniques, 96 ont fait

l'objet d'analyse du texte intégral. Finalement, 52 études ont été retenues comme étant éligibles et ont fait l'objet d'une analyse fine (Figure 1 des Annexes). Au total, on note que 47 des 52 études retenues (90,4%) ont été réalisées aux États-Unis (Tableau 1 des Annexes). La majorité des études ont été menées auprès de patients (71%) (Tableau 1 des Annexes). Les études ont utilisé une méthodologie à prédominance quantitative (67%) et par sondage (61 %) (Tableau 1 des Annexes).

Nous nous sommes concentrés lors de l'extraction des données, principalement sur les aspects en lien avec:

1) Les avantages et effets du DSPé, liés essentiellement : aux résultats cliniques des patients (6 études), à la prestation des services préventifs (4 études), à l'amélioration de la sécurité des patients (2 études), à l'autonomisation des patients (7 études), à la communication patient-professionnels de soins (6 études), au soutien à l'autogestion (2 études), à l'utilisation des services de santé et l'expérience de soins (6 études), à l'aspect de soins centrés sur le patient (2 études), à la qualité des soins (2 études) ainsi qu'au changement de processus (une étude) ;

2) L'acceptation, l'adoption et l'utilisation du DSPé auprès des patients (26 études), auprès des parents (4 études) et des médecins et autres professionnels (8 études) ;

3) Des questions liées à l'implantation de ces systèmes (15 études) traitant de la confidentialité, de la sécurité et de la vie privée des patients ainsi que des aspects liés aux notions de la littératie en santé et la fracture numérique ;

4) Les principales préoccupations envers l'utilisation ;

Une présentation des études analysées selon la catégorisation retenue se retrouve au Tableau 2 des Annexes.

Dans la section qui suit, nous allons présenter les résultats de notre revue liés à chacune des dimensions d'analyse retenue dans la catégorisation initiale.

5.1 Avantages et effets du DSPé

- **Amélioration des résultats cliniques :** Six études (deux essais cliniques randomisés (ECR) [48, 49], une étude à méthodologie mixte [50], une étude de cas [51] et deux études observationnelles [52, 53]) ont évalué les effets de l'utilisation de portails Internet sur la santé des patients. Quatre de ces études ont été réalisées auprès de personnes diabétiques [48, 50, 52, 53], une auprès de personnes vivant avec le VIH [51] et une auprès de personnes hypertendus [49]. Les études réalisées auprès des diabétiques notent une baisse significative des taux moyens d'hémoglobine glyquée (HbA1c) chez les utilisateurs d'un portail patient [48, 50, 52, 53], qui n'a pas été maintenue après 12 mois [48]. Cependant, seulement l'utilisation de la messagerie sécurisée (MS) a été significativement associée au contrôle glycémique [50]. Précisément, une plus grande utilisation de la MS pour planifier un rendez-vous était significativement associée à des valeurs d'HbA1c plus faibles [50]. Les deux études ayant évalué le rôle de l'utilisation du DSPé sur le taux du cholestérol des lipoprotéines de basse densité (LDL) notent des résultats divergents [48, 52]. Toutefois, une étude rétrospective a montré que des valeurs plus faibles de la pression artérielle (PA) ont été notées chez les adultes diabétiques utilisant un DSPé [53]. Par contre, aucun effet sur le contrôle de la PA n'a été noté dans deux autres études [48, 49]. De même, chez des patients vivant avec le VIH, aucune différence significative n'a été retrouvée entre les résultats cliniques des deux groupes (utilisateurs et non utilisateurs) [51].

- **Prestation de services préventifs requis :** Quatre études, soient trois ECR [54-56] et une étude observationnelle rétrospective [53], ont évalué la prestation de services préventifs chez les utilisateurs du DSPé. Ainsi, les ECR ont suggéré que l'utilisation de DSPé a été significativement associée à l'amélioration de la prestation de services préventifs fondés sur les guides de pratique et sur les taux de vaccination [54-56]. En outre, les utilisateurs étaient plus susceptibles d'être à jour

sur tous les services préventifs pour une longue durée (allant jusqu'à 16 mois) [54]. Particulièrement, les diabétiques utilisateurs d'un DSPé étaient plus susceptibles de bénéficier des tests recommandés (HbA1c, examen de la vue avec dilatation, vaccination antipneumococcique et test de la microalbuminurie) [53]. De plus, un ECR note que ces différences étaient significatives alors que les utilisateurs du DSPé ont effectué moins de visites en personne par rapport au groupe de contrôle [55]. Aussi, les enfants du groupe d'utilisateurs présentent un meilleur taux de vaccination pour les vaccins recommandés. Cependant, les résultats d'un ECR suggèrent que les effets de l'utilisation du DSPé pour la vaccination ne différaient pas significativement par rapport à d'autres interventions auprès de la population (contact par courrier postal) [56].

- **L'amélioration de la sécurité des patients :** Deux études récentes, un ECR [57] et une étude de cohorte [58], se sont intéressées à l'évaluation de la sécurité des patients en lien avec l'utilisation du DSPé. L'utilisation de la messagerie sécurisée d'un portail Web a permis de détecter des effets liés aux médicaments, cliniquement importants et classés graves, chez des patients ayant récemment quitté l'hôpital [58]. Cependant, les résultats de l'ECR n'ont pas montré de différence statistiquement significative entre les deux groupes en ce qui concerne les taux d'événements indésirables évitables ou pour des effets indésirables graves [57].

- **L'autonomisation des patients :** Le rôle du DSPé dans l'autonomisation des patients a été traité dans sept études [37, 49, 51, 59-62]. Des études qualitatives récentes notent qu'aussi bien les patients [51], les médecins [37, 61] que d'autres professionnels de la santé [60] estiment que le DSPé pourrait permettre l'autonomisation des patients. En effet, le DSPé présente un potentiel d'autonomisation qui est significativement corrélé à l'activation du patient [62]. Cependant, les résultats de la mesure empirique de ce concept à l'aide d'échelle de mesure validées (*Patient Activation Measure* et *Patient Empowerment Scale*), n'ont pas montré d'effet d'une ampleur suffisante pour impliquer des différences significatives entre les groupes (utilisateurs et non utilisateurs) [49].

- **Soutien à l'autogestion :** Deux études ont évalué l'utilisation du DSPé dans le soutien à l'autogestion [60, 63]. L'évaluation de l'expérience de soins chez des patients diabétiques et de leur gestion de leur maladie en ligne en collaboration avec les médecins, montre que l'utilisation du portail contribue à augmenter l'accès à l'information ainsi que l'engagement dans les soins [60, 63]. Aussi, le portail a été perçu très utile à l'amélioration de la prise de conscience des patients de leurs états de santé grâce à l'accès aux informations de santé crédibles [63]. En permettant le suivi de leur état de santé, l'outil a souvent encouragé une meilleure gestion de la maladie [60].

- **Amélioration de la communication patient-professionnels :** Six articles (2 enquêtes transversales [60, 64] et quatre études qualitatives [37, 60, 65, 66]) ont exploré comment le DSPé pouvait améliorer la communication entre les patients et les professionnels. Aussi bien les patients que les professionnels participant aux études qualitatives perçoivent que le DSPé pourrait contribuer à l'amélioration de la communication patient-professionnels [37, 60, 65, 66]. Pour les médecins, le DSPé pourrait devenir un outil importants pour la communication des patients avec les cliniciens et contribuer au renforcement du partenariat entre patient-professionnels [37]. Cependant, certains auteurs notent des préoccupations soulevées par les professionnels qui estiment que le DSPé peut être source de difficulté pour interpréter, documenter et répondre aux messages des patients [60]. Toutefois, une enquête transversale menée auprès d'utilisateurs et non utilisateurs de DSPé note qu'il n'y a pas de différence entre ces groupes concernant la valeur du DSPé pour la communication avec les médecins [64].

- **Utilisation des services de santé et expérience de soins :** Également, six études (3 ECR [49, 55, 57]; une enquête transversale [67] et deux études de cohorte [51, 68]) ont évalué l'utilisation des services de santé chez les utilisateurs de DSPé. Seuls Nagykaldis et collaborateurs ont noté que les utilisateurs avaient moins de visites durant la période étudiée par rapport à ceux

du groupe témoin, alors qu'ils bénéficiaient de façon significative en termes de prestation de services préventifs [55]. Cependant, deux ECR [57, 69] et une étude de cohorte [51] ne notent pas d'effet et suggèrent que l'accès au DSPé seul n'a pas permis de changer la fréquence d'utilisation des services de santé chez les utilisateurs (nombre et durée d'hospitalisations, visites en salle d'urgence et en ambulatoire et appels téléphoniques) [51, 57, 69]. Ce résultat a été aussi noté auprès des parents d'enfants atteints de maladies chroniques utilisant un portail patient intégré [67]. Paradoxalement, Palen et collaborateurs ont noté que les utilisateurs ont montré une augmentation significative des taux par an de visites de bureau et d'appels téléphoniques [68].

- **Soutien aux soins centrés sur le patient :** Cet aspect est évalué par deux études (un ECR [55] et une étude qualitative [70]) qui notent que l'utilisation du DSPé peut contribuer à des soins centrés sur le patient. En effet, une mesure à l'aide d'une échelle validée, note une augmentation significative de 0,32 point dans le score composite de centralité patient après l'utilisation du DSPé, tandis que, dans le groupe témoin, les patients affichent une baisse de 0,43 point du même score [55]. De même une étude qualitative réalisée auprès de responsables dans sept grandes organisations ayant implanté un DSPé aux Etats-Unis afin d'offrir des recommandations pratiques et de proposer un modèle de politique pour l'implantation du DSPé, soulignent que l'utilisation du DSPé dans ces organisations supporte bien l'approche de soins centrés sur le patient [70].

- **Effets liés à la qualité des soins :** Deux études ont évalué la qualité des soins liée à l'utilisation du DSPé (une étude de cohorte [53] et une enquête transversale [51]). Les adultes diabétiques montrent des valeurs de mesure de la qualité de soins plus positives chez les utilisateurs par rapport aux non-utilisateurs [53]. Toutefois, Gordon et collaborateurs notent qu'il n'y a pas de différences significatives dans la perception de la qualité des soins pour les utilisateurs vivant avec le VIH/SIDA [51].

- **Effets sur les processus cliniques :** Une étude qualitative menée par Nazi et collaborateurs note que le DSPé permettait la délégation de certaines fonctionnalités aux patients (entrée et mise à jour de données de santé, responsabilité de son propre suivi, accès à du matériel d'éducation à la santé) ainsi que l'intégration avec le système d'information clinique des cliniciens pourraient être source de changement de pratiques et de processus [66]. Ainsi, il est important de se préoccuper de l'alignement de ces systèmes avec le processus de soins et le travail des cliniciens lors de l'implantation. En effet, le manque d'alignement avec les processus (ex. obstacles à la circulation de l'information) pourrait constituer une barrière à l'endossement, l'adoption et l'utilisation du DSPé aussi bien par les patients que par les professionnels [66].

5.2 Principales préoccupations

Une multitude de préoccupations, perçues différemment par les patients et les professionnels, sont notées dans plusieurs des études retenues. La nature, la qualité et la gestion des données, le retentissement sur la relation patient-professionnels et sur la gestion de l'organisation de soins en constituent les trois principales.

- **La nature, la qualité et la gestion des données :** Les préoccupations en lien avec la donnée contenue dans le DSPé varient entre les patients et les professionnels. Concernant les perceptions des patients, les principaux soucis retenus sont : la cohérence, la clarté, la facilité d'interprétation des instructions, les tâches nécessaires, les difficultés liées aux tâches d'entrée des données et de l'organisation de l'information, les capacités cognitives utiles, le risque de mal interpréter les données, la sécurité et la confidentialité ainsi que le souci de partager ses propres données avec les fournisseurs privés de DSPé [61, 64, 66, 71-75]. Pour les professionnels, essentiellement les médecins, les inquiétudes envers la gestion de la donnée sont principalement : la qualité, la précision et l'organisation de l'information, l'accès du patient aux données du DME (surtout les notes des médecins), une interprétation erronée des données médicales par les patients,

la sécurité et la confidentialité et l'atteinte à la vie privée des patients, surtout en lien avec le partage de données avec des fournisseurs privés, des tâches longues avec augmentation de la charge de travail, l'absence de mode de rémunération et de remboursement pour le temps investi pour la consultation du DSPé, la responsabilité liée à l'obligation de consulter toutes les informations saisies par les patients et la responsabilité légale des médecins envers l'information contenue dans le DSPé [37, 60, 61, 64, 66, 71, 72, 76-78].

- **La relation patient-professionnels :** Plusieurs études soulèvent que l'utilisation du DSPé peut être source de changement du médium de la communication et de la relation du patient avec les professionnels, les rendant moins personnelles avec un risque accru pour le caractère thérapeutique de la relation patient-médecin [37, 50, 60, 61, 64-66, 78]. Cependant, la nature de la relation patient-professionnels est l'une des conditions d'adoption de ces outils. Ainsi, l'utilisation du DSPé peut affecter positivement la relation patient-professionnels et vice versa [65].

- **La gestion du travail dans l'organisation :** Les préoccupations d'ordre organisationnel sont surtout liées aux rôles, responsabilités et obligations envers le DSPé ainsi que les conséquences légales et juridiques liées à la saisie, la conservation et le partage des données contenues dans le DSPé, le besoin en pratique d'intégration réglementée et contrôlée avec le DME, l'augmentation de la charge de travail par ces tâches au détriment du temps personnel des médecins ou de la durée des rencontres face à face, l'absence de modèles de rémunération pour ces services ainsi que les coûts du système et le type de paiement pour les patients [37, 65, 78].

5.3 Attitudes et perceptions

Le groupe d'articles ayant discuté des attitudes et des perceptions des utilisateurs (patients [64, 76, 79-85], parents d'enfants [83] et professionnels [37, 60, 61, 78, 86]) était hétérogène. En général, les différents groupes manifestent des perceptions positives envers le DSPé. Les professionnels perçoivent qu'il pourrait être utile pour les deux catégories d'utilisateurs (patients et professionnels), mais restent moins susceptibles que les patients d'anticiper ses bénéfices et ses avantages [60].

Chez les patients : La plupart des participants estimait que les fonctionnalités offertes par le DSPé sont importantes ou très importantes, principalement pour réduire le risque d'oubli d'informations importantes (dates, actes médicaux ou médicaments, etc.), pour renforcer leur autonomie dans la gestion de leurs soins en procurant un sentiment de contrôle et d'indépendance, pour réduire les obstacles à la communication avec l'équipe de soins, en améliorant leur sécurité et en les aidant à être proactifs dans la prise de décisions.

Chez les professionnels de santé : Pour leur part, les professionnels perçoivent que le DSPé pourrait être utile et avoir un effet sur leurs décisions en matière de diagnostic et de traitement, pour vérifier ou corriger les erreurs dans leurs systèmes (DME), dans les urgences ou pour les patients en région et qu'il pourrait contribuer à l'autonomisation des patients et améliorer la communication avec eux pour de meilleurs résultats cliniques et de santé.

5.4 Adoption et utilisation du DSPé

Récemment, on note qu'un nombre important d'études, aux résultats très variés, se sont intéressées aux facteurs d'adoption du DSPé auprès des patients [50, 51, 59, 60, 62, 64, 67, 73, 76, 79, 83, 85, 87-95] et des professionnels de la santé [60, 61, 66, 78, 86, 93, 96].

Chez les patients : On note un persistant paradoxe entre l'intérêt déclaré des patients quant au DSPé et l'anticipation des bénéfices pour la prestation des soins et services avec un taux d'adoption relativement faible, sauf pour certaines organisations disposant de systèmes intégrés (ex. *My HealthVet* du Veteran Health Administration et *MyHealthManager* du Kaiser

Permanente). Nous relevons que plusieurs modèles théoriques sont utilisés pour évaluer l'adoption du DSPé auprès des patients (Social Cognitive Theory (SCT) [62], Grounded Theory approach [65, 66, 73], Technology Acceptance Model [65], Diffusion of Innovation model [64], RE-AIM evaluation framework [51], Unified Theory of Acceptance and Use of Technology [59], Personal Health Records Adoption Model [91], Personal Health Records Acceptance Model [62]). Ainsi, plusieurs catégories de déterminants pour les patients sont identifiés dans ces études : 1) *facteurs individuels* : caractéristiques sociodémographiques (éducation, âge, sexe ; revenu), état de la santé (contrôle de la maladie, maladies chroniques, polymédication, expérience d'utilisation des services de santé...), connaissances et compétences individuelles, compétences et expérience d'utilisation d'ordinateur/Internet, littératie en santé et littératie en e-santé, attitudes et attentes envers le DSPé; 2) *facteurs liés à l'environnement social et physique* (par exemple influence sociale, conditions facilitantes) ; 3) *facteurs liés à la technologie* : coût, facilité/complexité d'utilisation, fonctionnalités, utilité perçue, avantages relatifs ; 4) *facteurs liés aux professionnels et prestataires de soins* : qualité de la relation, confiance, accès, communication, soutien, endossement des prestataires du DSPé ; 5) *facteurs organisationnels* : soutien organisationnel, offre d'éducation et formation.

Pour les professionnels : En comparaison avec les patients, peu d'études se sont intéressées à l'adoption/utilisation du DSPé auprès des professionnels de la santé. Ces études ont identifié plusieurs catégories de facteurs : 1) *caractéristiques des médecins et des organisations*: spécialité, milieu de travail (rural), expérience avec le dossier de santé électronique en général, en pratique de groupe, milieu à forte charge de travail; 2) *caractéristiques des patients* : plus d'un problème médical, sévérité de la maladie ; et 3) *attitudes et perceptions des avantages potentiels* (par exemple : améliorer les relations avec les patients, améliorer la qualité des soins, économie du temps, améliorer la précision des données).

5.5 Questions liées à la mise en œuvre

Notre revue note que trois principales questions liées directement à l'utilisation du DSPé sont soulevées par plusieurs des études retenues. Ces questions pourraient avoir un impact particulier sur l'implantation de ces systèmes au Canada.

La confidentialité et la sécurité : Quatre études se sont intéressées à cet aspect de l'implantation du DSPé [61, 64, 74, 76]. Cet aspect est une préoccupation et un déterminant de l'adoption aussi bien chez les patients que les professionnels. Cependant, cet aspect entraîne moins d'inquiétudes chez les patients atteints de maladies chroniques ou d'invalidité. Généralement, cette préoccupation est perçue seulement sous l'angle technique, qui n'est pas souvent suffisant. D'autres auteurs suggèrent de s'y attaquer surtout et avant sous un angle juridique [66, 97].

La littératie en santé : Cet aspect de l'implantation a été traité dans quatre études [65, 75, 91, 98]. En effet, il est décrit que l'utilisation du DSPé pouvait améliorer la littératie en santé des patients, mais aussi, que la littératie en santé, surtout la littératie en e-santé (traduction de l'anglais de la « e-health literacy »), va influencer la décision d'adopter ou d'utiliser le DSPé.

La fracture numérique : Il semble très probable que les personnes défavorisées sur le plan éducatif, économique ou social seront moins susceptibles d'utiliser les DSPé, et que tous les avantages liés à l'utilisation soient inégalement répartis entre les patients (iniquité de santé) [64, 75, 77].

6 Discussion

L'objectif de cette revue de la littérature était de mettre à jour certains des aspects en lien avec l'implantation du DSPé. Dans ce travail, nous nous sommes inspirés de la catégorisation

réalisée par Archer et collaborateurs [47]. Cependant, notre revue a traité d'aspects sélectionnés et ciblés, ne considérant pas toutes les dimensions retenus par ces auteurs qui traitaient de l'ensemble des aspects des dossiers de santé personnels (DSP) pour les patients au-delà des préoccupations des décideurs impliqués dans la mise en œuvre de tels outils. En effet, notre mise à jour de la littérature a été réalisée dans le but d'orienter la prise de décisions en matière d'implantation du DSPÉ au Canada et de guider la façon dont ces systèmes peuvent être implantés pour de plus grand bénéfices pour le citoyen et le système de santé (amélioration de l'efficacité, la qualité et les coûts). Nous nous sommes concentrés principalement sur certains aspects jugés utiles pour répondre à nos objectifs, en mettant l'accent sur les expériences menées dans des pays similaires au Canada.

Notre travail comporte donc certaines différences par rapport à la revue originale de Archer et collaborateurs [47]. D'abord, nous nous sommes limités à la littérature traitant de la version électronique des DSP (la version papier étant exclue). De plus, nous traitons des dossiers électroniques permettant l'accès à l'information et aux données de santé ainsi que la communication des patients avec les professionnels. Plusieurs études soutiennent que ces DSPÉ sont les plus liés aux résultats positifs auprès des patients et à l'amélioration de la prestation des soins. En favorisant une participation active des patients à la gestion de leurs soins, ces dossiers peuvent entraîner un changement dans la façon dont la place et le rôle du patient sont envisagés dans le système de santé et l'organisation des soins. Particulièrement, ces dossiers sont responsables d'effets et d'impacts sur la pratique professionnelle, les processus, l'organisation des soins et le système de soins en général. De tels effets et impacts sont à l'origine de changements et de préoccupations qu'il est utile de prendre en considération lors de la prise de décision en lien avec leur implantation.

Concernant les effets de l'utilisation du DSPÉ liés à la santé les études réalisées se sont surtout intéressées à ces effets auprès de patients diabétiques. En effet, des études de cohortes et des essais cliniques récents montrent que le DSPÉ permet d'améliorer la prestation de soins préventifs auprès des diabétiques et des enfants, ainsi que l'amélioration de la qualité des soins avec des résultats cliniques positifs, essentiellement pour les diabétiques. Toutefois, seulement l'utilisation de la messagerie sécurisée a été significativement associée au contrôle glycémique chez des adultes diabétiques utilisant un DSPÉ [50] et pour la détection précoces des effets liés aux médicaments jugés cliniquement importants ou graves [58]. Cependant les résultats notés demeurent inspirants pour l'utilisation auprès de patients souffrant d'autres maladies chroniques et suggèrent la capacité des systèmes de DSPÉ à répondre aux besoins de ces personnes. Nous avons pu noter aussi que l'implantation du DSPÉ dans sept grandes organisations de prestation de soins aux États-Unis supporte une approche organisationnelle, celle de soins centrés sur le patient, devenue une dimension importante de la qualité et de la sécurité des soins au XXI^e siècle [99-102]. Par contre, notre revue a montré des résultats divergents quant au rôle du DSPÉ dans la rationalisation de l'utilisation des services de santé ou pour l'amélioration de la qualité des soins (sauf pour les patients diabétiques).

Par ailleurs, plusieurs auteurs soutiennent que les avantages liés à l'utilisation du DSPÉ sont étroitement liés aux fonctionnalités offertes par ces systèmes et à l'utilité perçue aussi bien par les patients que les cliniciens. Ainsi, plus sont complètes les fonctionnalités offertes par un DSPÉ, plus cet outil sera utile et utilisé par les utilisateurs. En effet, les patients affichent une forte intention d'utiliser cette technologie s'ils croient que l'utilisation permettrait d'améliorer leur capacité et leur efficacité dans la gestion de leur état de santé et de leurs maladies. D'autre part, les études suggèrent que l'utilité perçue du DSPÉ est un principal prédicteur de la réussite de sa mise en œuvre. Aussi, plusieurs facteurs d'ordre individuel sont liés à l'adoption et l'utilisation du DSPÉ. En plus des perceptions et des attitudes, les niveaux d'adoption et d'utilisation étaient liés aux caractéristiques sociodémographiques des utilisateurs, à leur état de santé, à leur confiance et expérience avec internet et avec l'utilisation de la technologie, à leurs expériences avec le système

de soins, et leurs relations avec leurs équipes de soins.

D'autre part, le DSPé pourrait être source de changement de pratiques et de processus cliniques, suggérant l'importance de se préoccuper de l'alignement de ces systèmes avec le processus de soins et le travail des cliniciens lors de l'implantation. Le manque d'alignement avec les processus pourrait constituer une barrière à l'endossement, l'adoption et l'utilisation du DSPé aussi bien par les patients que par les professionnels.

Notre analyse de la littérature suggère aussi que l'implantation du DSPé doit être abordée sous l'angle de l'interaction entre la *technologie*, les *acteurs* et les *contextes* particuliers pour s'assurer de toutes les conditions de réussite de l'implantation de tels outils dans le réseau de la santé. En effet, plusieurs expériences internationales ont démontré l'échec d'approche d'implantation se préoccupant seulement d'un aspect ou d'un niveau d'analyse sans se soucier de l'autre (p.ex. HealthSpace en Angleterre [97, 103]).

Cependant, on recense peu d'études ayant exploré la perspective des prestataires de soins de santé et aucune étude ne s'est attardée aux aspects financiers (financement et coûts), juridiques ni aux différents facteurs organisationnels liés à l'implantation et l'utilisation de ces outils. Ces dimensions constitueront d'importantes suggestions pour des pistes de recherches futures contribuant à la prise de décision en permettant de mieux cerner de principales questions d'implantation de ces systèmes au Canada.

7 Conclusion

Les résultats des études empiriques consultées concernant plusieurs questions et préoccupations autour de l'implantation du DSPé demeurent encore divergents. Ainsi, les preuves scientifiques utiles pour soutenir une implantation efficace du DSPé au Canada, sont, généralement, encore insuffisantes. Cependant, on note que les preuves scientifiques sont de plus en plus évidentes et mises en exergue pour certains aspects utiles, suggérant des bénéfices pour les patients, les professionnels, les organisation et le système de santé, particulièrement pour le suivi et la gestion de maladie et de l'état de santé de clientèles particulières (suivi des enfants), ou de catégories de patients considérées comme grandes consommatrices des ressources de soins dans le système de santé (ex. maladies chroniques).

Remerciements

Nous tenons à remercier toutes les personnes et les institutions qui ont contribué de près ou de loin à la réalisation de ce travail. Particulièrement, nous tenons à remercier Mr Erik Breton et Mme Julie Payne-Gagnon, de l'Axe santé publique et pratiques optimales en santé du Centre de Recherche du CHUQ, pour leurs commentaires. Nous tenons aussi à remercier le Programme stratégique de formation en recherche transdisciplinaire sur les interventions de santé publique : promotion, prévention et politiques publiques (4P) du Réseau de recherche en santé des populations du Québec (RRSPQ) pour sa contribution au financement du projet de doctorat du premier auteur ainsi que et les Instituts de recherche en santé du Canada (IRSC) pour le soutien financier au projet.

Références

1. Zelmer J: **Health information solutions for consumers: research and plans.** *Health Manage Forum* 2011, **24**(3):144-146; discussion 147-148.
2. Le commissaire à la santé et au bien-être: **Adopter une approche intégrée de prévention et de gestion des maladies chroniques : recommandations, enjeux et implications.** In *Rapport d'appréciation de la performance du système de santé et de services sociaux*. Québec; 2010.
3. Levesque J-F, Feldman D, Dufresne C, BERGERON P, PINARD B, GAGNE V: *Barrières et éléments facilitant l'implantation de modèles intégrés de prévention et de gestion des maladies chroniques. Volume 40.* Paris, FRANCE: Editions de la Société; française de santé; publique; 2009.
4. Department of Health: **The expert patient: a new approach to chronic disease management for the 21st century.** London: Department of Health; 2001.
5. Holman H, Lorig K: **Patients as partners in managing chronic disease - Partnership is a prerequisite for effective and efficient health care.** *Brit Med J* 2000, **320**(7234):526-527.
6. Coulter A: **Paternalism or partnership?** *Brit Med J* 1999, **319**(7212):719-720.
7. Bourbeau J, Collet JP, Schwartzman K, Ducruet T, Nault D, Bradley C: **Economic benefits of self-management education in COPD.** *Chest* 2006, **130**(6):1704-1711.
8. Holman H, Lorig K: **Patient self-management: a key to effectiveness and efficiency in care of chronic disease.** *Public Health Rep* 2004, **119**(3):239-243.
9. Bury M, Pink D: **The HSJ debate. Self-management of chronic disease doesn't work.** *The Health service journal* 2005, **115**(5947):18-19, 11.
10. Ekman I, Wolf A, Olsson LE, Taft C, Dudas K, Schaufelberger M, Swedberg K: **Effects of person-centred care in patients with chronic heart failure: the PCC-HF study.** *Eur Heart J* 2012, **33**(9):1112-1119.
11. Longtin Y, Sax H, Leape LL, Sheridan SE, Donaldson L, Pittet D: **Patient Participation: Current Knowledge and Applicability to Patient Safety.** *Mayo Clin Proc* 2010, **85**(1):53-62.
12. Kitson A, Marshall A, Bassett K, Zeitz K: **What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing.** *Journal of advanced nursing* 2012.
13. Commission européenne: **Participation des patients. Rapport complet.** Commission européenne-Direction générale de la Santé et des Consommateurs; 2012.
14. Ministère de la Santé et des Services Sociaux: **Cadre de référence pour la prévention et la gestion des maladies chroniques physiques en première ligne.** La Direction des communications du ministère de la Santé et des Services sociaux edition. Québec: MSSS; 2012:33.
15. **HIMSS Personal Health Records Definition and Position Statement** [<http://himss.files.cms-plus.com/HIMSSorg/Content/files/PHRDefinition071707.pdf>]
16. Falcão-Reis F, Correia ME: **Patient empowerment by the means of citizen-managed electronic health records: Web 2.0 health digital identity scenarios.** *Studies in Health Technology & Informatics* 2010, **156**:214-228.
17. Dorr D, Bonner LM, Cohen AN, Shoai RS, Perrin R, Chaney E, Young AS: **Informatics systems to promote improved care for chronic illness: a literature review.** *Journal of the American Medical Informatics Association : JAMIA* 2007, **14**(2):156-163.
18. Bitterman N, Lerner E, Bitterman H: **Evaluation of Data Display for Patient-Oriented Electronic Record of Anticoagulant Therapy.** *Telemedicine Journal and E-Health* 2010, **16**(7):799-806.
19. Brennan PF, Downs S, Casper G: **Project HealthDesign: rethinking the power and potential of personal health records.** *Journal of Biomedical Informatics* 2010, **43**(5 Suppl):S3-5.
20. Leonard KJ, Dalziel S: **Quand et comment la cybersanté est un bon investissement pour les patients qui prennent en charge une maladie chronique.** *Healthcare management forum / Canadian College of Health Service Executives = Forum gestion des soins de santé / Collège canadien des directeurs de services de santé* 2011, **24**(3):129-136.
21. Markle-Foundation: **Connecting for Health: A Public-Private Collaborative (July 2003).** . Volume 2011. Markle Foundation; 2003.
22. Blumenthal D, Glaser JP: **Information Technology Comes to Medicine.** *N Engl J Med* 2007,

23. Markle Foundation: **Connecting for health : The Personal Health Working Group final report.**; 2003.
24. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW: **A research agenda for personal health records (PHRs).** *Journal of the American Medical Informatics Association* 2008, **15**(6):729-736.
25. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ: **Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption.** *Journal of the American Medical Informatics Association* 2006, **13**(2):121-126.
26. Ahmadi M, Jeddi FR, Gohari MR, Sadoughi F: **A review of the personal health records in selected countries and Iran.** *J Med Syst* 2012, **36**(2):371-382.
27. Ertmer A, Uckert F: **User Acceptance of and Satisfaction with a Personal Electronic Health Record.** *Connecting Medical Informatics and Bio-Informatics* 2005, **116**:861-866.
28. Honeyman A, Cox B, Fisher B: **Potential impacts of patient access to their electronic care records.** *Inform Prim Care* 2005, **13**(1):55-60.
29. Mays D, Baker D, Butros A, Cowles KE: **Giving patients access to their medical records via the internet: the PCASSO experience.** *Journal of the American Medical Informatics Association : JAMIA* 2002, **9**(2):181-191.
30. Pyper C, Amery J, Watson M, Crook C: **Patients' experiences when accessing their on-line electronic patient records in primary care.** *British Journal of General Practice* 2004, **54**(498):38-43.
31. Ward L, Innes M: **Electronic medical summaries in general practice--considering the patient's contribution.** *Br J Gen Pract* 2003, **53**(489):293-297.
32. Winkelman WJ, Leonard KJ, Rossos PG: **Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness.** *Journal of the American Medical Informatics Association* 2005, **12**(3):306-314.
33. Nazi KM, Hogan TP, Wagner TH, McInnes DK, Smith BM, Haggstrom D, Chumbler NR, Gifford AL, Charters KG, Saleem JJ *et al*: **Embracing a Health Services Research Perspective on Personal Health Records: Lessons Learned from the VA My HealthVet System.** *Journal of General Internal Medicine* 2010, **25**:62-67.
34. Kahn JS, Hilton JF, Van Nunnery T, Leasure S, Bryant KM, Hare CB, Thom DH: **Personal health records in a public hospital: experience at the HIV/AIDS clinic at San Francisco General Hospital.** *Journal of the American Medical Informatics Association* 2010, **17**(2):224-228.
35. Kimmel Z, Greenes RA, Liederman E: **Personal health records.** *The Journal of medical practice management : MPM* 2005, **21**(3):147-152.
36. Ball MJ, Costin MY, Lehmann C: **The Personal Health Record: Consumers Banking on their Health.** *St Heal T* 2008, **134**:35-46.
37. Huba N, Zhang Y: **Designing Patient-Centered Personal Health Records (PHRs): Health Care Professionals' Perspective on Patient-Generated Data.** *J Med Syst* 2012, **36**(6):3893-3905.
38. TNS Qual+: **Participation des patients. Rapport complet.** la Commission européenne-Direction générale de la Santé et des Consommateurs; 2012.
39. Krist AH, Peele E, Woolf SH, Rothemich SF, Loomis JF, Longo DR, Kuzel AJ: **Designing a patient-centered personal health record to promote preventive care.** *BMC Med Inform Decis Mak* 2011, **11**(1).
40. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW: **A research agenda for personal health records (PHRs).** *Journal of the American Medical Informatics Association : JAMIA* 2008, **15**(6):729-736.
41. **Littératie en santé au Canada : Résultats initiaux de l'Enquête internationale sur l'alphabétisation et les compétences des adultes** [<http://www.ccl-cca.ca/pdfs/HealthLiteracy/LitteratieensanteauCanada.pdf>]
42. Department of Health: **The expert patient: a new approach to chronic disease management for the 21st century.** London: National Health System- NHS; 2001:38.
43. Assadi V, Hassanein K: **Consumer's acceptance and use of personal health record systems: a Theoretical Model** Edited by Systems tECoI. Verona. Italy; 2009:13.
44. Gerard MN, Cohen M, Greer-Smith R: **Personal touch: personal health records for consumers of healthcare.** *Journal of healthcare information management : JHIM* 2009, **23**(3):26-30.
45. Humphry JW, Thomas S, Mok M, Rubin H: **The electronic personal health record (EPHR): A new paradigm for ownership and use of health records.** *Journal of General Internal Medicine* 2005, **20**:50-

46. Raghupathi W, Tan J: **Strategic IT applications in health care.** *Commun ACM* 2002, **45**(12):56-61.
47. Archer N, Fevrier-Thomas U, Lokker C, McKibbin KA, Straus SE: **Personal health records: a scoping review.** *Journal Of The American Medical Informatics Association: JAMIA* 2011, **18**(4):515-522.
48. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, Hui SL, Hyde SM, Klieman LH, Mitchell CJ *et al*: **Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial.** *Journal of the American Medical Informatics Association : JAMIA* 2013, **20**(3):526-534.
49. Wagner PJ, Dias J, Howard S, Kintziger KW, Hudson MF, Seol YH, Sodomka P: **Personal health records and hypertension control: A randomized trial.** *Journal of the American Medical Informatics Association* 2012, **19**(4):626-634.
50. Wade-Vuturo AE, Mayberry LS, Osborn CY: **Secure messaging and diabetes management: experiences and perspectives of patient portal users.** *Journal of the American Medical Informatics Association : JAMIA* 2013, **20**(3):519-525.
51. Gordon P, Camhi E, Hesse R, Odlum M, Schnall R, Rodriguez M, Valdez E, Bakken S: **Processes and outcomes of developing a continuity of care document for use as a personal health record by people living with HIV/AIDS in New York City.** *Int J Med Inform* 2012, **81**(10):e63-e73.
52. Shaw RJ, Ferranti J: **Patient-provider internet portals - Patient outcomes and use.** *CIN - Computers Informatics Nursing* 2011, **29**(12):714-718.
53. Tenforde M, Nowacki A, Jain A, Hickner J: **The Association Between Personal Health Record Use and Diabetes Quality Measures.** *Journal of General Internal Medicine* 2012, **27**(4):420-424.
54. Krist AH, Woolf SH, Rothemich SF, Johnson RE, Peele JE, Cunningham TD, Longo DR, Bello GA, Matzke GR: **Interactive Preventive Health Record to Enhance Delivery of Recommended Care: A Randomized Trial.** *The Annals of Family Medicine* 2012, **10**(4):312-319.
55. Nagykaldis Z, Aspy CB, Chou A, Mold JW: **Impact of a Wellness Portal on the delivery of patient-centered preventive care.** *J Am Board Fam Med* 2012, **25**(2):158-167.
56. Otsuka SH, Tayal NH, Porter K, Embi PJ, Beatty SJ: **Improving herpes zoster vaccination rates through use of a clinical pharmacist and a personal health record.** *American Journal of Medicine* 2013, **126**(9):832.e831-832.e836.
57. Weingart SN, Carbo A, Tess A, Chiappetta L, Tutkus S, Morway L, Toth M, Davis RB, Phillips RS, Bates DW: **Using a patient internet portal to prevent adverse drug events: a randomized, controlled trial.** *Journal of patient safety* 2013, **9**(3):169-175.
58. Heyworth L, Paquin AM, Clark J, Kamenker V, Stewart M, Martin T, Simon SR: **Engaging patients in medication reconciliation via a patient portal following hospital discharge.** *J Am Med Inform Assoc* 2013.
59. Or CKL, Karsh BT, Severtson DJ, Burke LJ, Brown RL, Brennan PF: **Factors affecting home care patients' acceptance of a web-based interactive self-management technology.** *Journal of the American Medical Informatics Association* 2011, **18**(1):51-59.
60. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, Picton P, Seto E, Cafazzo J: **Improving diabetes management with a patient portal: Qualitative study of a diabetes self-management portal.** *J Med Internet Res* 2012, **14**(6).
61. Wynia MK, Torres GW, Lemieux J: **Many Physicians Are Willing To Use Patients' Electronic Personal Health Records, But Doctors Differ By Location, Gender, And Practice.** *Health Affairs* 2011, **30**(2):266-273.
62. Agarwal R, Anderson C, Zarate J, Ward C: **If We Offer it, Will They Accept? Factors Affecting Patient Use Intentions of Personal Health Records and Secure Messaging.** *J Med Internet Res* 2013, **15**(2).
63. Morton AA: **Examining Acceptance of an Integrated Personal Health Record (PHR).** University of Maryland, Baltimore; 2011.
64. Emani S, Yamin CK, Peters E, Karson AS, Lipsitz SR, Wald JS, Williams DH, Bates DW: **Patient perceptions of a personal health record: A test of the diffusion of innovation model.** *J Med Internet Res* 2012, **14**(6).
65. Day K, Gu Y: **Influencing factors for adopting personal health record (PHR).** *Studies in health technology and informatics* 2012, **178**:39-44.
66. Nazi KM: **The personal health record paradox: health care professionals' perspectives and the**

- information ecology of personal health record systems in organizational and clinical settings. *J Med Internet Res* 2013, **15**(4):e70-e70.**
67. Tom JO, Mangione-Smith R, Solomon C, Grossman DC: **Integrated personal health record use: association with parent-reported care experiences.** *Pediatrics* 2012, **130**(1):e183-190.
 68. Palen TE, Ross C, Powers JD, Stanley X: **Association of online patient access to clinicians and medical records with use of clinical services.** *JAMA - Journal of the American Medical Association* 2012, **308**(19):2012-2019.
 69. Wagner EH, Grothaus LC, Sandhu N, Galvin MS, McGregor M, Artz K, Coleman EA: **Chronic care clinics for diabetes in primary care: a system-wide randomized trial.** *Diabetes care* 2001, **24**(4):695-700.
 70. Reti SR, Feldman HJ, Ross SE, Safran C: **Improving personal health records for patient-centered care.** *Journal of the American Medical Informatics Association* 2010, **17**(2):192-195.
 71. Haggstrom DA, Saleem JJ, Russ AL, Jones J, Russell SA, Chumbler NR: **Lessons learned from usability testing of the VA's personal health record.** *Journal of the American Medical Informatics Association* 2011, **18**(SUPPL. 1):13-17.
 72. Segall N, Saville JG, L'Engle P, Carlson B, Wright MC, Schulman K, Tcheng JE: **Usability evaluation of a personal health record.** *AMIA Annu Symp Proc* 2011, **2011**:1233-1242.
 73. Kerns JW, Krist AH, Longo DR, Kuzel AJ, Woolf SH: **How patients want to engage with their personal health record: a qualitative study.** *BMJ Open* 2013, **3**(7).
 74. Luque AE, van Keken A, Winters P, Keefer MC, Sanders M, Fiscella K: **Barriers and Facilitators of Online Patient Portals to Personal Health Records Among Persons Living With HIV: Formative Research.** *J Med Internet Res* 2013, **15**(1).
 75. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, Schillinger D: **Social disparities in internet patient portal use in diabetes: Evidence that the digital divide extends beyond access.** *Journal of the American Medical Informatics Association* 2011, **18**(3):318-321.
 76. Lafky DB, Horan TA: **Personal health records: Consumer attitudes toward privacy and security of their personal health information.** *Health Inform J* 2011, **17**(1):63-71.
 77. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, Bates DW: **The digital divide in adoption and use of a personal health record.** *Arch Intern Med* 2011, **171**(6):568-574.
 78. Yau GL, Williams AS, Brown JB: **Family physicians' perspectives on personal health records: Qualitative study.** *Canadian family physician Medecin de famille canadien* 2011, **57**(5):e178-184.
 79. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW: **Patient reported barriers to enrolling in a patient portal.** *Journal of the American Medical Informatics Association* 2011, **18**:i8-i12.
 80. Patel VN, Abramson E, Edwards AM, Cheung MA, Dhopeswarkar RV, Kaushal R: **Consumer attitudes toward personal health records in a Beacon community.** *American Journal of Managed Care* 2011, **17**(4):e104-e120.
 81. Wakefield DS, Kruse RL, Wakefield BJ, Koopman RJ, Keplinger LE, Canfield SM, Mehr DR: **Consistency of patient preferences about a secure internet-based patient communications portal: contemplating, enrolling, and using.** *American journal of medical quality : the official journal of the American College of Medical Quality* 2012, **27**(6):494-502.
 82. Zulman DM, Nazi KM, Turvey CL, Wagner TH, Woods SS, An LC: **Patient interest in sharing personal health record information a web-based survey.** *Annals of Internal Medicine* 2011, **155**(12):805-810.
 83. Britto MT, Hesse EA, Kamdar OJ, Munafo JK: **Parents' perceptions of a patient portal for managing their child's chronic illness.** *The Journal of pediatrics* 2013, **163**(1):280-281 e281-282.
 84. Byczkowski TL, Munafo JK, Britto MT: **Variation in use of internet-based patient portals by parents of children with chronic disease.** *Archives of Pediatrics and Adolescent Medicine* 2011, **165**(5):405-411.
 85. Patel VN, Dhopeswarkar RV, Edwards A, Barron Y, Sparenborg J, Kaushal R: **Consumer Support for Health Information Exchange and Personal Health Records: A Regional Health Information Organization Survey.** *J Med Syst* 2012, **36**(3):1043-1052.
 86. Keplinger LE, Koopman RJ, Mehr DR, Kruse RL, Wakefield BJ, Canfield SM: **Patient portal implementation: resident and attending physician attitudes.** *Fam Med* 2013, **45**(5):335-340.
 87. Ancker JS, Barron Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, Calman N: **Use of an electronic patient portal among disadvantaged populations.** *J Gen Intern Med* 2011, **26**(10):1117-1123.

88. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW: **Disparities in enrollment and use of an electronic patient portal.** *Journal of General Internal Medicine* 2011, **26**(10):1112-1116.
89. Ketterer T, West DW, Sanders VP, Hossain J, Kondo MC, Sharif I: **Correlates of patient portal enrollment and activation in primary care pediatrics.** *Academic pediatrics* 2013, **13**(3):264-271.
90. Logue MD: **Validating and testing a model to predict adoption of electronic personal health record systems in the self-management of chronic illness in the older adult.** University of Arizona; 2011.
91. Logue MD, Effken JA: **An exploratory study of the personal health records adoption model in the older adult with chronic illness.** *Inform Prim Care* 2012, **20**(3):151-169.
92. Lyles CR, Sarkar U, Ralston JD, Adler N, Schillinger D, Moffet HH, Huang ES, Karter AJ: **Patient-provider communication and trust in relation to use of an online patient portal among diabetes patients: The Diabetes and Aging Study.** *Journal of the American Medical Informatics Association : JAMIA* 2013.
93. Menon AS, Greenwald S, Ma TJ, Kooshesh S, Duriseti R: **Patient and physician willingness to use personal health records in the emergency department.** *Western Journal of Emergency Medicine* 2012, **13**(2):172-175.
94. Sanders MR, Winters P, Fortuna RJ, Mendoza M, Berliant M, Clark L, Fiscella K: **Internet access and patient portal readiness among patients in a group of inner-city safety-net practices.** *The Journal of ambulatory care management* 2013, **36**(3):251-259.
95. Ronda MCM, Dijkhorst-Oei L-T, Gorter KJ, Beulens JWJ, Rutten GEHM: **Differences between diabetes patients who are interested or not in the use of a patient web portal.** *Diabetes Technol Ther* 2013, **15**(7):556-563.
96. Huba N, Zhang Y: **Designing Patient-Centered Personal Health Records (PHRs): Health Care Professionals' Perspective on Patient-Generated Data.** *J Med Syst* 2012, **36**(6):3893-3905.
97. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J: **Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace.** *BMJ (Clinical Research Ed)* 2010, **341**:c5814-c5814.
98. Noblin AM, Wan TTH, Fottler M: **The impact of health literacy on a patient's decision to adopt a personal health record.** *Perspect Health Inf Manag* 2012, **9**:1-13.
99. Stewart M: **Towards a global definition of patient centred care - The patient should be the judge of patient centred care.** *Brit Med J* 2001, **322**(7284):444-445.
100. Kohn L, Corrigan J, Donaldson M: **Crossing the quality chasm: A new health system for the 21st century.** Washington, DC: Committee on Quality of Health Care in America, Institute of Medicine 2001.
101. Institute of Medicine: **Crossing the Quality Chasm: A New Health System for the Twenty-First Century.** Washington: Institute of Medicine.; 2001.
102. Epping-Jordan JE, Pruitt SD, Bengoa R, Wagner EH: **Improving the quality of health care for chronic conditions.** *Qual Saf Health Care* 2004, **13**(4):299-305.
103. **The devil's in the detail. Final report of the independent evaluation of the Summary Care Record and HealthSpace programmes** [<http://www.ucl.ac.uk/news/scriefullreport.pdf>]

Adresse de correspondance

Centre de recherche du CHU de Québec
 10, rue de l'Espinau, D6-731
 Québec (QC), Canada
 G1L 3L5
el-kebir.ghandour.1@ulaval.ca

Annexes

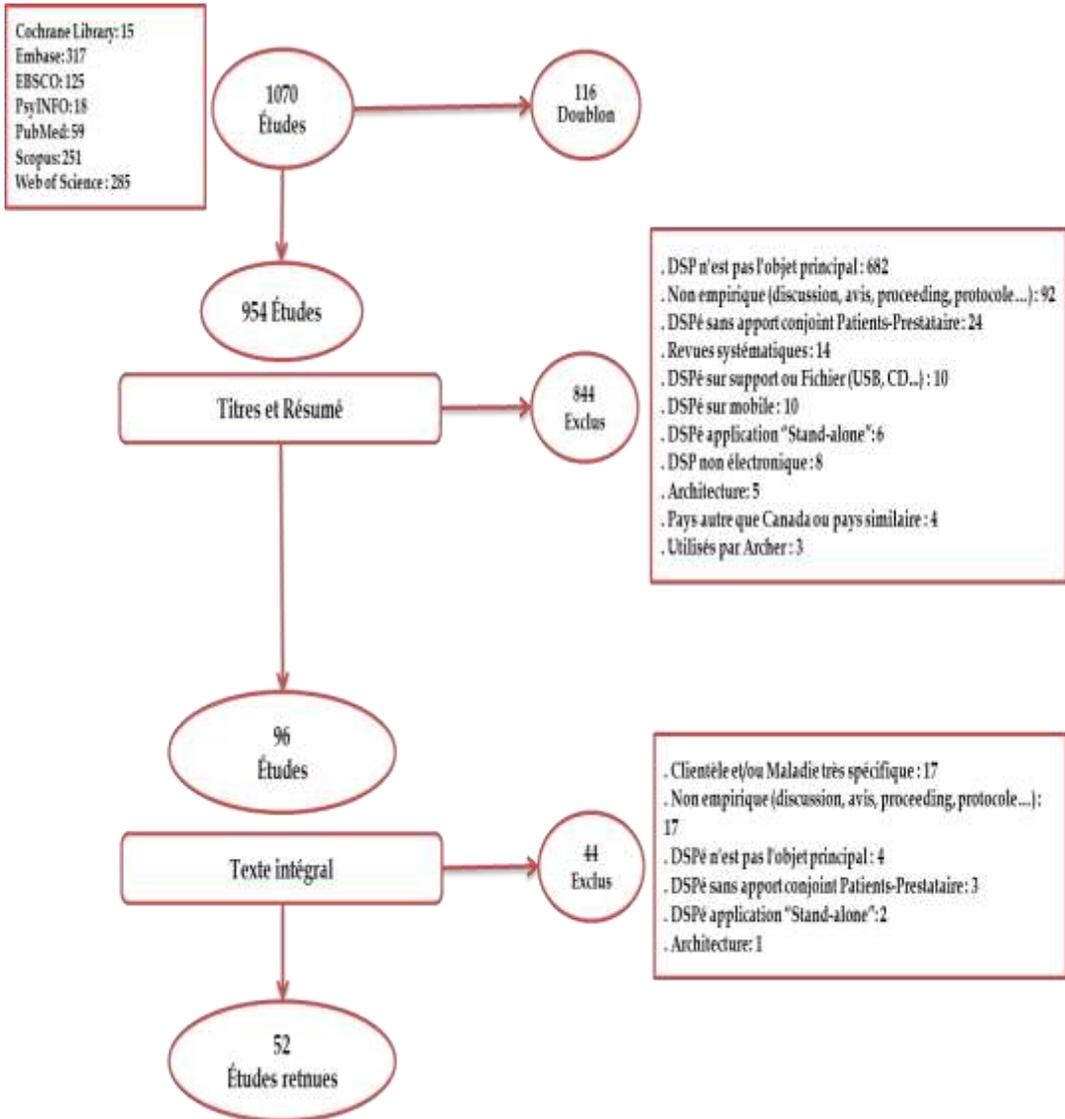


Figure 1 : Diagramme de sélection des études incluses

Annexes

Tableau 1 : Répartition des études selon le pays, les participants le devis et la méthodologie

Répartition des études selon le pays					
Pays	États-Unis		Canada		Autres Pays
Nombre (%)	47 (90,4%)		2 (3,8%)		3 (5,8%)
Répartition des études selon les participants					
Participants	Patients	Médecins/ autres professionnels	Parents	Patients/ professionnels	Experts
Nombre (%)	37 (71,2%)	6 (11,5%)	4 (7,7%)	3 (5,8%)	2 (3,8%)
Répartition des études selon le devis utilisé					
Devis	Sondage	Essai clinique randomisé	Étude de cohorte	Étude de cas	
Nombre (%)	32 (61,5%)	7 (13,5%)	4 (7,7%)	2 (3,8%)	
Répartition des études selon la méthodologie					
	Quantitative (+ECR)	Mixte		Qualitative	
Nombre (%)	35 (67,3%)	10 (19,2%)		7 (13,5%)	

Tableau 2 : Catégorisation des études retenues

QUESTION	NOMBRE D'ÉTUDES	DESIGN D'ÉTUDES ^{Nombre}	REFERENCES
<i>Résultats cliniques</i>	6	ECR ²	(Tang et al., 2013; Wagner et al., 2012)
		Sondage ²	(Shaw et al., 2011; Wade-Vuturo et al., 2013)
		Étude de cohorte ²	(Gordon et al., 2012; Tenforde et al., 2012)
<i>Prestation de services préventifs</i>	4	ECR ³	(Krist et al., 2012; Nagykaldi et al., 2012; Otsuka et al., 2013)
		Étude de cohorte ¹	(Tenforde et al., 2012)
Amélioration de la sécurité des patients	2	ECR ¹	(Weingart et al., 2013)
		Mixte ¹	(Heyworth et al., 2013)
Autonomisation des Patients	4	ECR ¹	(Wagner et al., 2012)
		Étude de cohorte ¹	(Gordon et al., 2012)
		Mixte ¹	(Urowitz et al., 2012)
		Étude de cohorte ¹	(Or et al., 2011)
Soutien à l'autogestion	2	Sondage ¹	(Morton, 2011)
		Mixte ¹	(Urowitz et al., 2012)
Amélioration de la communication patient–professionnels	6	Sondage ²	(Emani, et al., 2012; Wynia, et al., 2011)
		Qualitative ⁴	(Day, et al., 2012; Huba, et al., 2012; Kim M. Nazi, 2013; Urowitz, et al., 2012)
Utilisation et expérience des soins et services	6	Étude de cohorte ²	(Gordon et al., 2012; Palen et al., 2012)
		Sondage ¹	(Tom et al., 2012)
		ECR ³	(Nagykaldi et al., 2012; Wagner et al., 2012; Weingart et al., 2013)
Centralité des soins	2	Qualitative ¹	(Reti et al., 2010)
		ECR ¹	(Nagykaldi et al., 2012)
Effets liés à la qualité des soins	2	Étude de cohorte ¹	(Tenforde et al., 2012)
		Sondage ¹	(Gordon et al., 2012)

DSPÉ EFFETS/ BÉNÉFICES

Changement des processus	1	Qualitative ¹	(Nazi, 2013)
<i>Patients : Acceptation / Adoption et Utilisation</i>			
		Sondage ³	(Agarwal et al., 2013; Goel et al., 2011a; Morton, 2011)
Acceptation /Adoption	6	Observationnelle ¹	(Ancker et al., 2011; Or et al., 2011)
		Qualitative ¹	(Day et al., 2012)
		Étude de cas ¹	(Greenhalgh et al., 2010)
		Sondage ⁸	(Cho et al., 2010; Lyles et al., 2013; Menon et al., 2012; Nazi, 2010; Patel et al., 2012; Ronda et al., 2013; Seters et al., 2013; Sarkar et al., 2011)
Accès / Utilisation	15	Observationnelle ³	(Ancker et al., 2011; Goel et al., 2011b; Ketterer et al., 2013; Or et al., 2011)
		Qualitative ¹	(Kerns et al., 2013)
		Mixte ¹	(Lafky et al., 2011)
		Étude de Cohorte ²	(Gordon et al., 2012; Sarkar et al., 2011)
<i>Parents : Acceptation / Adoption et Utilisation</i>			
Acceptation /Adoption	1	Observationnelle ¹	(Ketterer et al., 2013)
Accès / Utilisation	2	Sondage ¹	(Tom et al., 2012)
		Observationnelle ¹	(Byczkowski et al., 2011)
<i>Médecins/ Professionnels : Adoption et Utilisation</i>			
		Qualitative ³	(Huba et al., 2012; Nazi, 2013; Yau et al., 2011)
	7	Sondage ⁴	(Keplinger et al., 2013; Menon et al., 2012; Urowitz et al., 2012; Wynia et al., 2011)
Sécurité et vie privée	3	Sondage ¹	(Emani et al., 2012; Wynia et al., 2011)
		Mixte ²	(Lafky et al., 2011; Luque et al., 2013)
Littérature en santé	1	Sondage ¹	(Noblin et al., 2012)
		Sondage ¹	(Sarkar et al., 2011)
Fracture numérique	2	Observationnelle ¹	(Yamin et al., 2011)

<i>Patients : Perceptions / Attitudes</i>	5	Sondage ⁵	(Emani et al., 2012; Goel et al., 2011a; Patel et al., 2011; Wakefield et al., 2012; Zulman et al., 2011)
<i>Médecins/ Professionnels : Perceptions / Attitudes</i>	5	Qualitative ² Sondage ³	(Huba et al., 2012; Yau et al., 2011) (Keplinger et al., 2013; Urowitz et al., 2012; Wynia et al., 2011)
<i>Parents : Perceptions / Attitudes</i>	1	Qualitative ¹	(Britto et al., 2013)

La satisfaction des utilisateurs des Systèmes d'Information Clinique du Groupe Hospitalier Paris Ouest (HUPO)

Clinical Information System User Satisfaction in the Paris-West University Hospital Group (HUPO)

Brahim Hadji^{1,2,3}, Isabelle Dupuis¹, Laurence Lenevent¹, Didier Heudes^{1,2,4}, Jean-François Wagner¹, Patrice Degoulet^{1,2,4}

¹Hôpital Européen George Pompidou (HEGP), Paris, France

²Inserm-UMRS 1138, CRC, Equipe 22, Paris, France

³Université Pierre et Marie Curie, Paris, France

⁴Université Paris Descartes, Paris, France

Résumé

Introduction- L'évaluation de la satisfaction des utilisateurs doit être effectuée régulièrement lors des différentes phases de déploiement d'un système d'information clinique (SIC). Cet article compare, en période de post-adoption tardive, la satisfaction des utilisateurs dans les trois hôpitaux du groupe hospitalier (GH) universitaire Paris-Ouest (HUPO), HEGP, Corentin-Celton (CCL) et Vaugirard (VGR). *Méthodes-* Deux enquêtes standardisées de satisfaction du SIC ont été réalisées en 2011 et 2013 dans les 3 sites du groupe. Le modèle utilisé, IT-PAM (Information Technology Post-Adoption Model), est bâti autour de 7 dimensions : caractéristiques des utilisateurs, qualité du SIC, utilité perçue, qualité du support utilisateurs, confirmation des attentes, satisfaction et intention de continuer. *Résultats-* L'analyse des résultats combinés des enquêtes 2011 et 2013 obtenus chez 519 utilisateurs montre une satisfaction plus grande à l'HEGP que dans les hôpitaux CCL et VGR ($p < 0,01$) associée à un plus grand déploiement et une plus grande utilisation du dossier patient électronique ($p < 0,001$). La confirmation des attentes est significativement plus faible chez les médecins et les personnels de sexe masculin que dans les autres catégories d'utilisateurs ($p < 0,001$). 74,7% des utilisateurs sont en faveur d'une homogénéisation complète du SIC sur les trois sites du GH. Dans les modèles de régression multiple, la qualité du SIC, l'utilité perçue et la confirmation des attentes sont les paramètres les plus liés à la satisfaction. *Conclusion-* La réalisation d'enquêtes périodique de satisfaction est indispensable pour améliorer l'utilisation hospitalière des SIC et leur amélioration par les fournisseurs de solutions informatiques.

Abstract

Introduction: Evaluation of user satisfaction should be regularly monitored during the various phases of deployment of a clinical information system (CIS). *Objective:* This paper compares in a

late post-adoption phase CIS satisfaction in the three hospitals of the Paris-West University Hospital group (HUPO): HEGP, Corentin-Celton (CCL), and Vaugirard (VGR). Methods: Two standardized satisfaction surveys were conducted in 2011 and 2013. The IT-PAM model used (Information Technology Post-Adoption Model) is built around 7 dimensions: user characteristics, CIS quality, user support, confirmation of expectations, satisfaction, and continuance intention. Results: The analysis of the combined surveys 2011-2013 obtained from 519 users shows greater user satisfaction at the HEGP compared to CCL-VGR ($p < 0.01$), associated with a greater deployment and use of the CIS functions ($p < 0.001$). Confirmation of expectations is lower among physicians and male users than among other categories ($p < 0.001$). 74.7% of users are in favor of a complete unification of the CIS in the three HUPO hospitals. In multi-regression models, satisfaction is positively and significantly correlated with CIS quality, perceived usefulness and confirmation of expectations. Conclusion: Repetition of satisfaction surveys is a necessary condition to improve CIS use within hospitals and their continuous improvement by solution providers.

Mots-clés : Acceptabilité des systèmes d'information cliniques – Modèle d'évaluation en post-adoption – Confirmation des attentes – Satisfaction – Intention de continuer.

Keywords: *Clinical Information System Acceptance – Post-adoption evaluation model – Confirmation of expectation – Satisfaction – Continuance Intention.*

1. Introduction

L'évaluation régulière de la satisfaction des utilisateurs est un élément essentiel de tout projet de mise en œuvre d'un système d'information clinique (SIC). Elle doit être effectuée régulièrement lors des phases de déploiement puis de consolidation d'un SIC, notamment au cours des changements majeurs de versions des logiciels. Elle doit faire partie d'une stratégie plus vaste d'évaluation permanente d'un système d'information [1–5]. Une confiance élevée dans un SIC et dans l'organisation logistique mise en place pour le déployer et le faire évoluer incitera les utilisateurs à utiliser au mieux ses fonctionnalités et donc à participer à son amélioration.

Plusieurs modèles d'évaluation de la satisfaction d'un SIC ont été proposés pour prédire leur utilisation, mesurer la satisfaction des utilisateurs ou prédire leur futur comportement (l'intention de continuer). Parmi les modèles fréquemment utilisés il faut citer le modèle ISC (Information Success Model) [6], le modèle d'acceptabilité TAM (Technology Acceptance Model) et ses extensions [4,7,8], le modèle ISCI (Information System Continuance Intention) [9], le modèle UTAUT (Unified Theory of Acceptance and Use of Technology) [10] et la théorie de confirmation des attentes ECT (Expectation Confirmation Theory) [9,11]. Le recueil des données sur chacune des dimensions peut être effectué par interviews individuels ou par groupes, par enquêtes (questionnaires), par des études ethnographiques (observations avec ou sans enregistrements vidéo) ou par la combinaison de ces techniques [12]. Les dimensions de l'évaluation retenues et l'importance donnée à chaque dimension doivent être ajustées par rapport à la phase du cycle de déploiement concernée. Lors des phases initiales de déploiement, il est important d'observer les utilisateurs au plus près de leur mode d'interaction avec un SIC afin d'évaluer la facilité d'utilisation et les modalités d'interaction avec le SIC (utilisabilité), les performances, mais également la qualité de la formation initiale et la réactivité du service

informatique (correction des anomalies, prise en compte des incidents matériels, etc.) et d'une façon plus générale la confirmation des attentes des utilisateurs [11]. Lors de l'utilisation en routine d'un SIC, la satisfaction peut dépendre de facteurs multiples comme les caractéristiques de l'utilisateur, la qualité du système et de son support (organisation, information, prise en compte des demandes d'évolution, rapidité de correction des anomalies, remplacement des matériels obsolètes, etc.). La répétition des évaluations au cours du temps est indispensable pour apprécier l'évolution des comportements des utilisateurs, l'évolution des fonctionnalités du SIC et la qualité des organisations mise en place. Dans ce travail, et à travers un outil d'enquête standardisée, nous essayons de mesurer et d'analyser les déterminants de la satisfaction et de l'intention de continuer à utiliser les différents SICs des trois hôpitaux du groupe Hospitalier Universitaire Paris Ouest (HUPO) sur la base de deux enquêtes réalisées en 2011 et 2013.

2. Matériels et méthodes

2.1. Les systèmes d'information clinique au sein du groupe HUPO

Les trois hôpitaux du groupe HUPO utilisent deux SICs différents, celui de l'HEGP reposant sur DxCare[®] /Medasys[®], et ceux de Corentin Celton (CCL) et Vaugirard (VGR) sur Actipidos[®]/GFI[®]. Le SIC de l'HEGP intègre, autour d'un dossier patient unique, la gestion des identités et mouvements, la planification des rendez-vous, la prescription connectée multimodale avec aide à la décision et intégration automatique des résultats et de la réalisation des actes dans le dossier électronique, un entrepôt clinique intégré basé sur i2b2 [13]. Le SIC des hôpitaux CCL et VGR fait appel à différents logiciels interfacés : Gilda[®] pour les identités, Actipidos[®] pour la pancarte et la prescription médicamenteuse et d'actes de soins, Agenda[®] pour les rendez-vous, Mediweb[®] pour le stockage de certains comptes rendus. Les résultats n'y sont pas directement insérés dans le dossier patient électronique mais accessibles à partir du serveur de résultats STARE[®].

2.2. Hypothèses de recherche

Les enquêtes ont été effectuées au travers de questionnaires semi-ouverts (50 questions fermées et deux questions ouvertes faisant appel à des commentaires libres). Les questions ont été sélectionnées à partir de construits précédemment publiés et validés dans la littérature et de deux études d'évaluation faites à l'HEGP en 2004 et 2008 [14,15]. Le nombre des questions a été volontairement limité pour permettre le remplissage d'un questionnaire en moins de 10 minutes. L'accès aux questionnaires est ouvert à tous les professionnels de santé du GH travaillant sur le SIC et en contact direct avec le dossier patient. Il est effectué au travers de l'Intranet du groupe, possibilité étant donnée aux utilisateurs de compléter leurs questionnaires en plusieurs étapes. La publicité est effectuée par affichage dans les services, courriers électroniques et rappels par emails. Les réponses ont été traitées de façon anonyme.

Les questionnaires d'évaluation sont basés sur le modèle IT-PAM (*Information Technology Post Adoption Model*), progressivement amélioré depuis sa première utilisation en 2004 à l'HEGP, lui-même considéré comme une extension des modèles ISC, TAM et ISCI [14,15].

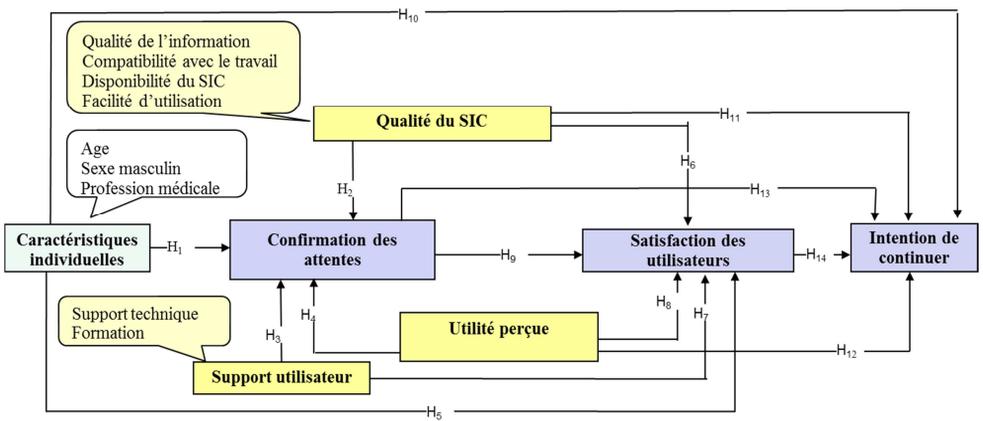


Figure 1 : Modèle IT-PAM 2 d'évaluation de la satisfaction en période de post-adoption.

La figure 1 résume les 14 hypothèses testées dans cet article. La confirmation des attentes est déterminée par les caractéristiques des utilisateurs (âge, sexe et profession) (H1), la qualité du SIC (H2), le support utilisateur (H3) et l'utilité perçue (H4). La satisfaction des utilisateurs est déterminée par les caractéristiques des utilisateurs (H5), la qualité du SIC (H6), le support utilisateur (H7), l'utilité perçue (H8) et la confirmation des attentes (H9). L'intention de continuer est influencée par les caractéristiques des utilisateurs (H10), la qualité du SIC (H11), l'utilité perçue (H12), la confirmation des attentes (H13) et la satisfaction des utilisateurs (H14).

2.3. Données et instruments des enquêtes

L'utilisation brute pour un groupe donné est la moyenne des réponses comprises entre 0 et 7 (0 = non utilisé, 1 = très rarement, 2 = rarement, 3 = plutôt rarement, 4 = occasionnellement, 5 = plutôt fréquemment, 6 = fréquemment, 7 = très fréquemment) pour chaque fonction considérée (n=15). Si la fonction est absente du SIC ou n'est pas déployée la valeur utilisée est zéro afin de mieux comparer les taux par sites et par professions. L'utilisation ajustée par profession est la moyenne des taux d'utilisation de base en ne considérant que les fonctions propres à la profession. Ainsi le taux de saisie des transmissions infirmières n'est pas comptabilisé chez les médecins. De même le taux de saisie de prescriptions d'imagerie ou de médicaments ne sont pas comptabilisés chez les infirmières. L'utilisation brute de prescriptions de médicaments ne sont pas comptabilisés chez les infirmières ou le personnel médico-administratif est ainsi un indicateur d'une délégitimation de saisie des médecins vers ces personnels.

L'échelle de Likert utilisée pour les réponses à la plupart des questions fermées des dimensions de la satisfaction (n=26), est comprise entre 1 et 7 (1 = fortement en désaccord, 2 = en désaccord, 3 = plutôt en désaccord, 4 = pas d'avis, 5 = plutôt d'accord, 6 = d'accord, 7 = entièrement d'accord).

Trois catégories d'utilisateurs concernées par le système du soin de patient ont été considérées : - personnels médicaux (MED) regroupant médecins et pharmaciens ; - personnels paramédicaux (PAR) regroupant aides-soignantes, infirmiers(ères) et cadres infirmiers ; - personnels médico-administratifs (PMA) regroupant secrétaires et assistantes sociales. Les utilisateurs n'ayant pas

déclaré une utilisation régulière d'au moins une des fonctions du SIC, ajustée selon les métiers ou n'ayant pas répondu à au moins une question portant sur les dimensions évalués du SIC sont exclus. Les résultats obtenus étant très similaires en 2011 et 2013, ont été combinées dans ce travail. Pour les utilisateurs ayant répondu à la fois en 2011 et en 2013 seule la moyenne des deux réponses a été conservée. Les analyses statistiques ont été réalisées avec R[®], Statview[®], et XLSTAT[®].

3. Résultats

3.1. Caractéristiques des utilisateurs

L'analyse porte sur 519 utilisateurs (taux de réponse = 18,7%), dont 100 à Corentin-Celton et Vaugirard (Tableau 1). La seule différence significative entre les deux groupes concerne l'âge moyen plus élevé chez les utilisateurs de Corentin Celton et Vaugirard que chez ceux de l'HEGP.

Tableau 1 – Caractéristiques des utilisateurs

	CCL-VGR (n=100)	HEGP (n=419)	Total (n=519)	<i>p-value</i>
Sexe masculin (%)	19,0	21,0	20,6	0,657
Profession médicale (%)	29,0	24,3	25,2	0,336
Age moyen (ET)	45,1 (10,4)	39,7 (11,6)	40,7 (11,6)	<0,001
Ancienneté sur le site (ET)	6,95 (4,69)	7,37 (4,49)	7,28 (4,53)	0,413
Activité temps plein (%)	84,8	88,6	87,9	0,307

ET = Ecart-type

3.2. L'utilisation du SIC

L'utilisation d'un SIC est fonction, d'une part, de la couverture fonctionnelle des produits déployés dans un site et, d'autre part, de la profession exercée. Elle a été évaluée, profession par profession, sur chacun des trois sites.

L'utilisation brute est globalement plus faible à Corentin Celton-Vaugirard qu'à l'HEGP, sites qui ne disposent pas, comme à l'HEGP, d'un dossier patient intégré dans lequel sont enregistrés les résultats d'examen et les différents comptes rendus (tableau 2). Les deux seuls domaines où les taux d'utilisation brute sont plus élevés, mais de façon non significative, à Corentin-Vaugirard qu'à l'HEGP sont la prise de rendez-vous et le codage PMSI, comparaisons devant également tenir compte pour leur interprétation du plus faible turn-over des malades dans les deux sites CCL-VGR de moyen et long séjour.

Le taux de délégation du codage PMSI et des prescriptions médicamenteuses des médecins vers le personnel paramédical est élevé à Corentin-Celton et Vaugirard comme à l'HEGP (tableau 3).

L'utilisation ajustée par la profession est plus élevée à l'HEGP qu'à CCL-VGR, et ceci quelle que soit la profession concernée (figure 2).

Tableau 2 - Utilisation brute par site

Fonctions (échelles de 0 à 7)*	CCL&VGR	HEGP	p-value
	(n=100)	(n=419)	
Identité mouvements	1,62 (2,32)	2,89 (2,71)	<0,001
Consultation ordonnances	2,01 (2,56)	2,84 (2,67)	<0,01
Consultation plans de soin	3,76 (2,72)	3,93 (2,85)	0,597
Consultation pancarte	3,29 (3,04)	3,75 (2,88)	0,154
Consultation transmissions infirmières	4,66 (2,64)	4,90 (2,61)	0,417
Consultation biologie	2,29 (2,82)	5,33 (2,24)	<0,001
Visualisation imagerie	1,77 (2,60)	3,76 (2,87)	<0,001
Consultation comptes rendus imagerie	1,25 (2,17)	3,31 (3,31)	<0,001
Saisie prescriptions biologie	1,26 (2,60)	1,50 (2,66)	0,404
Saisie prescriptions imagerie	1,12 (2,54)	1,64 (2,79)	0,146
Saisie prescriptions médicamenteuses	1,89 (2,94)	2,21 (2,98)	0,336
Planification des rendez-vous	2,32 (2,99)	1,98 (2,86)	0,296
Rédaction compte rendu hospitalisation	1,52 (2,47)	1,88 (2,78)	0,230
Codage PMSI	2,79 (3,17)	2,14 (2,97)	0,054
Utilisation moyenne	2,24 (1,48)	3,02 (1,39)	<0,001

* moyenne (Ecart-type)

Tableau 3 - Utilisation brute par catégorie d'utilisateurs

Fonctions (échelles de 0 à 7)*	MED		PAR	
	CCL-VGR	HEGP	CCL-VGR	HEGP
	(n=29)	(n=102)	(n=59)	(n=222)
Identité mouvements	0,62 (1,54)	1,22 (2,09)	2,21 (2,46)	4,05 (2,43)
Consultation ordonnances	3,22 (2,71)	3,94 (2,75)	1,70 (2,36)	2,10 (2,36)
Consultation plans de soin	4,00 (2,49)	2,93 (2,53)	4,39 (2,52)	5,37 (2,33)
Consultation pancartes	4,60 (2,66)	3,63 (2,91)	3,31 (3,03)	4,86 (2,66)
Consultation transmissions infirmières	4,59 (2,65)	4,56 (2,65)	5,42 (2,12)	6,06 (1,75)
Consultation biologie	5,43 (2,12)	5,86 (1,94)	1,20 (2,04)	5,44 (2,11)
Visualisation imagerie	4,12 (2,89)	5,06 (2,61)	0,84 (1,77)	3,39 (2,84)
Consultation comptes rendus imagerie	2,72 (2,81)	5,11 (2,60)	0,63 (1,45)	2,24 (2,48)
Saisie transmissions	2,12 (2,78)	3,35 (3,30)	2,21 (2,85)	4,08 (2,90)
Saisie prescriptions biologie	3,74 (3,35)	3,88 (3,11)	0,29 (1,33)	0,71 (1,95)
Saisie prescriptions imagerie	3,67 (3,37)	4,16 (3,18)	0,77 (1,11)	0,78 (2,08)
Saisie prescriptions médicamenteuses	3,35 (3,30)	2,69 (2,93)	1,56 (2,76)	2,68 (3,22)
Planification des rendez-vous	2,88 (3,02)	2,54 (3,01)	1,71 (2,77)	1,12 (2,25)
Saisie comptes rendus hospitalisation	1,48 (2,23)	2,84 (3,06)	1,32 (2,44)	0,80 (1,85)
Codage PMSI	3,48 (3,16)	3,01 (3,18)	2,36 (3,11)	1,54 (2,64)
Utilisation moyenne	3,34 (1,85)	3,61 (1,69)	1,96 (1,00)	3,05 (1,16)

* moyenne (Ecart-type)

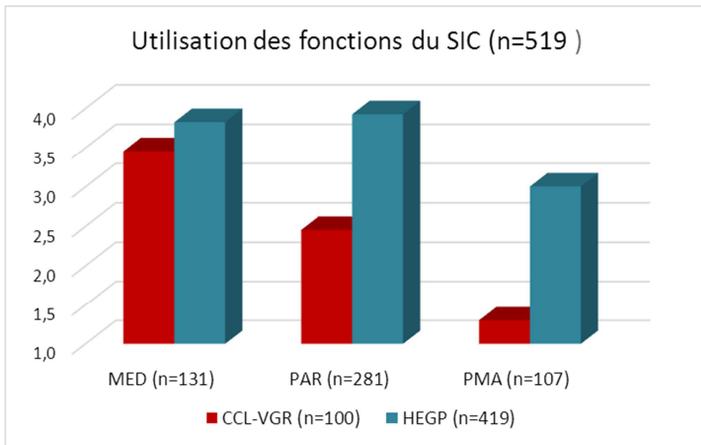


Figure 2 : Utilisation ajustée par la profession des différentes fonctions du SIC

La satisfaction globale est en moyenne de 4,89 (zone de satisfaction de l'échelle de Likert) sur le GH et significativement plus élevée ($p < 0,01$) à l'HEGP (4,98) qu'à Corentin-Celton et Vaugirard (4,50) (Figure 2). La cotation moyenne la plus élevée se retrouve chez les personnels médico-administratifs (5,10), la cotation la moins élevée chez les personnels médicaux (4,53). La cotation intermédiaire des personnels paramédicaux est de 4,96 et proche de celle des personnels médico-administratifs.

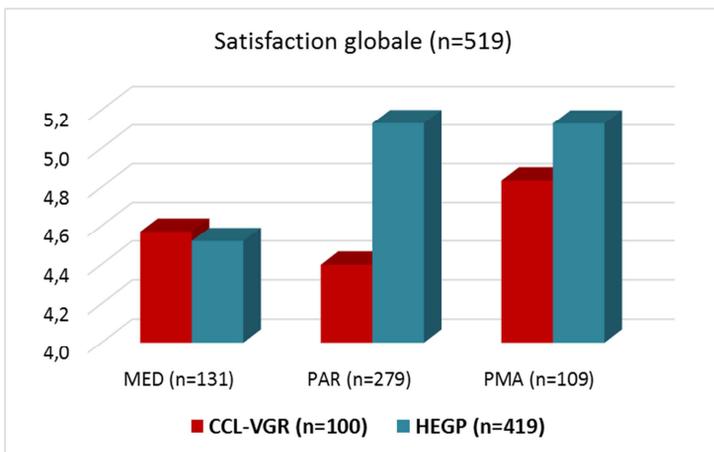


Figure 3 : Satisfaction globale par profession et par site

3.3. Les dimensions de la satisfaction

Sur les dimensions du modèle IT-PAM, tous les indicateurs, à l'exception de la qualité du support, sont supérieurs à l'HEGP à ceux de Corentin-Celton et Vaugirard. La différence n'est pas significative pour la qualité du support. La confirmation des attentes sur le SIC est en dessous de la moyenne à Corentin-Celton et Vaugirard. (Tableau 4).

Tableau 4 – Les dimensions de la satisfaction selon le site

Dimensions*	CCL-VGR (n=100)	HEGP (n=419)	p	Total (n=519)
Qualité du système	4,43 (1,10)	4,90 (1,04)	<0,001	4,81 (1,07)
<i>Qualité de l'information</i>	4,26 (1,41)	4,94 (1,27)		4,81 (1,32)
<i>Compatibilité avec le travail</i>	4,37 (1,37)	4,92 (1,35)		4,82 (1,37)
<i>Disponibilité</i>	4,54 (1,35)	4,99 (1,11)		4,91 (1,17)
<i>Facilité d'utilisation</i>	4,52 (1,28)	4,74 (1,33)		4,70 (1,32)
Qualité support utilisateur	4,52 (1,24)	4,35 (1,31)	0,262	4,40 (1,28)
<i>Support technique</i>	5,24 (1,39)	4,63 (1,49)		4,75 (1,49)
<i>Formation au SIC</i>	3,79 (1,56)	4,08 (1,53)		4,03 (1,54)
Utilité perçue	4,60 (1,30)	4,97 (1,25)	<0,01	4,90 (1,27)
Confirmation des attentes	3,90 (1,21)	4,40 (1,23)	<0,001	4,30 (1,24)
Satisfaction globale	4,50 (1,37)	4,98 (1,29)	<0,01	4,89 (1,32)
Intention de continuer	5,51 (1,28)	5,87 (1,10)	<0,01	5,80 (1,14)

* moyenne (Ecart-type)

Les indicateurs de satisfaction sont plus élevés chez les personnels médico-administratifs et soignants que chez les personnels médicaux. Le seul indicateur moyen inférieur à 4 est la confirmation des attentes chez les personnels médicaux. L'intention de continuer à utiliser le SIC dans l'avenir est très élevée quelle que soit l'hôpital et la catégorie professionnelle considérée (tableau 5).

Tableau 5 – Les dimensions de la satisfaction selon la profession

Dimensions*	MED (n=131)	PARA (n=281)	PMA (n=107)	p
Qualité du système	4,48 (1,24)	4,86 (1,01)	5,08 (0,91)	<0,001
Qualité du support utilisateur	4,36 (1,34)	4,27 (1,29)	4,72 (1,22)	<0,01
Utilité perçue	4,78 (1,54)	4,84 (1,21)	5,18 (0,98)	<0,05
Confirmation des attentes	3,91 (1,44)	4,44 (1,16)	4,42 (1,07)	<0,001
Satisfaction globale	4,53 (1,60)	4,98 (1,21)	5,10 (1,11)	<0,01
Intention de continuer	5,73 (1,35)	5,78 (1,09)	5,95 (0,99)	0,276

* moyenne (Ecart-type)

3.4. Les attentes des utilisateurs

La dernière partie de l'enquête était réservée aux attentes des utilisateurs, en particulier à leur souhaits d'évolution du SIC au sein du GH (tableaux 6 et 7).

Tableau 6 – Stratégie d'évolution du SIC par site

Stratégie d'évolution	CCL-VGR (n=97)	HEGP (n=390)	Total (n=487)
Statu quo (n=82)	20,6	15,9	16,8
Evolution à minima (n=41)	9,3	8,2	8,4
Convergence progressive (n=151)	33,0	30,5	31,0
Solution unique (n=213)	37,1	45,4	43,7
Total (%)	100	100	100

Toutes professions confondues, une très forte majorité d'utilisateurs (74,7 %) est en faveur d'une homogénéisation des fonctionnalités du SIC. Le souhait d'une solution unique est particulièrement marqué chez les personnels médicaux (88,0%) (Tableau 7).

Tableau 7 – Stratégie d'évolution du SIC par catégorie professionnelle

Stratégie d'évolution	MED (n=125)	PARA (n=260)	PMA (n=102)	Totaux (n=487)
Statu quo (n=82)	6,4	21,5	17,6	16,8
Evolution à minima (n=41)	5,6	10,4	6,9	8,4
Convergence progressive (n=151)	31,2	31,5	29,4	31,0
Solution unique (n=213)	56,8	36,5	46,1	43,7
Total (%)	100	100	100	100

De nombreux commentaires libres ont été recueillis à l'occasion des enquêtes. Parmi ceux-ci nous pouvons retenir les attentes portant sur l'amélioration de la rapidité de navigation, de l'ergonomie de l'interface, de l'accès à la formation pour les équipes de nuit ou encore le remplacement rapide des matériels les plus obsolètes par des matériels plus récents. A Corentin-Celton et Vaugirard, l'amélioration des performances du SIC est la principale préoccupation des utilisateurs. Elle n'apparaît qu'en troisième position à l'HEGP qui met en avant l'amélioration de l'ergonomie globale du SIC (réduction du nombre de « clics » pour atteindre une fonction donnée, par exemple).

3.5. Facteurs déterminants de la satisfaction et de l'intention de continuer

Les régressions simples montrent que la confirmation des attentes est corrélée positivement et significativement avec la qualité du SIC, la qualité du support et l'utilité perçue, négativement et significativement avec le sexe masculin et l'appartenance à une profession médicale (tableau 8). La satisfaction globale est corrélée positivement et significativement avec la qualité du SIC, la qualité du support, l'utilité perçue et la confirmation des attentes, négativement et significativement avec l'appartenance à une profession médicale. L'intention de continuer à utiliser le SIC est corrélée positivement et significativement avec la qualité du SIC, la qualité du support, l'utilité perçue, la confirmation des attentes, la satisfaction globale. Elle n'est pas corrélée significativement avec les caractéristiques individuelles (âge, sexe, profession). Après ajustement dans les modèles multidimensionnels, les caractéristiques individuelles des utilisateurs ne sont plus liées significativement à la satisfaction globale. La qualité du SIC, l'utilité perçue et la confirmation des attentes sont les paramètres les plus liés à la satisfaction. L'intention de continuer à utiliser le SIC est positivement corrélée avec la qualité du SIC, la qualité du support, l'utilité perçue et la satisfaction globale mais non avec les caractéristiques individuelles des utilisateurs.

Tableau 8 – Régressions des facteurs déterminants de la satisfaction

Dimension	Hyp.	Régression simple		Régression multiple	
		Coef.	<i>p</i>	Coef.	<i>p</i>
Confirmation des attentes					
Terme constant				0,314	0,163
Âge	H1	-0,002	0,682	0,002	0,616
Sexe masculin	H1	-0,440	<0,01	-0,230	<0,05
Profession médicale	H1	-0,526	<0,001	-0,148	0,132
Qualité du SIC	H2	0,812	<0,001	0,636	<0,001
Qualité du support	H3	0,345	<0,001	-0,021	0,557
Utilité perçue	H4	0,573	<0,001	0,212	<0,001
<i>R² ajusté (p)</i>				0,521	<0,001
Satisfaction globale					
Terme constant				-0,368	0,028
Âge	H1	-0,004	0,458	-0,002	0,534
Sexe masculin	H1	-0,248	0,083	0,037	0,619
Profession médicale	H1	-0,478	<0,001	-0,009	0,906
Qualité du SIC	H6	1,065	<0,001	0,844	<0,001
Qualité du support	H7	0,479	<0,001	0,040	0,123
Utilité perçue	H8	0,670	<0,001	0,069	<0,05
Confirmation des attentes	H9	0,741	<0,001	0,172	<0,001
<i>R² ajusté</i>				0,767	<0,001
Intention de continuer					
Terme constant				2,555	<0,001
Âge	H1	0,002	0,617	0,002	0,995
Sexe masculin	H1	-0,103	0,408	-0,034	0,748
Profession médicale	H1	-0,096	0,406	0,143	0,164
Qualité du SIC	H11	0,602	<0,001	0,156	0,050
Utilité perçue	H12	0,515	<0,001	0,277	<0,001
Confirmation des attentes	H13	0,446	<0,001	0,071	0,133
Satisfaction globale	H14	0,487	<0,001	0,164	<0,01
<i>R² ajusté</i>				0,391	<0,001

4. Discussion

Alors que les trois hôpitaux du groupe hospitalier universitaire Paris-Ouest utilisent deux SICs différents en matière de dossier patient électronique, la même enquête de satisfaction des utilisateurs a pu être réalisée par auto-questionnaires en 2011 et 2013. Les résultats très concordants entre les sites et lors des deux enquêtes ont permis, d'une part, de regrouper les données de 2011 et 2013 et, d'autre part, de regrouper les données de Corentin-Celton et Vaugirard (CCL-VGR) qui partagent le logiciel Actipidos/GFI et de les comparer à celles de l'HEGP utilisateur de DxCare/Medasys. Les trois sites partagent par contre le même système de gestion des laboratoires et de l'imagerie.

La mesure du degré d'utilisation doit faire partie de toute évaluation. Elle est la plus difficile à quantifier car les fonctions utilisées dépendent à la fois des produits concernés (la fonction est ou

n'est pas disponible dans le progiciel) mais également de l'ancienneté du déploiement et du métier des professionnels (chaque catégorie de personnel est plus ou moins concernée par chaque catégorie de fonction). Dans ce travail, et en se basant sur le model IT-PAM, les résultats de l'enquête montrent des taux d'utilisation brute et ajustée par la profession plus élevés à HEGP qu'à CCL-VGR associées à une plus grande satisfaction globale et à une plus grande intention de continuer à utiliser le SIC dans les années à venir. S'il ne peut être déduit de cette enquête une quelconque relation causale entre richesse fonctionnelle et satisfaction des utilisateurs, l'un des résultats les plus important est sans doute le désir de l'ensemble des utilisateurs d'homogénéisation du SIC sur l'ensemble du groupe HUPO.

Les différences observées entre les catégories professionnelles correspondent à celles retrouvées dans la littérature [5,15,16], les SICs étant développés d'avantage en direction des personnels paramédicaux et administratifs que des médecins. En particulier, l'utilisation de la prescription connectée, si elle apporte des garanties de sécurité pour les personnels infirmiers, demande toujours plus de temps (15 à 30%) pour les médecins que la prescription sur papier [12]. Le taux élevé de prescription déléguée par les médecins vers les personnels paramédicaux, et ceci dans les trois sites, pose un problème de responsabilité en cas d'accident iatrogénique et doit faire l'objet d'une attention toute particulière. Elle implique en particulier que soit vérifiée que toute prescription relevant d'une responsabilité médicale mais saisie par un personnel paramédical soit suivie d'une validation médicale enregistrée dans le dossier électronique comme le permet le logiciel DxCare.

L'amélioration de la qualité des interfaces est la demande la plus fréquemment citée par les utilisateurs de SICs. Elle n'est pas spécifique des logiciels utilisés (DxCare ou Actipidos dans le GH). Elle semble d'autant plus justifiée que les utilisateurs se sont habitués très rapidement à l'utilisation de tablettes ou de téléphones intelligents dans leur vie professionnelle ou familiale. Ces développements sont planifiés par les principaux fournisseurs de progiciels présents sur le marché, mais malheureusement en décalage par rapport à la rapidité d'évolution du marché global de la bureautique ou de la téléphonie mobile.

Le faible taux de réponses observé est habituel dans les enquêtes répétitives en ligne et limite la généralisation des résultats mais probablement moins l'analyse évolutive des données et des corrélations entre les dimensions de la satisfaction. Si les niveaux de satisfaction diffèrent d'un site à l'autre du fait en particulier de l'utilisation de SIC différents, une grande homogénéité entre les indicateurs d'un même site a été observée d'une enquête à l'autre. Ces premiers résultats laissent penser que de telles comparaisons pourraient être étendues à d'autres hôpitaux, d'autres SIC et d'autres niveaux de déploiement. A terme, les indicateurs de satisfaction et plus généralement d'évaluation des SICs devraient être utiles aux sites pour améliorer l'utilisation et la satisfaction de leur SICs comme aux fournisseurs pour améliorer leurs produits.

Références

- [1] Ammenwerth E, Mansmann U, Iller C, *et al.* Factors affecting and affected by user acceptance of computer-based nursing documentation: results of a two-year study. *J Am Med Inform Assoc JAMIA* 2003;10:69–84.
- [2] Degoulet P. *Hospital Information Systems*. In: Venot A, Burgun A, Quantin C (eds). *Medical Informatics, eHealth Fundamentals and Applications*. Paris: Springer-Verlag 2014; pp 289-313.

- [3] Van der Meijden MJ, Tange HJ, Troost J, *et al.* Determinants of Success of Inpatient Clinical Information Systems: A Literature Review. *JAMIA* 2003;10:235–43.
- [4] Holden RJ, Karsh B-T. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010;43:159–72.
- [5] Pynoo B, Devolder P, Voet T, *et al.* Assessing hospital physicians' acceptance of clinical information systems: a review of the relevant literature. *Psychol Belg* 2013; 53(2):15-31.
- [6] Delone WH. The DeLone and McLean model of information systems success: a ten-year update. *J Manag Inf Syst* 2003;19:9–30.
- [7] Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q* 1989;13:319-40.
- [8] Lee Y, Kozar KA, Larsen KRT. The Technology Acceptance Model: Past, Present, and Future. *Commun Assoc Inf Syst* 2003;12: Article 50.
- [9] Bhattacharjee A. Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Q* 2001;25:351-70.
- [10] Venkatesh V, Morris M, Davis G, *et al.* User Acceptance of Information Technology: Toward a Unified View. *Manag Inf Syst Q* 2003;27:425-78.
- [11] Oliver RL. A Cognitive Model of Antecedents and Consequences of Satisfaction Decisions. *J Mark Res* 1980;17:460–9.
- [12] Beuscart-Zéphir MC, Pelayo S, Anceaux F, *et al.* Impact of CPOE on doctor-nurse cooperation for the medication ordering and administration process. *Int J Med Inf* 2005;74:629–41.
- [13] Zapletal E, Rodon N, *et al.* Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. *Stud Health Technol Inform* 2010 ; 160:193–7.
- [14] Palm J-M, Colombet I, Sicotte C, *et al.* Determinants of User Satisfaction with a Clinical Information System. *AMIA Annu Symp Proc* 2006; 2006:614–8.
- [15] Palm J-M, Dart T, Dupuis I, *et al.* Clinical information system post-adoption evaluation at the georges pompidou university hospital. *AMIA Symp* 2010; 2010:582–6.
- [16] Abdekhoda M, Ahmadi M, Dehnad A, Hosseini AF. Information technology acceptance in health information management. *Methods Inf Med.* 2014; 53(1):14-20.

Adresse de correspondance

Brahim Hadji Ing MSc, PhD Cand
Hôpital Européen George Pompidou
20 Rue Leblanc, 75015 Paris
Tel : +33 (0)1 56 09 20 30
Fax : +33 (0)1 56 09 20 30
brahim.hadji@egp.aphp.fr

Indicateurs en biologie et en imagerie au sein des systèmes d'information de santé

Impact indicators for radiology and biology information systems

Mher B. Joulakian¹, Nicolas Griffon^{1,2}, Matthieu Schuers¹, Eric Lepage², Céline Savoy-Collet³, Soumaya Skalli⁴, Philippe Massari¹, Stéfan J. Darmoni^{1,2}

¹ Service d'Informatique Biomédicale, CHU de Rouen & CISMeF, TIBS, LITIS EA 4108, Université de Rouen, France

² LIMICS, INSERM U1142, Université Paris 13, Université Paris Descartes

³ Service d'Imagerie Médicale, CHU de Rouen, France

⁴ Service de Bactériologie, CHU de Rouen, France

Résumé

Introduction : les bénéfices des systèmes d'information de production de soin (SIPS) ont été beaucoup étudiés. Toutefois, les résultats sont hétérogènes et les méthodes d'évaluation ne sont pas adaptées au système de santé français. L'objectif de cette étude est de développer un outil de description des SIPS et de recenser les indicateurs couramment utilisés pour en mesurer les bénéfices en biologie et en imagerie. Ces outils permettront de faire des comparaisons entre différents établissements de santé. Méthodes : La description fine des processus d'un centre hospitalier a servi de base pour la création d'un outil de description des SIPS. Une revue de la littérature a permis de recenser les indicateurs utilisés au niveau international pour évaluer les bénéfices des SIPS de biologie et d'imagerie. Résultats : L'outil de cartographie a été utilisé avec succès par 4 hôpitaux pour décrire leur SIPS. De nombreux indicateurs ont été relevés dans la littérature. Certains ont été sélectionnés et parfois modifiés pour permettre d'évaluer précisément chaque sous-processus. Conclusion : Les éléments nécessaires à l'évaluation des bénéfices des différents SIPS sont en place.

Abstract

Introduction: Health information system (HIS) has been widely studied in order to evaluate their effectiveness. However the results are heterogeneous and the methods are rarely adapted to the french health system. This work aims at developing a tool for HIS description and listing indicators frequently used to assess its benefits in the fields of laboratory and radiology. Those tools will allow inter HIS comparison. Methods: Processes were finely described in a university hospital in order to create a processes description tool. A short literature review allowed us to list the most commonly used indicators in the field of research concerning the HIS benefits. Results:

The process describing tool was successfully used by four hospitals to describe their HIS. Several indicators have been selected in the international literature and some of them have been modified to describe more precisely the sub processes. Conclusion: All the items for the evaluation of the benefits of the HIS are ready.

Mots clés : Dossiers médicaux électroniques ; Systèmes d'information hospitaliers ; Systèmes d'information de laboratoire d'analyses médicales ; Systèmes d'information de radiologie ; Indicateurs qualité santé.

Keywords: *Electronic health records; Hospital information systems; Clinical laboratory information systems; Radiology information systems; Quality indicators, Health care.*

1 Introduction

Comme le souligne le rapport de l'Agence Nationale d'Appui à la Performance des établissements de santé et médico-sociaux (ANAP) [1], de très nombreux travaux ont été menés dans le domaine de l'évaluation des Systèmes d'Information en Santé (SIS) [1]. Ainsi, plusieurs rapports nord-américains [3][4], en particulier les travaux de l'Agency for Healthcare Research & Quality (AHRQ) [5][6] montrent les bénéfices des Technologies de l'Information et de la Communication (TIC). Ceux-ci portent principalement sur l'amélioration de la qualité des soins, de la productivité hospitalière et de la coordination des soins [7]. Des travaux ont également été menés sous l'impulsion de la Communauté Européenne, basée sur une revue extensive de preuves scientifiques. Ils concluent à l'apport des TIC dans le cadre de la prévention des erreurs médicamenteuses et la réduction des effets indésirables [8].

L'analyse de la littérature montre que si de nombreux travaux ont été menés, l'évaluation de l'impact des SIS reste un domaine important de recherche. En effet, les études ont généralement été effectuées de façon descriptive et mono-disciplinaire. Elles sont fondées sur des approches thématiques au sein de domaines tels que l'informatique, les sciences cognitives, l'économie et l'organisation. Le plus souvent elles ne sont pas applicables à un environnement multi-acteurs. De ce fait, elles n'intègrent pas l'analyse des conséquences de la mise en œuvre d'une informatisation au sein d'un système impliquant de multiples participants au processus de soins. Si de nombreuses revues systématiques démontrent le bénéfice des TIC, les conclusions restent discutées du fait de certaines études discordantes, de leur approche unidimensionnelle, de la conduite de ces projets sur les centres hospitalo-universitaires, rendant ces résultats non généralisables. De plus, peu d'études rapportent des résultats objectifs portant sur les conditions de travail, l'intérêt public et social, l'interopérabilité et le partage de l'information, le parcours de soins, l'accès du patient à l'information médicale. Sur ces points, il est à rappeler que les différences culturelles et les modes d'organisation différents entre les pays rendent plus difficile la transposition des méthodes et des résultats d'une étude particulière, la nécessité d'outiller le système de santé français est donc majeure. Le projet « Évaluation de la création de valeur par l'usage des systèmes d'information de production de soins » (UPRES) a pour objectif de mettre au point des outils d'évaluation de l'impact des Systèmes d'Information de Production de Soins (SIPS), d'un point de vue quantitatif et qualitatif pour les établissements de santé, mais aussi en termes d'organisation et d'interaction avec l'ensemble des professionnels de santé. Ce travail de recherche se décline ainsi selon 3 axes : (a) l'impact quantitatif et qualitatif du déploiement des

circuits de la biologie et de l'imagerie sur les délais et la qualité des soins, sur lequel cet article se focalise ; (b) l'impact quantitatif et qualitatif de la production des comptes-rendus d'hospitalisation sur le processus de sortie du patient ; (c) les impacts organisationnels, internes et externes à l'hôpital, du déploiement d'un dossier patient informatisé.

Un élément essentiel des systèmes d'informations hospitalier est l'informatisation des circuits de biologie et d'imagerie, de la prescription au retour des résultats au médecin susceptible de modifier la prise en charge du patient. L'informatisation du circuit des examens biologiques a débuté de longue date (1989 au CHU de Rouen [10]), mais dans la plupart des cas cette informatisation n'est que partielle, se limitant souvent à l'informatisation des laboratoires et au retour des résultats d'examens biologiques. Le niveau d'informatisation et l'intégration à l'ensemble du Système d'Information Hospitalier (SIH) sont très variables d'un établissement à l'autre. L'informatisation du circuit des examens d'imagerie a été, le plus souvent, plus tardive et est aussi très variable en fonction des établissements [11].

Ces différences de niveaux d'informatisation et d'intégration des circuits des examens biologiques et des examens d'imagerie permettent de proposer une méthode générale fondée sur des comparaisons inter-établissements. Les modes d'informatisation différents ou l'absence d'informatisation de certaines phases des processus selon les établissements permettront d'étudier l'impact de l'informatisation et de son intégration

Ce travail a pour objectif de construire un outil permettant de décrire simplement le niveau d'informatisation des différents sous-processus des circuits de biologie et d'imagerie, et de proposer des indicateurs permettant d'évaluer chacun des sous-processus en fonction de son niveau d'informatisation, en comparaison avec d'autres établissements informatisés différemment.

2 Matériel et méthodes

Ce travail de recherche repose sur un processus d'évaluation aussi pragmatique que possible pour pouvoir être appliqué par tout type d'établissement de santé mais également par toute structure de santé souhaitant évaluer la valeur ajoutée de ses investissements financiers ou bien de ses recommandations, comme cela peut être le cas d'une Agence Régionale de Santé (ARS), de la Direction Générale de l'Offre de Soins (DGOS), ou bien de la Haute Autorité de Santé (HAS).

L'étude comprend quatre phases : une phase initiale d'étude des processus et de l'informatisation des circuits des examens de biologie et d'imagerie, ainsi que le choix des indicateurs (qui est le sujet central de cet article), suivie de deux phases de mesure des indicateurs, et enfin une phase d'évaluation des impacts de l'informatisation à partir des données recueillis lors des phases précédentes. La phase 2 commencera en juin 2014.

2.1 Description des processus

Pour chacun des circuits, un responsable a coordonné l'analyse des processus qui est réalisée en collaboration avec les experts métiers de deux des sites pilotes (AP-HP et CHU de Rouen). Pour chaque établissement, on a réalisé une description détaillée des processus et des flux d'information à supporter par le système d'information. La mise en commun des processus a permis de construire un outil permettant théoriquement de décrire la plupart des circuits existant

dans les établissements de santé en France. Cet outil a été proposé aux hôpitaux partenaires du projet (voir remerciement) qui ont ainsi pu cartographier l'informatisation de leurs circuits de façon relativement aisée : pour chaque processus ou sous processus, chaque établissement a pu préciser le niveau d'informatisation, les flux de données inter-étapes et les données susceptibles d'être utilisées pour le calcul d'indicateur. Cette cartographie permettra d'identifier certaines phases des processus dont le mode ou le niveau d'informatisation est différent entre les établissements et pour lesquelles on pourra, par comparaison d'indicateurs, évaluer l'impact de l'informatisation.

2.2 Détermination des indicateurs

Le répertoire d'indicateur doit permettre d'évaluer l'impact sur :

- l'établissement de santé
- l'extra-hospitalier, en particulier la médecine de ville
- la coopération sanitaire
- le patient

Une revue de la littérature a été menée pour identifier les indicateurs utilisés pour évaluer les sous-processus de biologie et d'imagerie. Les principaux articles récents sur le sujet ont été sélectionnés dans PubMed. Les listes de références ont été consultées pour enrichir cette bibliographie.

A partir de la liste des indicateurs extrait de la littérature, nous avons sélectionné ceux qui étaient a priori applicables quel que soit le niveau d'informatisation des sous processus concernés, qui soit relativement faciles à mesurer quel que soit l'établissement de santé considéré et qui semblaient pertinents par rapport à l'objectif de mesure d'impact comparative entre plusieurs établissements. Chaque indicateur est décrit selon une fiche type proposée par l'ANAP.

3 Résultats

3.1 Description des processus

La description des processus a permis d'aboutir à un schéma partagé entre le CHU de Rouen et l'AP-HP présenté en figure 1 et a priori compatible avec les autres structures participant à ce projet. L'outil de cartographie des circuits de biologie et d'imagerie est directement inspiré de cette description des processus. La figure 2 présente un aperçu de cet outil. L'informatisation des circuits des quatre partenaires a été cartographiée à l'aide de cet outil [60], chaque partenaire précisant, pour chaque étape, l'informatisation ou la non informatisation, les flux de données inter-étapes, et les données disponibles susceptibles d'être utilisées pour le calcul des indicateurs (dates et heures, connexions, codes des analyses...).

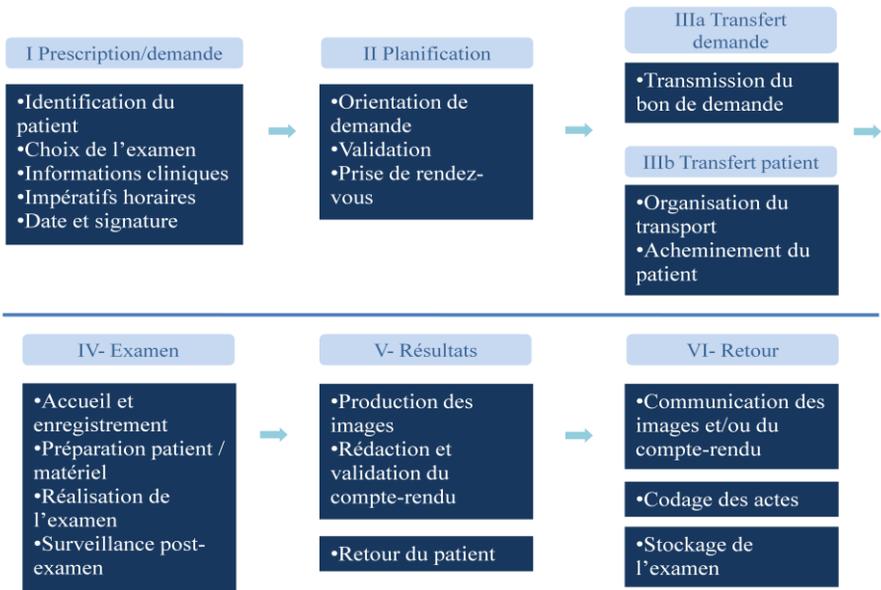


Figure 1 : Description des processus pour le circuit d'imagerie.



Figure 2 : Cartographie (partielle) du circuit de biologie aux CHU de Rouen et de Lille.

3.2 Détermination des indicateurs

L'analyse de la littérature [12]-[59] a révélé que les variables d'évaluation les plus couramment retrouvées appartiennent aux catégories suivantes : (a) indicateurs temporels (de délais et de temps consacré à la réalisation d'une tâche donnée), (b) les indicateurs de qualité, (c) les indicateurs de conformité.

Indicateurs de volume (dans la période de l'étude)* [48]-[50]	Nombre total d'examens réalisés - stratification par type [48]-[50]
	Nombre d'examens par séjour rapporté à la durée du séjour
	Nombre d'examens par médecin prescripteur
	Nombre d'examens par médecin radiologue
	Nombre d'examens par modalité
Indicateurs de qualité	Taux de conformité des bons de demande selon les critères IPAQSS-CDEI édité par la HAS [46]
	Adéquation des prescriptions aux Guide de Bon Usage (GBU) [49],[50]
	Taux d'images visualisés par les cliniciens [52]
	Taux de comptes-rendus lus par les cliniciens [52]
	Taux d'examens perdus dans l'historique des patients [55]
	Doses d'irradiation reçus par les patients, cumulées et par examen [54]
Indicateurs temporels* [47],[49],[51],[56],[59]	Délai total du circuit [47],[49],[51]
	Délai total du circuit dans le cadre de l'urgence
	Temps passé à la prescription d'un examen
	Délai entre la prescription et la réception de la demande dans le service d'imagerie [49]
	Temps de transmission des demandes entre les services cliniques et le service d'imagerie
	Délai entre la prise en compte de la demande et la programmation du rendez-vous avec le service clinique
	Délai entre la date de programmation du rendez-vous avec le service clinique et la date prévue de l'examen

Tableau 1 : Liste (partielle) des indicateurs pressentis pour le circuit d'imagerie.

* Au sein de ces champs les indicateurs pressentis ne sont pas exactement les mêmes que ceux trouvés dans les articles dont ils sont inspirés, c'est pourquoi nous avons référencé l'ensemble du champ plutôt que les indicateurs individuellement. Cependant, lorsque l'indicateur n'a pas été modifié, nous l'avons référencé directement.

Sur la base de cette revue de la littérature nous avons élaboré un panel d'indicateurs pressentis pour l'imagerie qui sont disponibles dans le **Erreur ! Source du renvoi introuvable.** Les indicateurs de biologie sont assez similaires à ceux de l'imagerie.

Dans le **Erreur ! Référence non valide pour un signet.**, nous proposons une matrice croisant ces indicateurs avec les processus vus précédemment (voir figure 1). Cette matrice va permettre de sélectionner les indicateurs pertinents en fonction de la situation des établissements sélectionnés dans l'étude : à titre d'exemple, l'adéquation des prescriptions aux recommandations est susceptible d'être modifiée par le niveau d'informatisation de la prescription.

Tableau 2 : Matrice (partielle) de croisement de l'impact potentiel, sur les indicateurs, de l'informatisation des différents processus du circuit de l'imagerie.

Impact potentiel de l'informatisation du sous processus sur l'indicateur :	Nombre total d'examens sur une période donnée	Nombre d'examens par patient et par séjour rapporté à la durée du séjour	Taux de conformité des bons de demandes selon les critères IPAQSS-CDEI édité par la HAS.	Adéquation des prescriptions aux Guides de Bon Usages (GBU).	Délai total (De la prescription à la transmission des résultats)	Délai prescription à la réception de la demande dans le service d'imagerie	Délai entre la réception de la demande et la programmation du rendez-vous avec le service clinique	Délai entre la prise en compte du rendez-vous dans le service clinique et la programmation du transport du patient avec le service des transports
Aucun ■								
Possible ■								
Probable ■								
PRESCRIPTION	3	3	3	3	3	3		
DEMANDE			3		3	3		
TRANSMISSION DEMANDE			3		3	3	3	
PRISE DE RENDEZ-VOUS	3			3	3		3	3
ORGANISATION ACHEMINEMENT DU PATIENT	3				3			3
ENREGISTREMENT DEMANDE DANS LE RIS					3			
REALISATION DE L'EXAMEN	3				3			
PRODUCTION DES IMAGES	3				3			
DICTEE COMPTE-RENDU					3			
REDACTION COMPTE-RENDU					3			
TRANSMISSION DES RESULTATS	3				3			
PRISE EN COMPTE PAR LES CLINIENS	3	3			3			

4 Discussion

Nous avons présenté dans ce travail la première étape du projet UPRES : la création d'un outil de cartographie et les principaux indicateurs des circuits d'imagerie et de biologie étudiés dans la littérature.

L'outil de cartographie a été facilement adopté par les partenaires du projet et semble donc aisé à transposer à une pratique plus généralisée d'évaluation de l'impact de l'informatisation. Les indicateurs identifiés ne seront pas tous nécessairement utilisés dans le cadre du projet UPRES, et d'autres seront très vraisemblablement évoqués, néanmoins, ce tour d'horizon des éléments qui permettent d'évaluer l'impact de l'informatisation des circuits d'imagerie, nous évite d'omettre certains aspects de la question et assure une plus grande comparabilité des résultats à la littérature internationale.

L'analyse de la littérature pour la biologie nous a montré les principaux impacts des logiciels d'aide à la prescription sur le circuit des examens biologiques : la réduction du délai total du circuit, la diminution des erreurs de transcription, un gain de temps des infirmières, une diminution de la phase d'enregistrement au niveau du laboratoire, la suppression de la réimpression des demandes, la réduction des demandes ambiguës et des demandes non conformes, la réduction des redondances, une meilleure utilisation des résultats, un meilleur respect des recommandations des laboratoires et une plus grande facilité de création de formulaires contenant des actes biologiques. [40] En marge de l'informatisation, l'utilisation du

pneumatique ou d'autres moyens automatisés du transport des prélèvements a également ont montré leur impact sur le délai total du circuit [41].

L'analyse de la littérature pour les examens d'imagerie a mis en évidence les principaux impacts attendus : l'impact de l'informatisation peut sembler globalement positif, comme peuvent le suggérer les diminutions systématiques concernant le délai de réalisation totale du circuit [47], [59]. Par contre, l'absence d'étude globale concernant les autres indicateurs nous contraint à analyser l'impact de l'informatisation circonscrit à une ou deux étapes du circuit. Ce faisant, nous nous apercevons qu'une grande part des études montre un effet positif sur les indicateurs d'évaluation utilisés. Concernant la phase de prescription, l'informatisation entraîne une baisse du nombre d'examens consommés [48]-[50], une plus grande adéquation des prescriptions aux guides de bonnes pratiques [49], [50], une disparition de l'étape d'envoi de la demande aux services d'imagerie [47] ce qui entraîne une diminution du délai ainsi qu'une économie en personnel par rapport à une demande manu-portée, postée ou faxée. Il existe cependant un impact négatif de l'informatisation dans cette phase et cela concerne le temps que le clinicien passe à remplir la demande électronique qui semble être plus long que pour une demande manuscrite [48], mais il faut considérer le fait que l'évaluation s'est faite dans une phase précoce du déploiement de l'outil informatique, les utilisateurs pouvaient donc ne pas en être assez familiers.

Concernant les phases autour de la réalisation de l'examen, nous n'avons pas trouvé d'étude évaluant l'impact de l'informatisation, que cela soit pour la planification du rendez-vous, pour l'organisation et le transport du patient, ou bien pour l'accueil de celui-ci durant lequel des problèmes peuvent se poser, notamment en termes d'identité-vigilance, d'erreur d'examen, ou bien de temps d'attente. Dans la phase allant de l'acquisition des images jusqu'à la production du compte rendu final, l'informatisation semble avoir un impact très positif. Le temps global de cette phase semble systématiquement diminué [47][58][59], la mise à disposition immédiate des images élimine l'étape de transport des films vers les services cliniques [51]. De plus, à part quelques éléments qui peuvent poser problème, telle la reconnaissance vocale, les outils informatiques semblent assez bien accueillis par les radiologues [57]-[59]. Sur le versant des services cliniques, l'informatisation semble avoir aussi un impact positif, qu'il s'agisse de la meilleure performance des cliniciens à faire un bon diagnostic sur images numériques que sur films [55], ou du temps gagné par ceux-ci à retrouver les examens de leurs patients, surtout lorsqu'il s'agit d'examens anciens. De plus les enquêtes de satisfaction semblent montrer que les cliniciens ont une assez bon ressenti du passage à l'informatique [55].

Lors de notre analyse de la littérature nous nous sommes aussi rendu compte que l'informatisation du circuit de l'imagerie permet de mesurer de nouveaux indicateurs inaccessibles auparavant, c'est le cas du taux de CR consultés par les cliniciens dans l'étude de Hurlen P. [52], où l'on voit qu'il existe une part non négligeable de CR ignorés. Ce dernier point illustre l'intersection entre les axes 1 et 2 de ce projet, et la possibilité d'envisager de nouvelles hypothèses par rapport à celles posées au départ du projet.

Remerciements

Les auteurs remercient : la DGOS et l'ANAP qui ont en partie financé le projet UPRES dans le cadre du PREPS 2012, les autres membres du consortium UPRES (équipe des CHU de Lille, Nice et de l'APHP impliquées dans le domaine des systèmes d'information de santé, le

laboratoire de biostatistique du CHU de Rouen , l'unité INSERM U1142, le Centre de Recherche pour l'Étude et l'Observation des Conditions de vie (CREDOC) et une équipe du Centre de Gestion de l'École des Mines de Paris) pour avoir validé et utilisé les outils de cartographie produits à Rouen, Naémé Nekooguyan, Vincent Gallais, Christophe André, Sylvie Cormont, Déa Giardella, Marie Gueudin, Élodie Dehedin et Céline Lefevre qui ont travaillé à la description des processus dans le cadre du projet TeRSan (ANR-11-TECS-019-03).

Références

- [1] Alain D, Douesneau C. Piloter les projets SI par la valeur : Méthode et expérimentations, ANAP 2012.
- [2] Ammenwerth E, de Keizer N. An inventory of evaluation studies of information technology in health care. *Methods Inf Med* 2005;44:44e56.
- [3] President's Information Technology Advisory Committee. Revolutionizing Healthcare through information technology, 2004 http://www.nitrd.gov/pitac/reports/20040721_hit_report.pdf
- [4] President's Council of Advisors on Science and Technology. Report to the president realizing the full potential of health information technology to improve healthcare for Americans: the path forward. December 2010 <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-health-it-report.pdf>
- [5] Shekelle P. G., Morton S. C., Keeler E. B., Costs and Benefits of Health Information Technology, Southern California Evidence-based Practice Center, Agency for Healthcare Research and Quality. Rockville, MD: s.n., April 2006, Evidence Report/Technology Assessment, No. 132, Publication No. 06-E006.
- [6] Goldzweig Lubick C., Towfigh A., Maglione M., Shekelle P. G., "Costs and Benefits of Health Information Technology: New Trends from Literature", *Health Affairs*, Millwood, 2009, Vol. 2, 28, p. w282-w293.
- [7] Schoen C, Osborn R, Doty MM, Squires D, Peugh J, Applebaum S: A survey of primary care physicians in eleven countries, Perspectives on care, costs and experiences. *Health Aff* 2009, 28(6):21171-21183.
- [8] Stroetman V. N, Thierry J. P., Stroetman K. A., Dobrev A., eHealth for Safety - Impact of ICT on Patient Safety and Risk Management, European Commission. Luxembourg : Office for Official Publications of the European Communities, 2007.
- [9] Finkelstein J, Knight A, Marinopoulos S, Gibbons MC, Berger Z, Aboumatar H, et al. Enabling patient-centered care through health information technology. *Evid Rep Technol Assess (Full Rep)*. juin 2012;(206):1-1531.
- [10] Darmoni, SJ.; Massari, P; Allaire, P; Caffarel, JL; Monconduit, M; Baldenweck, M & Hecketsweiler, P. REX: Extralaboratory Cycle Of Information Processing From Laboratories To Care Units. Proc. Annual International Conference of the IEEE Vol.14 Engineering in Medicine and Biology Society, Volume 3, Pages 1253-1254, 1992.
- [11] Griffon, N; Savoye-Collet, C; Massari, P; Christel, D & Darmoni, SJ. An interface terminology for medical imaging ordering purposes. *AMIA* 2012;2012:1237-43.
- [12] Bonner A., Shrufi A., Rozen S., A Database Benchmark for High-Throughput Workflow Management. In: P. Apers et al. (eds.): Proc. EDBT'96, Avignon, 1996.
- [13] Reichert M., Dadam P., Towards Process – oriented Hospital Information Systems: Some Insights into Requirements, Technical Challenges and Possible Solutions, Jahrestagung der GMDS (GMDS'98), Bremen, 1998.

- [14] Mohammad N., Mohammad Y., Verdier C., Auray J P., Evaluation of health information system: reasons for failure, ICSSHC Congress, 3-5 September, Lyon, France, September 2008.
- [15] Gremy F., Fessler J.M., Bonnin M., Information systems evaluation and subjectivity. International Journal of Medical Informatics Volume 56, Issues 1-3, December 1999, Pages 13-23.
- [16] Levy A Y., Srivastava D., Kirk T., Data Model and Query Evaluation in Global Information Systems. Journal of Intelligent Information Systems, 5, 121-143 (1995), Boston, USA.
- [17] Walsham G., Sahay S., GIS for district-level administration in India: Problems and opportunities. Mis Quarterly 23(1): 39-65, (1999).
- [18] Hu D., Using actor – network theory to understand inter – organizational network aspects for strategic information systems planning, Master thesis, University of Twente, Netherlands, 2011.
- [19] Tomasi E., Facchini LA., Santos Maia MF., Health information technology in primary health care in developing countries: a literature review. Bulletin of the World Health Organization. 2004;82:867-874.
- [20] Smits M T., Luijsterburg J C., Van Ham M A., Dynamics of Inter –Organizational Information Systems in Health Care from an Actor Network Perspective, Journal of Philosophical logic, 2009.
- [21] Littlejohns P., Wyatt J. C., Garvican L., Evaluating computerised health information systems: hard lessons still to be learnt, Information in practice, BMJ Volume 326, 19 April 2003.
- [22] Guide des indicateurs des pré-requis et des domaines prioritaires du socle commun. Site Ministère de la santé française (URL : http://www.sante.gouv.fr/IMG/pdf/DGOS_Guide_d_indicateurs_Programme_Hopital_Numerique_-_avril_2012-2.pdf ; accès le 30/12/2013).
- [23] Creighton C. A literature review on communication between picture archiving and communication systems and radiology information systems and/or hospital information systems. J Digit Imaging. août 1999;12(3):138-143.
- [24] Rapport d'information n° 62 « L'informatisation dans le secteur de la santé » de M. Jean-Jacques Jégou, fait au nom de la commission des finances, déposé le 3 novembre 2005
- [25] Haug PJ, Pryor TA, Frederick PR. Integrating Radiology and Hospital Information Systems: the advantage of shared data. Proc Annu Symp Comput Appl Med Care. 1992;187-191.
- [26] Atlas 2013 des SIH – état des lieux des systèmes d'informations hospitaliers (mai 2013) / DGOS – ATIH (URL: http://www.sante.gouv.fr/IMG/pdf/Atlas_SIH-2.pdf, accès le 31/12/13).
- [27] Indicateurs IPAQSS – HAS: Conformité des Demandes des Examens d'Imagerie (URL: http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-07/ipaqss_fiche_cdei2011.pdf, accès le 31/12/13).
- [28] AHRQ Publication No. 12-E005-EF. Enabling Patient-Centered Care Through Health Information Technology, Rockville, MD: Agency for Healthcare Research and Quality; June 2012. (URL: <http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1158>; accès le 2/1/2014).
- [29] Nicolle C. Prescription connectée impact en biologie hospitalière. SPECTRA BIOLOGIE, 2007 Dec, 163, 54-59
- [30] Hawkins RC. Laboratory turnaround time. Clin Biochem Rev. 2007 Nov;28(4):179-94.
- [31] Breil B, Fritz F, Thiemann V, Dugas M. Mapping turnaround times (TAT) to a generic timeline: a systematic review of TAT definitions in clinical domains. BMC Med Inform Decis Mak. 2011;11:34
- [32] College of American Pathologists. Past Q-Probes Studies. (URL: http://www.cap.org/apps/cap.portal?nfpb=true&cntvwrPtl{actionOverride=%2Fportlets%2FcontentViewer%2Fshow&windowLabel=cntvwrPtl&cntvwrPtl{actionForm.contentReference}=q_probes%2Ftat.html&state=maximize_d&pageLabel=cntvwr}); accédé le 7 janvier 2014.

- [33] Steindel SJ, Howanitz PJ. Physician Satisfaction and Emergency Department Laboratory Test Turnaround Time. Observations Based on College of American Pathologists Q-Probes Studies. *Arch Pathol Lab Med* 2001;125:863-871
- [34] Roubille M et al. Étude sur les délais de rendu de résultats d'examens biologiques demandés en urgence dans les laboratoires hospitaliers. *Ann Biol Clin* 2010 ; 68 (6) : 741-6
- [35] Howanitz JH, Howanitz PJ. Laboratory Results, Timeliness as a Quality Attribute and Strategy. *Am J Clin Pathol* 2001;116:311-315
- [36] Vollmer RT. Analysis of turnaround times in pathology: an approach using failure time analysis. *Am J Clin Pathol* 2006;126:215-20
- [37] Westbrook JI, Georgiou A, Dimos A, Germanos T. Computerised pathology test order entry reduces laboratory turnaround times and influences tests ordered by hospital clinicians: a controlled before and after study. *J Clin Pathol.* 2006 May;59(5):533-6.
- [38] Jian WS, Hsu MH, Sukati H, Syed-Abdul S, Scholl J, Dube N, Hsu CK, Wu TJ, Lin V, Chi T, Chang P, Li YC. LabPush: a pilot study of providing remote clinics with laboratory results via short message service (SMS) in Swaziland, Africa. *PLoS One.* 2012;7(9):e44462.
- [39] Georgiou A, Williamson M, Westbrook JI, Ray S. The impact of computerized physician order entry systems on pathology services: a systematic review. *Int J Med Inform.* 2007 Jul;76(7):514-29.
- [40] Baron JM, Dighe AS. Computerized provider order entry in the clinicallaboratory. *J Pathol Inform.* 2011;2:35.
- [41] Guss DA, Chan TC, Killeen JP. The impact of a pneumatic tube and computerized physician order management on laboratory turnaround time. *Ann Emerg Med.* 2008Feb;51(2):181-5.
- [42] Creighton C. A literature review on communication between picture archiving and communication systems and radiology information systems and/or hospital information systems. *J Digit Imaging.* août 1999;12(3):138-143.
- [43] Rapport d'information n° 62 « L'informatisation dans le secteur de la santé » de M. Jean-Jacques Jégou, fait au nom de la commission des finances, déposé le 3 novembre 2005
- [44] Haug PJ, Pryor TA, Frederick PR. Integrating Radiology and Hospital Information Systems: the advantage of shared data. *Proc Annu Symp Comput Appl Med Care.* 1992;187-191.
- [45] Atlas 2013 des SIH – état des lieux des systèmes d'informations hospitaliers (mai 2013) / DGOS – ATIH (URL: http://www.sante.gouv.fr/IMG/pdf/Atlas_SIH-2.pdf, accès le 31/12/13).
- [46] Indicateurs IPAQSS – HAS: Conformité des Demandes des Examens d'Imagerie (URL: http://www.has-sante.fr/portail/upload/docs/application/pdf/2011-07/ipaqss_fiche_cdei2011.pdf, accès le 31/12/13).
- [47] Inamura K, Umeda T, Harauchi H, Kondoh H, Hasegawa T, Kozuka T, Takeda H, Inoue M. Time and flow study results before and after installation of a hospital information system and radiology information system and before clinical use of a picture archiving and communication system. *J Digit Imaging.* févr 1997;10(1):1-9.
- [48] Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, Morton SC, Shekelle PG. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med.* 16 mai 2006;144(10):742-752.
- [49] Georgiou A, Prgomet M, Markewycz A, Adams E, Westbrook JI. The impact of computerized provider order entry systems on medical-imaging services: a systematic review. *J Am Med Inform Assoc.* 5 janv 2011;18(3):335-340.
- [50] Chin HL, Wallace P. Embedding guidelines into direct physician order entry: simple methods, powerful results. *Proc AMIA Symp.* 1999;221-225.
- [51] Mariani C, Tronchi A, Oncini L, Pirani O, Murri R. Analysis of the X-ray work flow in two diagnostic imaging departments with and without a RIS/PACS system. *J Digit Imaging.* 2006;19 Suppl 1:18-28.

- [52] Hurlen P, Østbye T, Borthne A, Dahl FA, Gulbrandsen P. Do clinicians read our reports? Integrating the radiology information system with the electronic patient record: experiences from the first 2 years. *Eur Radiol.* janv 2009;19(1):31-36.
- [53] Weatherburn G, Bryan S, Nicholas A, Cocks R. The effect of a picture archiving and communications system (PACS) on diagnostic performance in the accident and emergency department. *J Accid Emerg Med.* mai 2000;17(3):180-184.
- [54] Weatherburn GC, Bryan S. The effect of a picture archiving and communication system (PACS) on patient radiation doses for examination of the lateral lumbar spine. *Br J Radiol.* juin 1999;72(858):534-545.
- [55] Bryan S, Weatherburn GC, Watkins JR, Buxton MJ. The benefits of hospital-wide picture archiving and communication systems: a survey of clinical users of radiology services. *Br J Radiol.* mai 1999;72(857):469-478.
- [56] Bryan S, Weatherburn G, Watkins J, Roddie M, Keen J, Muris N, Buxton MJ. Radiology report times: impact of picture archiving and communication systems. *AJR Am J Roentgenol.* mai 1998;170(5):1153-1159.
- [57] Ramaswamy MR, Chaljub G, Esch O, Fanning DD, vanSonnenberg E. Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact. *AJR Am J Roentgenol.* mars 2000;174(3):617-622.
- [58] Seltzer SE, Kelly P, Adams DF, Chiango BF, Viera MA, Fener E, et al. Expediting the turnaround of radiology reports in a teaching hospital setting. *AJR Am J Roentgenol.* avr 1997;168(4):889-893.
- [59] Lepanto L. Impact of electronic signature on radiology report turnaround time. *J Digit Imaging.* sept 2003;16(3):306-309.
- [60] Livrable du projet UPRES remis au Conseil Scientifique de l'ANAP, janvier 2014.

Adresse de correspondance :

Stéfan Darmoni, Equipe CISMef, Cour Leschevin, Porte 21, 3ème étage, 1 rue de Germont
76031 Rouen Cedex; Courriel : stefan.darmoni@chu-rouen.fr

Moteur de recherche sémantique au sein du dossier du patient informatisé : langage de requêtes spécifique

Romain Lelong¹ Tayeb Merabti¹ Julien Grosjean¹ Mher B. Joulakian¹ Nicolas Griffon^{1,2} Badisse Dahamna¹ Marc Cuggia³ Suzanne Pereira⁴ Natalia Grabar⁵ Frantz Thiessard⁶ Philippe Massari¹ Stefan J. Darmoni^{1,2}

¹Service d'Informatique Biomédicale, CHU de ROUEN, Haute-Normandie & TIBS, LITIS EA 4108, France

²LIMICS, INSERM, U1142, Paris, France.

³Inserm U936 – Université de Rennes 1, France

⁴Société Vidal, 92 Issy les Moulineaux, France

⁵STL UMR8163 CNRS, Université Lille 1&3, France

⁶LESIM Université Bordeaux II (Victor Segalen), France

Résumé

La recherche d'information (RI) au sein du Dossier du Patient Informatisé (DPI) doit fournir aux professionnels de santé la bonne information à la bonne personne, au bon moment, et au bon endroit, et de réduire les tâches lourdes de recherche d'information manuelle dans des dossiers papiers voire informatiques. Dans ce contexte, l'objectif de ce travail a été de décrire les fonctionnalités d'un moteur de recherche sémantique au sein d'un DPI. Dans ce papier, nous décrivons un langage de requête orienté objet, conçu pour la consultation de données, flexible et évolutif permettant de prendre en charge n'importe quelle modèle de données. Ce moteur permet des requêtes sur des données structurées et non structurées, sur un patient unique à visée soin ou N patients à visée épidémiologique. Nous avons testé différents types de requêtes sur une base de test de 2 000 patients anonymisés, contenant environ 200 000 comptes rendus.

Abstract

Information Retrieval (IR) in the Electronic Health Record (EHR) should provide healthcare professionals with the right information to the right person at the right time and place and should reduce the hard tasks of manual information retrieval from papers or from computer. In this context, the objective of this study was to describe the features of a semantic search engine implemented in an EHR. In this paper, we describe a flexible and scalable object-oriented query language designed for retrieving and viewing data which support any data model. This search engine deals with structured and unstructured data, on a unique patient in the context of care, and on N patients in the context of epidemiology. In this study, we tested several types of queries on a test databases containing 2,000 anonymized patients and about 200,000 records.

Mots-clés : Dossier du patient informatisé ; recherche d'information ; indexation automatique

Keywords: *Electronic Health Record; information retrieval; automatic indexing*

1 Introduction

Le Dossier du Patient Informatisé (DPI) est une « version informatisée du *dossier du patient papier* » [1]. Hebda and Czar [2] décrivent le DPI comme une ressource d'informations informatisées utilisées en santé pour capturer des données du patient. L'International Organization for Standardization (ISO) a défini le DPI comme « un outil de dépôt d'informations de santé dans une forme informatisable, archivée, et transmissible à des utilisateurs authentifiés ». Son objectif principal est de garantir un soin de qualité, efficace et intégré ; le DPI contient des informations à la fois rétrospectives, actuelles et prospectives [3], utiles à tous les professionnels de santé, avec des prescriptions, de la planification et des évaluations [4]. Ces informations permettent par exemple l'aide à la décision et ou la création de cohortes.

Dans ce contexte, l'objectif de la recherche d'information (RI) au sein du DPI est de fournir aux professionnels de santé la bonne information à la bonne personne, au bon moment et au bon endroit [5]. Dans la pratique, utiliser un outil de recherche dans le DPI doit permettre de réduire les tâches lourdes de recherche d'information manuelle dans des dossiers papiers voire informatiques, et par ce biais, réutiliser ce temps de professionnel de santé pour améliorer la qualité des soins. Pour González-González et al. [6], les professionnels de santé ont besoin de différentes informations et de connaissances pour réaliser leurs tâches :

- Information sur le patient (information sur l'œil du patient s'il est diabétique) ;
- Connaissances (par exemple, sur les recommandations sur une pathologie donnée).

Terry et al. [7] ont décrit cinq options de recherche d'information dans le DPI :

1. Requêtes prédéfinies : l'utilisateur choisit une requête dans un menu ;
2. Requêtes simples personnalisables : l'utilisateur peut écrire une requête simple pour obtenir des résultats souvent hétérogènes, mais pas nécessairement tous pertinents ;
3. Requêtes avancées personnalisables : l'utilisateur peut saisir une grande variété d'informations dans sa requête, le plus souvent séparés par des opérateurs Booléens (ET, OU, SAUF) ;
4. Interface de langage de requêtes structuré : utilisant une interface spécifique pour saisir les requêtes ;
5. Analyse et extraction d'information avec des outils de bases de données : ces outils de bases de données fournissent le plus haut niveau de possibilité pour réaliser des requêtes complexes.

Ces cinq niveaux sont différents en termes d'usage et de complexité. Les trois premiers niveaux apparaissent comme des solutions assez pauvres pour l'utilisation de recherche d'information dans le DPI, car les questions de recherche des professionnels de santé sont de plus en plus complexes.

Dans ce contexte, l'objectif de ce travail a été de décrire les fonctionnalités d'un moteur de recherche sémantique au sein d'un DPI. Le langage de requêtes est plus complexe que le mode d'interrogation classique de recherche d'information documentaire, du fait de l'existence de plusieurs niveaux hiérarchiques (patient, établissement, séjour), puis le niveau plus classique (actes, procédures, codage PSMI, examens biologiques, métadonnées d'un compte-rendu). Le modèle de données utilisé a permis de définir un langage proche de la représentation médicale de la prise en charge d'un patient. Toutes les informations contenues dans le DPI peuvent ainsi pouvoir être affichées à ces différents niveaux. Ce moteur de recherche est en cours de développement dans le cadre du projet RAVEL [8], financé par le programme TecSan de l'Agence Nationale de la Recherche (ANR).

2 État de l'art

Plusieurs outils et plateformes pour la recherche dans le DPI ont été proposés. Nous intégrons ici les systèmes orientés population fondés sur un entrepôt de données, et les systèmes de recherche d'information dans le dossier (mono)patient. Ces outils sont souvent différents suivant le type de données à rechercher : structurées ou non structurées [9].

Dans les systèmes de recherche d'information (SRI) dans le dossier (mono)patient, plusieurs outils ont été décrits dans la littérature : CISearch [10] est l'outil développé et implémenté au sein du DPI de l'hôpital universitaire de Columbia, aux Etats-Unis. Il permet à l'utilisateur d'effectuer des recherches dans l'ensemble des notes en texte libre (comptes-rendus de radiologie, d'anatomopathologie, résumés de sortie, notes de soins...) du dossier médical qu'il consulte. Il exploite quelques fonctionnalités de Lucene. MIRS (Medical Information Retrieval System) [11] est également fondé sur Lucene. Citons également le projet LERUDI [12] (dont l'équipe Rouennaise était partenaire sur les terminologies de santé) qui avait comme objectif la RI au sein du DMP (Dossier Médical Partagé) dans un contexte d'urgence, avec une approche sémantique et la création d'une ontologie de domaine.

Dans les entrepôts de données permettant la recherche d'information (multi)patients, I2B2 (Informatics for Integrating Biology and the Bedside) [13] est une plateforme open source développée aux Etats-Unis et dédiée à la recherche translationnelle. I2B2 est implémenté dans plusieurs pays, dans environ 70 CHU, et peut déjà être considéré comme un standard de facto. I2B2 stocke les données dans un entrepôt de données centré sur l'exploitation de données structurées. Un des composants les plus visibles d'I2B2 est un outil open source de sélection des patients appelé « i2b2 workbench », qui est un outil modulaire, facile à utiliser, et qui permet l'interrogation et la visualisation graphique des données cliniques [14]. La plateforme utilise des données biologiques et d'autres données génomiques (surcouche Transmart).

La fonctionnalité de recherche d'information est aussi proposée par le système de requêtes d'OpenEHR [15], qui repose sur un langage dédié, dépendant de sa structure orientée "archétypes". Spécialement conçu pour interroger ce type de modèle sur le Dossier Patient Informatisé, l'AQL (Archetype Query Language)¹ se veut un langage sémantique et indépendant du système. Stanford Translational Research Integrated Database Environment (STRIDE) propose un outil de requêtes nommé « Anonymous Patient Cohort Tool » dédié à la création de cohortes de patients [16]. Le moteur de recherche EMERSE [17] permet de rechercher des termes en plein texte avec des options avancées, adaptées au DPI (exemple : recherches avec troncatures, synonymie, etc.). XOntoRank [18] est un moteur de recherche permettant de faire une RI sémantique dans des documents médicaux structurés conformes au standard HL7 CDA. Ces documents ont la particularité de contenir à la fois des données codées, structurées et des données textuelles. L'outil utilisé comprend deux phases :

- Une indexation sémantique à l'aide de la nomenclature SNOMED
- Une phase de requête dans laquelle les concepts SNOMED des termes extraits de la requête utilisateur sont mappés avec la base de documents XML indexés.

Contrairement aux premiers outils de recherches appliquées sur des données structurées, ces types d'outils exploitent le contenu textuel des DPI. La recherche en utilisant ces outils peut se faire de différentes manières: une recherche plein texte [19] ou une recherche fondée sur les métadonnées décrivant la sémantique du contenu textuel [20], après utilisation d'outils de traitement automatique de langues (TAL).

¹ <http://www.openehr.org/wiki/display/spec/Archetype+Query+Language+Description>

Roogle [21] est une plateforme française du CHU de Rennes dédiée à la recherche d'information au sein de DPI. Cet outil a été à l'origine du projet RAVEL [8]. Cette plateforme est constituée d'un entrepôt de données stockant les données du DPI (données biologiques [LOINC], données d'actes [CCAM], des comptes-rendus médicaux de radiologie, d'anatomopathologie et des courriers de sortie) et d'un ensemble d'outils de RI combinant des méthodes de RI sémantique basées sur l'exploitation des métadonnées spécifiques aux documents (métadonnées issues des systèmes d'information clinique, métadonnées sur la structure logique du document), et des méthodes de RI plein texte exploitant le contenu textuel. A ce jour, Roogle permet une RI à la fois sur données non structurées et données structurées.

D'autres outils linguistiques permettent la recherche sur des données textuelles, par exemple : Currie et al. [22] proposent une approche linguistique (variation lexicale des termes, prise en compte du contexte) pour analyser les documents médicaux afin d'identifier les patients qui ont des problèmes de cœur et qui fument. Jain et al. [23] proposent une méthode d'expansion d'une requête en utilisant plusieurs sources de connaissances, incluant les relations sémantiques fondées sur des ontologies, des méthodes d'apprentissage supervisé des co-occurrences d'un terme. Plaza and Díaz [24] utilisent l'outil MetaMap pour exploiter les relations sémantiques d'UMLS afin de rechercher des cas similaires de DPI [25] pour bénéficier de la puissance du métathésaurus de l'Unified Medical Language System (UMLS) [26].

3 Matériel

3.1 Le modèle de données

L'utilisation d'un DPI est devenue une pratique courante dans tous les hôpitaux. Les modèles de ces DPI varient d'un hôpital à l'autre, mais ils sont le plus souvent complexes : à titre d'exemple, le DPI du CHU de Rouen contient plus de 100 tables. Le modèle de données de notre moteur de recherche décrit dans [9] est volontairement compact (onze tables) pour minimiser les temps de réponse de la RI (voir Figure 1).

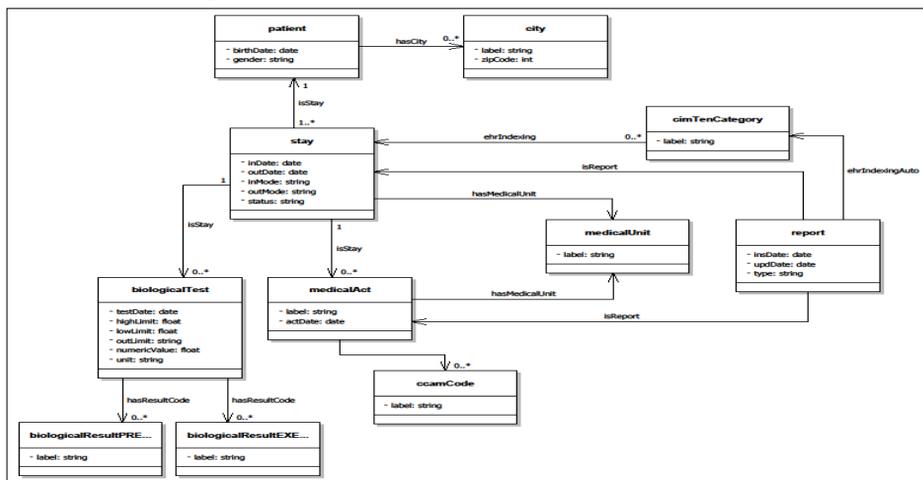


Figure 1 : Schéma du modèle de données du moteur de recherche

4 Méthode

4.1 Le langage de requête

4.1.1 Description :

L'intérêt de proposer un nouveau langage de requêtes est de faciliter la consultation et la recherche d'information dans le DPI et ainsi proposer une alternative aux langages de type SQL utilisés dans de nombreuses structures. Le principe est de rendre invisible la couche SQL afin de proposer une syntaxe la plus simple et intuitive et assez riche pour construire des requêtes complexes sans avoir besoin d'autres connaissances que la connaissance des « entités existantes » dans la base de données : patient, séjour...

Le langage de requêtes proposé dans ce travail à trois caractéristiques importantes :

- Il s'agit d'un langage de requêtes orienté objet avec un motif de syntaxe sur un modèle conceptuel de données. En revanche, il est conçu uniquement pour la consultation de données.
- Ce langage de requêtes est flexible et évolutif : il détecte automatiquement les entités conceptuelles de la base de données et donc il permet de prendre en compte automatiquement de nouvelles « entités », « attribut d'entités » ou « relations à d'autres entités » sans modification préalable du langage de la requête. Cette fonctionnalité importante nous a permis une extension aisée vers les données omiques (génomiques, métabolomiques, protéomiques, méthylation...) [27].
- Ce langage de requêtes a des capacités d'interrogation complète, c'est-à-dire que toutes les données incluses dans la base de données sont interrogeables. Il permet l'interrogation de données : symboliques (présence ou absence d'un diagnostic), numériques (comme les examens biologiques), impliquant de nouveaux opérateurs (>, <, =) par rapport aux opérateurs booléens plus classiques (ET, OU, SAUF) dans la recherche d'informations documentaires, et chronologiques (dernier examen échographique par exemple).

Ce langage permet l'affichage de toutes les informations contenues dans le DPI aux différents niveaux de celui-ci : patient, séjour ou le niveau le plus bas (actes, diagnostics, examens biologiques, dates, etc... qui est le niveau classique d'une RI documentaire).

4.1.2 Syntaxe :

Les principaux composants du langage sont les unités imbriquées de type :

ENTITE (CLAUSE de CONTRAINTES)

- **ENTITE** peut correspondre à n'importe quel type d'entité du modèle conceptuel de données (par exemple : patient, séjour, unité médicale...).
- **CLAUSE de CONTRAINTES** est une expression booléenne qui utilise des opérateurs booléens en combinaison avec des parenthèses pour lier logiquement les contraintes entre elles.

Par exemple, l'expression : patient(dateNaissancePatient='01/01/1937' ET sexe='M'), montre l'utilisation des attributs « **dateNaissancePatient** » et « **sexe** » de l'entité « **patient** », où les expressions symboliques comme **sexe='M'** et expressions temporelles **dateNaissancePatient='01/01/1937'** représentent les contraintes reliées entre elles avec l'opérateur booléen « **ET** ».

L'avantage d'une telle représentation est que les opérateurs booléens, les parenthèses, les comparateurs logiques sont les seules variables définies dans la grammaire du langage contrairement aux entités qui sont générées dans la grammaire suivant la base de données utilisée.

4.1.3 Les contraintes :

D'une manière plus simple, les contraintes unitaires sont l'expression d'une restriction d'un attribut direct de l'objet. Ainsi, dans le langage proposé, trois types de données sont traitées : symboliques ou textuelles, numériques, et temporelles, permettant une gestion chronologique (voir Tableau 1).

Tableau 1 : Exemples des types de données traitées dans les contraintes simples

Type de données	Exemple	Description
Symbolique	patient(sexe="M")	Patient de sexe masculin
Numérique	analyse(valeurNumericAnalyse >6 ET valeurNumericAnalyse <=6.25)	Test biologique dont la valeur est comprise entre 6 et 6.25
Date	sejour(dateEntreeSejour=2010-03-10)	Le séjour du 10/03/2010

D'un autre côté la puissance de ce langage de requêtes vient du fait que nous pouvons combiner plusieurs contraintes imbriquées à l'intérieur d'une même contrainte (voir Tableau 2). Pour les valeurs numériques, il est également possible d'exprimer un examen biologique en fonction d'une référence aux bornes inférieures et supérieures à la normale, présentes pour chaque analyse biologique : par exemple, rechercher pour un patient donné toutes les glycémies supérieures à 1,5 fois la normale (sous-entendu supérieur à 1,5 la borne supérieure). Notons ici que cette notion de valeur de la borne supérieure (ou inférieure) peut évoluer au cours du temps. Elle sera prise en compte par notre modèle de données et notre outil de recherche.

Tableau 2 : Exemples avancés sur l'utilisation des contraintes simples

Exemple	Description
Patient(analyse(codeEXEResultatBiologique (label="Phosphore") ET 0,81<valeurNumericAnalyse <= 1,17))	Les patients qui ont une analyse indexée par le terme EXE « Phosphore » et dont la valeur est comprise entre 0,81 et 1,17
<i>analyse(codeEXEResultatBiologique(label="Sodium") ET valeurNumeriqueAnalyse<borneInfAnalyse ET patient(id="DM_PAT_125"))</i>	Pour un patient donné (n° 125), affichez toutes les hyponatrémies

4.2 Le moteur de recherche

Par rapport à l'analyse de l'existant, nous avons imaginé notre outil de recherche à la fois comme le plus générique possible et avec une forte contrainte en termes de temps de réponse. Cet outil s'appuie sur nos travaux précédents sur la RI documentaire, où nous avons créé un moteur sémantique pour retrouver un type d'objets unique : les ressources Web[29]. La première phase de

la généricité avait permis la réalisation d'un moteur de recherche permettant de retrouver tout type d'objets (concept ontologique, article scientifique, ressource documentaire), mais toujours limité à un parcours de profondeur un. Dans le moteur actuel, cette limitation a été levée et nous pouvons désormais parcourir un arbre de profondeur N (voir Figure 3). Dans l'hypothèse où l'utilisateur n'aurait pas précisé le chemin à parcourir dans sa requête, un algorithme détermine le chemin optimal. Pour toutes ces raisons, nous avons élaboré un langage de requêtes complexe (niveaux 3 et 4 de Terry et al. [7]).

Ce moteur de recherche permet la recherche d'informations sur des données structurées (essentiellement numériques, mais aussi parfois symboliques, comme le sexe), mais aussi sur des données non structurées (issues essentiellement des différents comptes-rendus d'un DPI). Cette fonctionnalité importante a été rendue possible par l'intégration dans le moteur de recherche des outils de TAL développés par les équipes du Vidal (SP) et de Lille (NG). Pour chaque compte-rendu (CR), les outils TAL sont exécutés, aboutissant un ensemble de métadonnées détaillant les concepts médicaux reconnus en positif, en négatif ou en incertain ; ces outils TAL gérant la négation, souvent présente dans les CR (voir Figure 2). Les concepts peuvent être reconnus s'ils sont présents dans une ou plus des 55 terminologies et ontologies de santé du portail terminologique de santé [28]. Dans cet exemple (Figure 2), l'utilisateur choisit la recherche simple (niveau 2 de Terry et al. [7]) et saisit « gonalgie », qui est reconnu par trois terminologies (SNOMED International, SNOMED CT et MedDRA). Par défaut, la recherche s'effectue également sur le PMSI, mais comme ce concept n'existe ni dans la CIM10 ni dans CCAM, cette recherche s'effectue exclusivement dans les CR, indexées par les outils TAL. Le concept en rouge dans cet exemple « M65.9 - synovite et ténosynovite, sans précision » indique sa négation. A noter que l'outil RAVEL permet l'intégration de différents outils TAL : celui du projet RAVEL et celui développé au CHU de Rouen (ECMT dans le projet SYNODOS[30]).

The screenshot displays the RAVEL search interface. The main window shows a search for 'gonalgie' with a 'Recherche simple' button. Below, a 'Recherche avancée' section includes a list of search criteria and a table of results. A modal window titled 'Comptes-rendus' is open, showing details for a 'CR RAVEL' from 2000-01-01. The modal contains a list of annotations, including 'Liquide', 'articulation_sai', 'élément chimique', 'CORTEX/CL', 'ENERGEL', 'SEROPLEX', 'genou_droit', 'genou_sai', 'index_sai', 'M65.9 - synovite et ténosynovite, sans précision', and 'Vitesse de sédimentation'. The main window also shows a table of search results with columns for 'Unité médicale', 'Date', 'CR', and 'Patient'.

Unité médicale	Date	CR	Patient
Compte-rendu de séjour (ab)	2008-10-15 00:00:00	CR 1051	0143.072
Compte-rendu de séjour (ab)	2007-05-21 00:00:00	CR 1067	
Compte-rendu de séjour (ab)	2006-11-10 00:00:00	CR 1031	
Compte-rendu de séjour (ab)	2003-11-20 00:00:00	CR 1017	
Compte-rendu de séjour (ab)	2003-08-22 00:00:00	CR 1063	
Compte-rendu de séjour v1.	2001-02-28 00:00:00	CR 1017	
CR RAVEL	2000-01-01 00:00:00	CR 104	
CR RAVEL	2000-01-01 00:00:00	CR 104	
CR RAVEL	2000-01-01 00:00:00	CR 104	
CR RAVEL	2000-01-01 00:00:00	CR 105	
Courrier type par défaut	1997-06-01 00:00:00	CR 1017	

Figure 2: Copie d'écran de l'intégration des outils TAL dans le moteur de recherche RAVEL

A l'inverse, la recherche simple sur « infarctus du myocarde » s'effectuera à la fois sur les données structurées (ici, au sein du PMSI, dans les diagnostics de la CIM10 car l'outil reconnaît une maladie) et sur les données non structurées (au sein des différents CR du patient). Comme certaines expressions dans les CR ne sont pas reconnues comme des concepts médicaux de notre portail terminologique, nous avons également développé une recherche en texte intégral pour pallier ce manque : dans la recherche simple, une requête sera d'abord analysée pour rechercher les concepts médicaux issus du portail, ensuite seulement les expressions non reconnues seront recherchées en texte intégral. Dans notre langage de requêtes, la recherche en texte intégral s'écrit comme tel : `compteRendu(FILE.f_html()="expression")`,

Notre moteur est générique, car il permet une recherche d'information dans le DPI, au niveau d'un patient unique, essentiellement à visée « soin », de prise en charge effective de ce patient (le projet RAVEL se limite à ce cas d'usage), mais aussi au niveau de plusieurs patients (potentiellement tous les patients d'un établissement). Dans ce dernier cas, les objectifs sont plus variés : épidémiologie (création de cohortes par exemple), recherche clinique (détection de critères d'inclusion d'une étude), mais aussi calculs d'indicateurs de qualité (pour le circuit de biologie, temps global entre la prescription d'un examen et son retour au lit du malade).

Pour permettre la recherche à l'intérieur du DPI, nous avons développé un moteur de recherche permettant une interprétation des requêtes pour extraire les données correspondantes. Ainsi, le processus global du traitement d'une requête en langue naturelle par le moteur peut être divisé en trois phases :

4.2.1 Le parseur de requête

La première étape consiste à une analyse syntaxique de la requête en utilisant la grammaire définie par le langage de requêtes. Ainsi, cette étape permet de valider la requête (conforme ou non à la grammaire) et de déviser la requête en plusieurs « unités lexicales et syntaxiques » élémentaires.

4.2.2 La représentation en arbre de requête

L'objectif de cette étape est d'extraire les éléments significatifs de la requête. Dans cette phase, une représentation sous forme d'arbre de la requête est générée par l'analyse de sa structure hiérarchique (voir *Figure 3*). Pour cause, la structure de l'arbre est bien adaptée pour maintenir une bonne représentation de la structure logique de la requête surtout pour les éléments booléens. Cette structuration est également conforme au raisonnement du professionnel de santé.

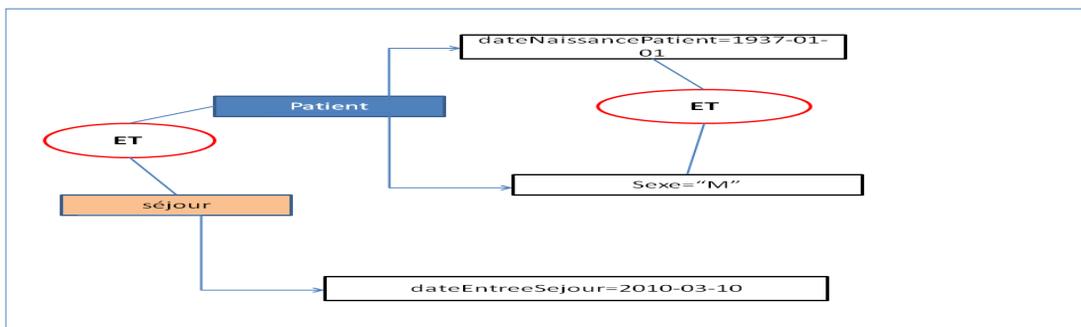


Figure 3 : Représentation en arbre de la requête
 « `patient(dateNaissancePatient='01/01/1937' ET Sexe='M') AND`
`sejour(dateEntreeSejour = 2010-03-10)` »

4.2.3 La construction SQL de la requête

L'arbre construit dans l'étape précédente est une représentation complète de tous les aspects de la requête et, par conséquent, peut être utilisé pour générer de manière récursive la requête SQL appropriée qui correspond à la requête demandée.

5 Résultats

En amont du projet RAVEL, nous avons construit au CHU de Rouen une base de test de 2 000 patients anonymisés, contenant environ 200 000 CR pour le projet LERUDI, financé par l'ASIP Santé. Toutes les requêtes présentées dans ce travail ont été testées sur cette base, en attendant l'intégration des données du CHU de Bordeaux. Le Tableau 3 affiche les différentes requêtes gérées par le moteur de recherche du projet RAVEL.

Tableau 3 : Exemple de requêtes du moteur de recherche RAVEL

Exemple	Description
<code>sejour(patient(id="DM_PAT_21") AND acte(label="PRELEVEMENT SANGUIN"))</code>	Les séjours du patient DM_PAT_21 dans lesquels un acte de prélèvement sanguin à eu lieu
<code>uniteMedicale(sejour(patient(id="DM_PAT_21") AND acte(label="PRELEVEMENT SANGUIN")))</code>	Les unités médicales des séjours du patient DM_PAT_21 dans lesquels ont eu lieu les actes de prélèvement sanguin
<code>analyse(patient(id="DM_PAT_1078") AND prestationCodeResultatBiologique(label="Plaquettes") AND 10*valeurNumeriqueAnalyse<borneInfAnalyse)</code>	Les analyses de plaquettes du patient DM_PAT_1078 dont le résultat est plus de 10 fois inférieures à la normale
<code>acte(codeCCAM(id="CCA_AM_EQQM006") AND dateActe="MAX")</code>	Le dernier acte de code CCAM AM_EQQM006

Nous avons ensuite tenté de répondre au cas d'utilisation définis par notre collègue de Bordeaux (FT). Par exemple, nous sommes capables de répondre aux problèmes suivants des professionnels de santé :

- Fournir au cours du temps une visualisation globale du taux de polynucléaires d'un patient donné porteur d'une polyarthrite rhumatoïde ;
- Fournir tous les comptes-rendus d'imagerie qui affichent la « notion de métastases », avec la requête en texte intégral suivante : `compteRendu(FILE.f_html()="notion de métastases")`, puisque cette expression « notion de métastases » n'est pas un concept reconnu par notre portail terminologique, alors que le terme « métastase » aurait bien entendu été reconnu.

Nous avons signalé que notre outil de recherche savait gérer des requêtes pour un patient donné et aussi pour plusieurs patients (voir *Figure 4*).

Utilisez Ctrl-Espace pour voir les propositions de mots réservés

1 acte(patient(id="001_Pat_43") AND label="CHOLECYSTECTOMIE")

OK

1 entrées trouvées (0,07 s)

Items per page: 20 Page: 1 / 1

Type d'acte	Date d'acte	Unité médicale	CCAM	Patient
CHOLECYSTECTOMIE	2002/01/21	Chirurgie Digestive		43

Utilisez Ctrl-Espace pour voir les propositions de mots réservés

1 acte(label="CHOLECYSTECTOMIE")

OK

26 entrées trouvées (0,28 s)

Items per page: 20 Page: 1 / 2

Type d'acte	Date d'acte	Unité médicale	CCAM	Patient
CHOLECYSTECTOMIE	2005/08/08	Chirurgie Digestive		821
CHOLECYSTECTOMIE	2005/07/01	Chirurgie Digestive		1848
CHOLECYSTECTOMIE	2005/06/06	Chirurgie Digestive		1781
CHOLECYSTECTOMIE	2005/04/08	Chirurgie Digestive		1815
CHOLECYSTECTOMIE	2004/11/18	Chirurgie Digestive		566
CHOLECYSTECTOMIE	2004/10/22	Chirurgie Digestive		837
CHOLECYSTECTOMIE	2003/06/09	Chirurgie Digestive		1100
CHOLECYSTECTOMIE	2002/12/24	Chirurgie Digestive		1199
CHOLECYSTECTOMIE	2002/07/18	Chirurgie Digestive		228
CHOLECYSTECTOMIE	2002/04/26	Chirurgie Digestive		467
CHOLECYSTECTOMIE	2002/01/21	Chirurgie Digestive		43
CHOLECYSTECTOMIE	2000/10/12	Chirurgie Digestive		1567
CHOLECYSTECTOMIE	1999/08/23	Chirurgie Digestive		722
CHOLECYSTECTOMIE	1999/06/14	Chirurgie Digestive		1361
CHOLECYSTECTOMIE	1999/05/24	Chirurgie Digestive		80
CHOLECYSTECTOMIE	1996/08/09	Chirurgie Digestive		799
CHOLECYSTECTOMIE	1994/03/17	Chirurgie Digestive		1311
CHOLECYSTECTOMIE	1993/09/20	Chirurgie Digestive		1255
CHOLECYSTECTOMIE	1992/08/11	Chirurgie Digestive		1155
CHOLECYSTECTOMIE	1992/01/30	Chirurgie Digestive		804

Figure 4 : Exemple d'une requête similaire pour les deux niveaux

En termes de temps de réponse, les deux contextes sont très différents. Pour un patient unique, nous avons fixé le maximum acceptable pour un professionnel de santé toujours pressé à deux secondes (il s'agit d'un temps serveur qui ne tient pas en compte le temps d'affichage qui varie selon la qualité de l'ordinateur et du réseau). Pour toutes les requêtes testées sur notre échantillon de 2 000 patients, ce temps a toujours été inférieur à ce seuil. En revanche, le temps d'affichage des examens biologiques reste encore trop long (de l'ordre de 10 secondes), mais il s'agit alors d'un problème de visualisation et non de recherche d'information proprement dit.

Voici également comment s'écrivent deux requêtes pourtant sur la même notion, mais s'affichant à deux niveaux (patient et séjour). Dans notre exemple, il s'agit de rechercher

- soit le dernier séjour dans le département d'imagerie médicale ; la requête est alors : `sejour(uniteMedicale(label="Imagerie Médicale") AND dateEntreeSejour="MAX")` ;
- soit les patients ayant eu un dernier séjour dans le département d'imagerie médicale ; la requête est alors : `patient(sejour(uniteMedicale(label="Imagerie Médicale") AND dateEntreeSejour="MAX"))`

Le sens médical de ces deux requêtes est très différent, et sera sans doute utilisé plutôt dans le service d'imagerie ou plutôt dans les services de soins.

6 Discussion

Nous avons présenté dans ce travail les principales fonctionnalités du moteur de recherche du projet RAVEL, en insistant sur son langage de requêtes que nous jugeons générique et puissant. Notre approche est générique, permettant une recherche d'information sémantique multi-terminologique, à la fois sur un patient unique à visée soin et sur plusieurs patients à visée épidémiologiques. Cet outil de recherche permet également des requêtes à plusieurs niveaux du

DPI : patient, séjour, et niveau le plus bas, c'est-à-dire toutes les données du DPI (actes, prescription, PMSI, biologie, compte-rendus). Une recherche en texte intégrale a été ajoutée également. Enfin, nous avons étendu notre approche aux données omiques.

Néanmoins, nous envisageons, suite au projet, d'effectuer une évaluation comparative de l'expressivité et de précision du langage avec le standard de facto qu'est I2B2 [13], [14] et Transmart, sa surcouche omique. I2B2 étant installé dans plusieurs hôpitaux français, une comparaison de la puissance respective des langages de requêtes de ces deux outils pourra ainsi être menée. Dans tous les cas de figure, un parseur pourrait également être développé entre le modèle de notre outil [8] et celui d'I2B2, permettant aux données de basculer entre les deux outils pour unir si nécessaire les résultats des deux langages de requêtes.

Limites :

A ce jour, le langage de requêtes permet de décrire les niveaux de 2 à 5 de Terry et al [7]. Nous n'avons pas imaginé jusqu'à présent de niveau 1, avec une liste prédéfinie de requêtes, mais nous le discuterons avec les différents professionnels de santé du projet RAVEL. Le langage de requêtes est complexe à manipuler de notre point de vue. Néanmoins, nous l'avons démontré à plusieurs médecins de santé publique, qui disent être favorables à son usage après formation rudimentaire. Les documentalistes, les professionnels des sciences de l'information, les informaticiens sont également à même de l'utiliser dans les mêmes conditions. Quant au professionnel de santé, il est impératif de prévoir plusieurs interfaces, en dehors de la recherche simple : une des pistes d'inspiration pourrait être l'interface d'I2B2.

Cet outil de recherche ne permet pas de répondre à toutes les questions des cas d'usage établis dans le projet RAVEL. Par exemple, à ce jour, nous ne savons pas répondre à la question : « faire le comptage des petites articulations touchées lors d'une polyarthrite rhumatoïde (c'est-à-dire rouges ou gonflées) ».

Notre modèle d'information ne permet pas encore de gérer le niveau « établissement », dans la mesure où ne gérons pour l'instant que des données d'un seul établissement de santé.

En perspective, et en extension du projet RAVEL au CHU de Rouen, nous allons mener une étude de passage à l'échelle de notre outil de recherche sur les 65 000 patients de cet établissement passés dans le service de dermatologie depuis 1992. L'objectif sera de tester cet outil dans un contexte multi-patients pour créer une cohorte.

7 Conclusion

Nous avons présenté dans ce travail l'outil de recherche pour retrouver une information de santé au sein du dossier du patient informatisé, en insistant sur sa généricité et son langage de requêtes puissant mais difficile à manier pour un professionnel de santé.

Remerciements

Ce travail a été en partie financé par l'Agence Nationale de la Recherche (ANR) et la Direction Générale de l'Armement (DGA), sous le numéro Tecsan ANR-11-TECS-012.

Références

- [1] Sewell J. *Thede: Informatics and Nursing: Opportunities and Challenges*. 3rd ed. Philadelphia, PA: Lippincott, Williams, & Wilkins; 2013.
- [2] Hebda T, Czar P. *Handbook of Informatics for Nurses and Healthcare Professionals*. 5th ed.

- Boston, MA: Pearson; 2012.
- [3] International Organisation for Standardisation (ISO). 20514 Draft Technical Report: EHR Definition, Scope and Context.
 - [4] Garde T, Knaup T, Hovenga E, Herd S: Towards semantic interoperability for electronic health records. *Methods Inf. Med.* 2007; 46 (3): 332–343.
 - [5] Ondo K, Wagner J, Gale K. The electronic medical record: Hype or reality? *Journal of Healthcare Information Management*, 2002; 17(4):2.
 - [6] González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, Hernández-Fernández T. Information needs and information-seeking behavior of primary care physicians. *Ann Fam Med.* 2007; 5: 345-352.
 - [7] Terry AL, Chevendra V, Thind A, Stewart M, Marshall JN, Cejic C. Using your electronic medical record for research: a primer for avoiding pitfalls. *Fam Pract.* 2010;27(1):121-6.
 - [8] Thiessard F, Mouglin F, Diallo G, Jouhet V, Cossin S, Garcelon N, Campillo B, Jouini W, Grosjean J, Massari P, Griffon N, Dupuch M, Tayalati F, Dugas E, Balvet A, Grabar N, Pereira S, Frandji B, Darmoni S, Cuggia M. RAVEL: retrieval and visualization in ELeCtronic health records. *Stud Health Technol Inform.* 2012;180:194-8.
 - [9] Dirieh Dibad AD. *Recherche d'Information Multi Terminologique au sein d'un Dossier Patient Informatisé*. Thèse de doctorat. Université de Rouen, 2012.
 - [10] Natarajan K, Stein D, Jain S, Elhadad N. An analysis of clinical queries in an electronic health record search utility. *Int J M Inform.* 2010 Jul;79(7):515-22.
 - [11] Spat S, Cadonna B, Rakovac I, Gütl C, Leitner H, Stark G, Beck P. Enhanced information retrieval from narrative German-language clinical text documents using automated document classification. *Stud Health Technol Inform.* 2008;136:473-8.
 - [12] Charlet J, Declerck G, Dhombres F, Gayet P, Miroux P, Vandenbussche PY. Construire une ontologie médicale pour la recherche d'information : problématiques terminologiques et de modélisation. 23èmes journées francophones d'Ingénierie des connaissances, Paris : France, 2012.
 - [13] Informatics for Integrating Biology and the Bedside. 2008 [<https://www.i2b2.org>].
 - [14] Deshmukh V, Meystre S, Mitchell J. Evaluating the i2b2 system for clinical research. *BMC Medical Research Methodology.* 2009;9(1):70.
 - [15] Kalra D, Beale T, Heard S. The openEHR foundation. *Studies in health technology and informatics.* 2005; 115, 153-173.
 - [16] Lowe H, Ferris, T, Hernandez P, Weber S. Stride—an integrated standards-based translational research informatics platform. In: *Proceedings of the AMIA Annual Symposium.* 2009;391.
 - [17] Hanauer D.A. EMERSE: The Electronic Medical Record Search Engine. *AMIA Annu Symp Proc.* 2006; 2006: 941.
 - [18] Farfan F, Hristidis V, Ranganathan A, Weiner M. Xontorank: Ontology aware search of electronic medical records. In *Proceedings of the 25th International Conference on Data Engineering*, 2009; 820–831. IEEE.
 - [19] Cuggia M, Bayat S, Garcelon N, Sanders L, Rouget F, Coursin A, Pladys P. A full-text information retrieval system for an epidemiological registry. *Studies In Health Technology And Informatics.* 2010. 160(Pt 1):491–495.
 - [20] Chung J, Murphy S. Concept-Value Pair Extraction from Semi-Structured Clinical Narrative: A Case Study Using Echocardiogram Reports. In *Proceedings of the AMIA Annual*

Symposium. 2005; 131–135.

- [21] Cuggia, M., Garcelon, N., Campillo-Gimenez, B., Bernicot, T., Laurent, J. F., Garin, E., Happe, A., Duvauferrier, R. (2011). Roogle: an information retrieval engine for clinical data warehouse. *Stud Health Technol Inform*, 169, 584-588.
- [22] Currie, A., Cohan, J., and Zlatic, L. Information retrieval of electronic medical records. *Computational Linguistics and Intelligent Text Processing*, 2001;460–471.
- [23] Jain H, Thao C, Zhao H. Enhancing electronic medical record retrieval through semantic query expansion. *Information Systems and E-Business Management*. 2010; 1–17.
- [24] Plaza L, Díaz A. Retrieval of similar electronic health records using UMLS concept graphs. *Natural Language Processing and Information Systems*. 2010; 296–303.
- [25] Aronson A.R. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. *Proc. AMIA Symp*. 2001 ;17-21.
- [26] Lindberg DAB, Humphreys BL, McCray AT. The Unified Medical Language System. *Meth Inform Med*. 1993; 32:281–91.
- [27] Cabot C, Grosjean J, Lelong R, Lefebvre A, Lecroq T, Soualmia LF, Darmoni, SJ. Omic Data Modelling for Information Retrieval. *Proceedings of the 2nd International Work-Conference on Bioinformatics and Biomedical Engineering, IWBBIO*. 2014 (in press).
- [28] Grosjean J, Merabti T, Griffon N, Dahamna B, Darmoni SJ. Teaching medicine with a terminology/ontology portal. *Stud Health Technol Inform*. 2012;180:949-53.
- [29] Darmoni SJ, Thirion B, Leroy JP, Douyère M, Lacoste B, Godard C, Rigolle I, Brisou M, Videau S, Goupy E, Piot J, Quéré M, Ouazir S, Abdulrab H. Doc'CISMEF: a search tool based on "encapsulated" MeSH thesaurus. *Stud Health Technol Inform*. 2001;84(Pt 1):314-8.
- [30] Dupuch M, Segond F, Bittar A, Dini L, Soualmia LF, Darmoni SJ, Gicquel Q, Metzger MH. Separate the grain from the chaff: make the best use of language and knowledge technologies to model textual medical data extracted from electronic health records. *Proceedings of the LTC'13 6th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*. 2013 (in press.).

Adresse de correspondance

Stéfan Darmoni, Service d'Informatique Biomédicale, Cour Leschevin, Porte 21, 3ème étage, 1 rue de Germont 76031 Rouen Cedex; Courriel : stefan.darmoni@chu-rouen.fr

Usage secondaire du dossier médical informatisé à des fins épidémiologiques et d'évaluation de la qualité des soins : le projet SYNODOS

A secondary use of the electronic health record for epidemiological studies and evaluation of health care quality: the SYNODOS project

Marie-Hélène Metzger^{1,2}, Quentin Gicquel¹, Domoina Rabarijaona³,
Frédérique Segond³, André Bittar⁴, Luca Dini⁴, Lina F. Soualmia⁵, Stéfan J.
Darmoni⁵

¹Université Lyon I – CNRS-UMR 5558 / LBBE, Villeurbanne, France

²Hôpital de la Croix-Rousse, UHE, Lyon, France

³Objet Direct, Grenoble, France

⁴Holmes Semantic Solutions, Grenoble, France

⁵CISMef, LITIS EA 4108 - Université de Rouen, France

Résumé

Contexte : L'objectif du projet SYNODOS est de développer une solution générique permettant l'extraction sémantique de données du dossier patient informatisé et d'organiser cette information médicale de manière à ce qu'elle puisse être utilisée à des fins d'études épidémiologiques ou d'évaluation de la qualité des soins.

Méthodes : SYNODOS propose une architecture modulaire qui fait une distinction claire entre les règles linguistiques et les règles du système expert. Les différents modules sont en cours de développement : une interface entre le serveur multi-terminologique et l'analyseur sémantique au cours de la phase d'extraction, règles linguistiques pour extraire les expressions temporelles, une interface entre le moteur linguistique et la base de connaissances.

Résultats: Les résultats du projet seront un système opérationnel intégrant les différents modules technologiques décrits ci-dessus. Le projet évaluera également la qualité des informations extraites dans deux domaines d'application : les infections associées aux soins et le cancer. Conclusion : Nous avons décrit un projet dont l'originalité réside dans le développement d'une solution unique intégrant différentes technologies nécessaires à la production d'indicateurs épidémiologiques dans le contexte de l'activité hospitalière.

Abstract

Background: The purpose of the SYNODOS project is to develop a generic solution for extracting semantics out of medical data and organize this medical information in such a way that it could be used to support epidemiological studies or evaluate healthcare quality.

Methods: SYNODOS proposes a general modular architecture that makes a clear distinction between linguistic rules and expert system rules. An interface between the semantic analyzer and multi-terminology server upstream during the extraction phase, linguistic rules to extract temporal expressions as well as an interface between linguistic engine and knowledge representation are being developed.

Results: Project outcomes will be an operational system integrating the various technological modules described above. The project will also evaluate the quality of extracted information in two domains: hospital acquired infections and cancer.

Conclusion: We have described a project whose originality lies in the integration into a single solution of the various technologies needed to produce epidemiological indicators in the context of hospital activity.

Mots-clés : Épidémiologie; Systèmes d'aide à la décision clinique ; Systèmes informatisés de dossiers médicaux ; Évaluation de programme.

Keywords: Epidemiology; Decision support systems, clinical; Medical records systems, computerized; Program evaluation.

1 Introduction

Le dossier patient informatisé (DPI) constitue une source potentielle très importante de données dans des domaines aussi variés que l'aide à la décision médicale, la médecine factuelle, la surveillance épidémiologique ou la fouille de données. Toutefois très peu de données sont structurées et codées dans les DPI pour permettre ce type d'utilisation. Á l'heure actuelle, seules les données nécessaires à la tarification à l'activité sont systématiquement structurées et codées dans les dossiers médicaux hospitaliers. Même si d'importants efforts ont été réalisés ces dernières années pour permettre l'utilisation épidémiologique des données du Programme de Médicalisation des Systèmes d'Information (PMSI), notamment en cancérologie, différents travaux montrent les difficultés liées à l'utilisation de ces données médico-économiques à des fins épidémiologiques [1]. Des travaux plus récents tentent de combiner ces données à d'autres sources de données structurées telles que les codes anatomo-pathologiques ADICAP [2]. Toutefois, la plupart des données du DPI sont disponibles dans des documents non structurés en langage naturel, ce qui nécessite la mise au point d'outils de traitement de cette information. Le projet SYNODOS vise à développer une solution de traitement du langage naturel des dossiers médicaux à des fins d'utilisation épidémiologique et d'évaluation de la qualité des soins.

2 Etat de l'art

De nombreux travaux sont en cours pour développer des méthodes de phénotypage à partir d'une utilisation secondaire des données du DPI. Ces méthodes intègrent notamment le traitement du langage naturel (ex : projet eMERGE [3]). Il n'existe pas à l'heure actuelle de langage formel ni d'approche standardisé pour extraire ces phénotypes [3]. Un modèle a été développé aux USA (National Quality Forum's Quality Data Model) qui donne une structure décrivant les concepts

cliniques dans un format standardisé pour la production automatique d'indicateurs de qualité. Ce modèle a été évalué et est utilisé dans le projet eMERGE [4] ou encore par le Consortium SHARPN qui développe une plateforme d'usage secondaire des données du DPI, reposant sur l'utilisation de diverses ressources (terminologiques, modèles de représentation) [5]. Dans le cadre du projet SYNODOS, nous avons développé un modèle conceptuel de données à visée générique et dont l'évaluation est prévue à l'issue du développement de la solution [6]. Par ailleurs une fois le phénotype obtenu par ces techniques, la 2^{ème} étape consiste à exploiter les données extraites pour par exemple, mesurer des associations ou sélectionner des patients éligibles. Il n'existe pas à notre connaissance de solution de langue française intégrant l'ensemble de ces étapes dans une solution unique et conçue pour être utilisée en production hospitalière. Le challenge du projet SYNODOS est de réussir à rassembler toutes les technologies nécessaires à l'exploitation de ces données pour divers usages en milieu hospitalier. Il s'agit donc d'un projet de recherche très appliqué.

L'objectif de cet article est de décrire le processus général de développement de la solution informatique SYNODOS.

3 Matériel et méthodes

Le projet SYNODOS (URL : <http://www.synodos.fr>) réunit 2 laboratoires publics de recherche, l'une experte dans le domaine de la recherche en Informatique Médicale (CISMeF) et l'autre dans le domaine de l'épidémiologie (LBBE), et 2 industriels, l'un spécialisé dans le développement de logiciels et ressources linguistiques (Holmes Semantic Solutions) et l'autre dans l'intégration de solutions de Business Intelligence et technologies web (Objet Direct). Le projet financé par l'Agence Nationale de Recherche (programme TecSan 2012) a été décomposé en 6 tâches scientifiques : 1) définition de l'architecture générale de la solution 2) développement du traitement sémantique des données textuelles médicales 3) interfaçage de l'extracteur de concepts multi-terminologique avec l'analyseur sémantique 4) développement d'un système générateur de règles expertes 5) intégration des différents modules constitutifs de la solution et enfin 6) évaluation des performances de la solution.

Le projet a démarré en octobre 2012 et se déroulera jusqu'à septembre 2015.

3.1 Architecture générale de la solution

La solution SYNODOS est une application web (SYNODOS – Médiateur) qui utilise les services de 2 serveurs distants :

- le serveur terminologique du CISMeF, existant, qui sera enrichi de nouvelles terminologies liées aux besoins spécifiques du projet
- le serveur sémantique, spécifique à SYNODOS, reposant sur les technologies de Holmes Semantic Solutions.

L'accès aux serveurs distants se fera par un « protocole de transfert hypertexte sécurisé » (HTTPS) de telle sorte qu'aucune donnée ne soit transmise en clair. Toutes les fonctionnalités de SYNODOS seront accessibles depuis le médiateur qui fournira l'interface unique de l'application. Différents logiciels seront utilisés pour le développement des modules composant la solution SYNODOS. Le projet s'appuiera sur : un langage de développement (Java), un framework web, un serveur d'applications, un SGBD relationnel, éventuellement une base NoSQL, un moteur BRMS (Business Rules Management System - Drools [7]) et des outils du web sémantique. La liste exacte des composants logiciels n'est pas finalisée, faisant partie du travail de développement en cours.

3.2 L'analyseur sémantique

L'analyseur sémantique est une plateforme de traitement du langage naturel développé par Holmes Semantic Solutions qui combine différents types de modules, appliqués de façon incrémentale dans le processus de traitement : modules symboliques (à base de règles), modules statistiques, et modules à base d'apprentissage automatique.

3.3 L'extracteur de concepts multi-terminologique

Le portail multi-terminologique développé par l'équipe CISMef contient 55 terminologies ou ontologies médicales, correspondant à 500,000 concepts médicaux en langue française et 1,5 million en langue anglaise (URL : <http://www.hetop.eu>). A partir de ce portail, un extracteur de concepts multi-terminologique (ECMT V2) a été développé pour indexer les termes médicaux rencontrés dans les documents textuels. Ce traitement est disponible en service web (SOAP ou REST).

Différentes terminologies médicales ont été sélectionnées pour normaliser le langage médical naturel dans le cadre du projet : la CIM-10 (Classification Internationale des Maladies, 10^{ème} revision), le thesaurus MeSH® (Medical Subject Headings) pour l'indexation des concepts diagnostiques ou symptomatologiques, la classification ATC (Anatomical Therapeutic and Chemical Classification) pour le codage des médicaments, la SNOMED International etc. Différentes techniques de mises en correspondance ont été employées dans ce projet : (a) mise en correspondance conceptuelle basée sur le CUI du Metathesaurus® de l'UMLS (Unified Modelling Language System); (b) mise en correspondance basée sur le traitement du langage naturel; (c) mise en correspondance manuelle ou supervisée par les experts terminologiques de l'équipe CISMef.

3.4 Développement d'un système générateur de règles expertes

Pour la finalité de la solution SYNODOS qui consiste à exploiter les données extraites des documents textuels à des fins épidémiologiques ou d'aide à la décision, une simple normalisation des concepts médicaux n'est pas suffisante. En effet, pour une utilisation pertinente des concepts extraits, il est nécessaire d'y associer un étiquetage temporel (s'agit-il d'une donnée reliée dans le parcours du patient à un antécédent, un motif d'entrée, une évolution en cours d'hospitalisation ?). Il est également nécessaire de créer des relations permettant de lier les concepts entre eux : par exemple dans la phrase suivante : « on a assisté à une chute de la gastrinémie », il est indispensable de relier le concept « gastrinémie » à celui de « chute » pour exploiter correctement cette information. Il faudra donc tout d'abord créer une règle qui permet d'affecter un concept du type « gastrinémie » à l'élément « type d'examen biologique » et un concept de type « chute » à l'élément « résultat d'examen biologique » de la base de faits. Ces règles s'appuieront sur l'utilisation des données sémantiques (par exemple négation du type « absence » associé au concept médical) et terminologiques (par exemple le type sémantique UMLS du concept) rattachées à chacun de ces concepts. La relation entre ces 2 concepts sera ensuite établie par l'utilisation des relations construites dans la base de connaissances (relations hiérarchiques ou transversales). Par ailleurs de nouveaux faits seront inférés par combinaison de « faits certains ». Par exemple, la présence d'un écoulement purulent de cicatrice après intervention chirurgicale, permettra d'inférer le fait qu'un patient est porteur d'une infection nosocomiale alors que ce dernier concept n'est pas mentionné dans le DPI. Des relations entre les différents concepts médicaux extraits, entre étiquettes temporelles ou relations sémantiques établies lors du traitement sémantique vont ainsi devoir être construites via un système expert développé par Objet Direct.

Le système expert sera composé d'une base de faits comprenant :

- les faits « certains », provenant du traitement terminologique et sémantique mais également des métadonnées récupérées dans le système hospitalier (ex : âge du patient, spécialité du service, type de compte-rendu traité,...)
- Les faits « inférés », c'est-à-dire générés par le raisonnement effectué par le moteur d'inférence. Pour effectuer ce raisonnement, le moteur d'inférence s'appuiera sur une base de règles. La base de règles sera élaborée par le LBBE et consistera en règles probabilistes, produites par technique de fouille de données et par règles heuristiques.

3.5 Apprentissage et évaluation des performances de la solution SYNODOS

Deux corpus de données textuelles sont utilisés dans le cadre du développement de la solution informatique SYNODOS. Le 1^{er} corpus a été constitué dans le cadre d'un précédent projet de recherche ALADIN-DTH [8] visant à développer un outil de repérage sémantique des infections nosocomiales dans les documents textuels médicaux. Dans le cadre de ce projet, 1,600 documents médicaux textuels composés de comptes rendus d'hospitalisation, lettres de sortie (et comptes rendus opératoires en chirurgie) ont été annotés par des médecins enquêteurs. Cette étude a été réalisée dans 4 centres hospitaliers universitaires français (Lille, Lyon, Nice, Rouen) sur la période allant de 2009 à 2011. Les dossiers médicaux des services de spécialités suivantes avaient été inclus : chirurgie digestive, neurochirurgie, chirurgie orthopédique et réanimation adulte.

La méthode de sélection et d'annotation de ces comptes rendus a été décrite en détail dans une autre publication [9]. Ce corpus de données textuelles a été divisé par tirage aléatoire des dossiers en 2 jeux : un jeu d'apprentissage et un jeu test. L'annotation consistait à fournir une référence pour la détection des infections nosocomiales.

Un 2^{ème} corpus de données textuelles sera constitué dans le cadre de ce nouveau projet et concerne un domaine médical très différent : la prise en charge diagnostique du cancer du côlon. Le choix de ce domaine médical vise à évaluer le caractère générique de la solution. Le nouveau corpus sera composé de 300 documents textuels médicaux, sélectionnés dans le centre anti-cancer régional de Rhône-Alpes chez des patients pris en charge pour un nouveau cancer du côlon. Le corpus sera constitué des documents textuels traçant la prise en charge du cancer du côlon : le compte rendu hospitalier, le compte rendu d'anatomo-pathologie, comptes rendus d'imagerie, comptes rendus de consultation, comptes rendus opératoires, compte rendu de concertation pluridisciplinaire). Un tirage aléatoire de 100 dossiers permettra de développer les algorithmes de repérage du cancer du côlon ainsi que les éléments diagnostiques nécessaires au calcul de différents délais de prise en charge.

Une annotation complémentaire d'un sous-groupe de dossiers est en cours afin de fournir la référence pour l'apprentissage et l'évaluation du peuplement de la base de faits. L'annotation est semi-automatique et réalisée à l'aide d'une application développée par le LBBE sur les logiciels R et MS Access. Cette annotation permet de reconstituer l'histoire médicale du patient, d'identifier le motif d'entrée, la prise en charge initiale et l'évolution. Il s'agit d'une restructuration a posteriori de la trajectoire de soins du patient effectuée dans le cadre de la prise en charge de l'événement de santé d'intérêt pour l'étude (par exemple prise en charge pour pose d'une prothèse de hanche et suivi dans les 12 mois). Cette annotation permet donc d'étiqueter les événements de santé temporellement dans cette trajectoire (s'agit-il d'un antécédent, d'un motif d'entrée,...), de distinguer les répétitions textuelles se rapportant au même événement de santé de celles correspondant à de réelles répétitions du même événement de santé dans la trajectoire du patient. La base d'annotation de l'application constitue le modèle conceptuel de la base de faits [6].

Les résultats aboutissant au repérage de l'infection nosocomiale par la solution SYNODOS seront validés par les médecins investigateurs sur un échantillon de 400 comptes rendus d'infections nosocomiales (soit 100 rapports par spécialité) et sur 400 comptes rendus de patients sans infection nosocomiale (soit 100 rapports par spécialité). Les résultats obtenus en termes de sensibilité / spécificité seront considérés comme les performances définitives de l'outil. Pour l'estimation du délai de prise en charge du cancer du côlon, le délai estimé par méthode automatique sera comparé au délai estimé par méthode manuelle sur 200 documents textuels. L'indicateur d'évaluation retenu sera la différence moyenne de délai de prise en charge du cancer du côlon obtenu entre la solution SYNODOS et la méthode manuelle de référence.

4 Résultats

4.1 Prétraitement des données hospitalières pour l'exploitation SYNODOS

La solution SYNODOS sera installée dans la zone démilitarisée du réseau sécurisé de l'établissement de santé ou chez un hébergeur agréé de données de santé afin de bénéficier de toute la sécurité liée au traitement de données médicales (cf. Figure 1).

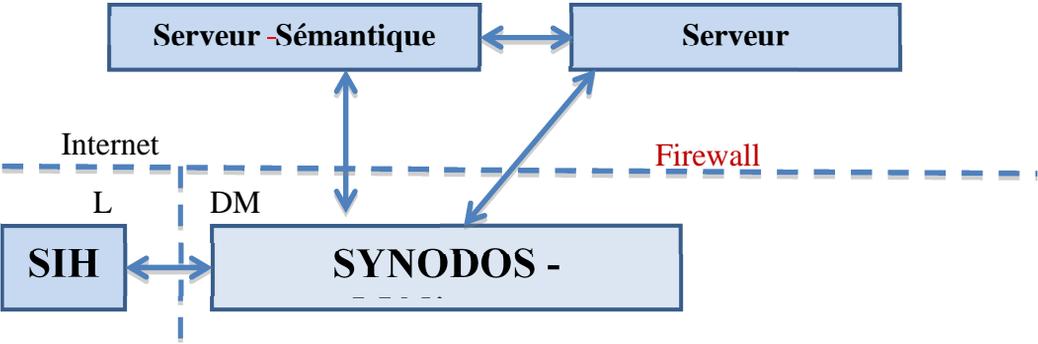


Figure 1 : Interface système d'information hospitalier, médiateur SYNODOS et serveurs Web distants

Objet Direct fournira une machine virtuelle prête à l'emploi, qu'il suffira de déployer sur le superviseur de l'établissement. Un module d'importation des données provenant du système d'information hospitalier récupèrera les métadonnées et les données textuelles « poussées » par le SIH de l'établissement. Le module procède ensuite à une anonymisation automatisée des données nominatives ou indirectement nominatives (ex : noms de personnes, noms de lieux, numéros de téléphone, adresses, adresse mail, etc.) afin de permettre les échanges avec les services web distants et afin également de permettre un usage épidémiologique des données. Ce module permettra également de désanonymiser le document après traitement pour certains types d'usage (par exemple surveillance intra-hospitalière des infections nosocomiales, où le dossier patient une fois repérée par le système doit être validé par l'équipe d'hygiène).

4.2 Indexation terminologique et analyse sémantique des documents textuels

Une fois les documents textuels formatés avec leurs métadonnées hospitalières dans un fichier XML, un 1^{er} traitement linguistique va être réalisé : détection et découpage des phrases par l'analyseur sémantique. Puis l'analyseur envoie une requête en mode SOAP au serveur terminologique afin que chaque concept de chaque phrase puisse être indexé. L'ECMT renvoie alors dans un fichier XML pour chaque concept médical reconnu, plusieurs métadonnées qui seront nécessaires à la suite du traitement (cf. Figure 2) : a) l'offset du 1^{er} caractère du concept médical b) le « concept unique identifiant (CUI) » et le « type sémantique » si le concept médical est inclus dans le Metathesaurus de l'UMLS, c) sa spécialité médicale identifiée par des liens sémantiques établis manuellement par l'équipe CISMéF ou par l'utilisation de terminologies contenant ce niveau de granularité. Le serveur renvoie la liste des termes indexés à l'analyseur sémantique selon un protocole HTTPS. L'analyse est effectuée phrase par phrase afin d'obtenir des délais de réponse compatibles avec la suite du traitement par la solution SYNODOS.

Après l'indexation, la suite du traitement linguistique consiste en la tokénisation, l'étiquetage morpho-syntaxique, la lemmatisation, l'analyse morphologique et l'analyse syntaxique en dépendances. L'analyseur propose une représentation graphique des dépendances syntaxiques, en permettant une visualisation et donc un contrôle de la qualité (cf. Figure 3).

```
- <return>
- <cis:sentence idpatient="0" idsentence="0" iddoc="0">
  <cis:text>le patient est arrivé aux urgences sous oxygène après avoir fait une crise d'asthme.</cis:text>
- <cis:indexations>
- <cis:indexation end="84" idterm="9114" idcismef="ADI_PA_9114" offset="78" start="78" ter="ADI">
  - <cis:term>
    <cis:label lang="fr">ASTHME</cis:label>
  </cis:term>
  <cis:categorization />
</cis:indexation>
- <cis:indexation end="48" idterm="V03AN01" idcismef="ATC_CD_V03AN01" offset="41" start="41" ter="ATC">
  - <cis:term>
    <cis:label lang="fr">V03AN01 - oxygène</cis:label>
  </cis:term>
  <cis:categorization />
```

Figure 2 : extrait du fichier XML produit en sortie de l'extracteur multi-terminologique CISMéF

L'analyse sémantique proprement dite se base ensuite sur le résultat de cette analyse syntaxique, ainsi que sur des connaissances ontologiques ou terminologiques. Cette analyse consiste à représenter les éléments de sens, permettant d'utiliser l'information extraite de façon adéquate : qui a fait quoi ? quand ? comment ? où ?

Pour représenter le résultat de l'analyse sémantique, l'analyseur permet de définir des relations (« prédicats logiques ou faits »). Par exemple pour la phrase « le patient est arrivé aux urgences sous oxygène après avoir fait une crise d'asthme », la Figure 3 représente l'analyse sémantique produite. Les nœuds correspondent aux entités (objets, endroits, personnes, etc.) et les arcs aux relations sémantiques définies entre ces entités.

L'analyseur sémantique envoie ensuite selon le protocole HTTPS les données issues du traitement sémantique sur un répertoire de stockage de la solution SYNODOS. Ces données sont stockées sous forme de fichiers XML avec un document textuel stocké par fichier.

4.3 Le système expert

Deux niveaux de règles expertes vont devoir être élaborés : 1) des règles expertes, permettant de traiter les données issues des 2 modules (terminologique et sémantique) et des métadonnées hospitalières afin de les ordonner dans la base de faits. Ce type de règles a été nommé « règles de transition ». Il s'agit par exemple de créer une relation entre un diagnostic médical tel que « pneumopathie à pneumocoques » et le statut « antécédent » de cette pathologie notifiée dans le compte-rendu; 2) des règles expertes permettant d'enrichir la base de faits avec de nouveaux faits, tels que par exemple une détection d'infection nosocomiale.

Les règles de transition font appel à des règles expertes de type « linguistique » ou à des règles expertes de type « médical ». Certaines règles nécessitent de combiner les 2 approches mais il est tenté de dissocier au maximum les 2 types de règles afin de répondre à l'un des défis du projet, à savoir permettre aux utilisateurs médicaux d'adapter eux-mêmes la solution au domaine d'application qui les concerne. La base de faits est en cours de développement et consiste à reproduire le modèle conceptuel élaboré par le LBBE lors de la procédure d'annotation manuelle, détaillée dans Gicquel et al. [6].

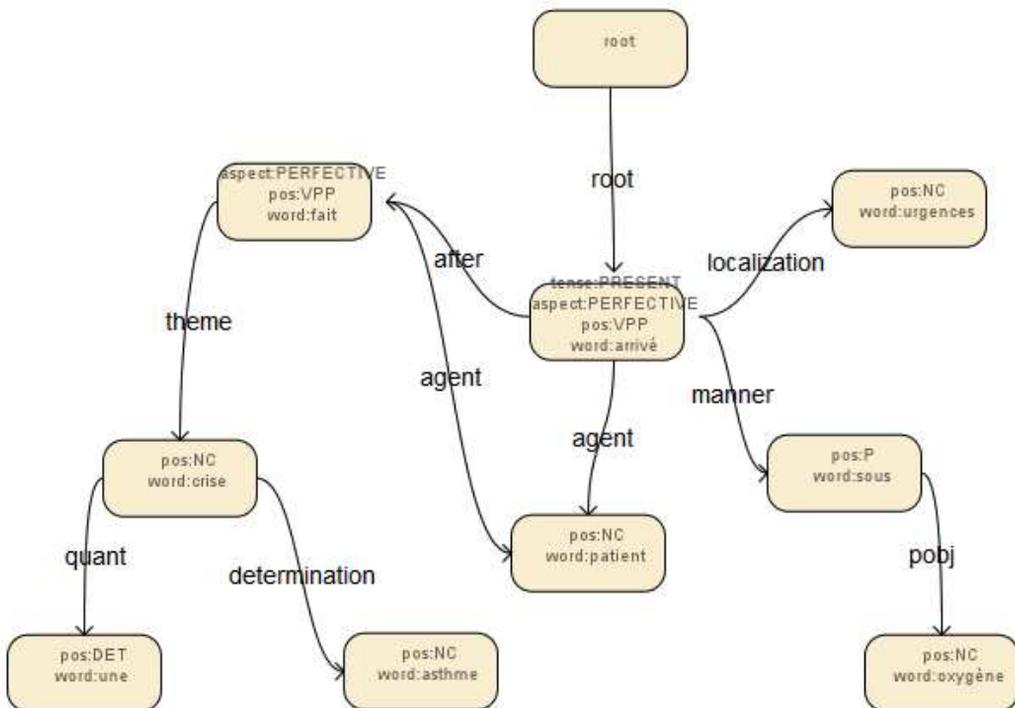


Figure 3 : Exemple de graphe sémantique

Deux approches ont été retenues pour le développement du module « système expert ». Une approche traditionnelle (classique) via l'utilisation d'un outil du marché (BRMS: Business Rules Management System) et une approche orientée Recherche qui étudie le Web sémantique et ce qu'il peut apporter dans un système expert.

5 Discussion

Le projet SYNODOS est donc un projet à visée très opérationnelle et dont l'un des objectifs majeurs est d'intégrer et d'interfacer des technologies encore très expérimentales (serveur terminologique, analyseur sémantique, systèmes expert) dans une solution utilisable en production hospitalière. Le développement n'est pas encore terminé à ce stade pour en évaluer les performances globales mais il est prévu de tester les fonctionnalités de la solution SYNODOS en mode production dès 2015 au Centre Léon Bérard (centre régional de cancérologie de Rhône-Alpes).

L'intérêt d'un tel outil dans le fonctionnement hospitalier est multiple : phénotypage des patients dans le cadre d'études épidémiologiques, screening de sujets éligibles pour diverses interventions, amélioration de la qualité, aide à la décision thérapeutique, évaluation de la prise en charge etc.

Ce projet présente toutefois un certain nombre de limites lié aux verrous technologiques existants : les ressources terminologiques mises à disposition pour l'indexation automatique des concepts médicaux bien que très riches, ne répondent pas de façon exhaustive et avec le niveau de précision souhaité à tous les domaines d'application. Par exemple, dans le cadre du domaine d'application sur la détection des infections nosocomiales, il a été mis en évidence la nécessité de développer une terminologie spécifique additionnelle, qui n'a pas pu être générée de manière automatisée avec les techniques classiques d'ingénierie des connaissances. Cette terminologie est donc enrichie manuellement en fonction de l'annotation manuelle du corpus, ce qui pose le problème de sa complétion. Selon les domaines d'application, les futurs utilisateurs de la solution pourront être confrontés à cette limite. L'étiquetage temporel des concepts médicaux dans le parcours de soins du patient est un autre élément indispensable à l'exploitation ultérieure de l'information médicale extraite. La pertinence de cet étiquetage temporel sera évalué dans le cadre du projet ; elle dépend à la fois de la qualité de l'analyse sémantique des documents mais aussi de la qualité des « règles de transition ». Plus généralement, les 2 challenges les plus importants du projet sont d'une part définir un modèle conceptuel suffisamment générique et d'autre part obtenir des « règles de transition » très diverses pour la représentation des données dans la base de faits afin que leur analyse ultérieure corresponde de la manière la plus exacte possible à la situation médicale décrite dans le DPI. Afin d'enrichir le modèle et les règles en fonction des utilisations ultérieures par domaine d'application, il est prévu une interface utilisateur permettant d'ajouter des éléments à la base de connaissances et de nouvelles règles expertes. L'évaluation de la solution SYNODOS prévue en 2015, permettra de préciser l'impact réel de ces limites.

D'autres limites sont inhérentes aux propriétés de ce type de données (incomplètes, complexes, biaisées). Hripcsak et al. ont notamment souligné les limites de l'usage secondaire du DPI à des fins de phénotypage [10]. Les perspectives seront d'étudier l'impact de ces propriétés sur la qualité du phénotypage.

6 Conclusion

La solution SYNODOS constitue un projet innovant dans la mesure où diverses technologies terminologiques, sémantiques et expertes sont assemblées pour traiter et analyser l'information médicale très riche des dossiers médicaux mais dont les données sont essentiellement textuelles. La conception de la solution que nous venons de décrire, vise à exploiter ces données pour tout domaine d'application médicale. Ce caractère générique devrait être assuré d'un point de vue fonctionnel par le fait que ce sont les utilisateurs de la solution qui choisiront et adapteront eux-mêmes les événements médicaux d'intérêt qu'ils souhaitent analyser dans leurs dossiers médicaux.

Remerciements

Ces travaux ont été financés par l'Agence Nationale de Recherche, dans le cadre d'un programme TECSAN (projet SYNODOS n° ANR-12-TECS-0006).

Références

- [1] Olive F, Gomez F, Schott AM, Remontet L, Bossard N, et al. (2011) [Critical analysis of French DRG based information system (PMSI) databases for the epidemiology of cancer: a longitudinal approach becomes possible]. *Rev Epidemiol Sante Publique* 59: 53-58.
- [2] Defossez G, Rollet A, Dameron O, Ingrand P Temporal representation of care trajectories of cancer patients using data from a regional information system: an application in breast cancer. *BMC Med Inform Decis Mak* 14: 24.
- [3] Gottesman O, Kuivaniemi H, Tromp G, Faucett WA, Li R, et al. (2013) The Electronic Medical Records and Genomics (eMERGE) Network: past, present, and future. *Genet Med* 15: 761-771.
- [4] Thompson WK, Rasmussen LV, Pacheco JA, Peissig PL, Denny JC, et al. An evaluation of the NQF Quality Data Model for representing Electronic Health Record driven phenotyping algorithms. *AMIA Annu Symp Proc* 2012: 911-920.
- [5] Pathak J, Bailey KR, Beebe CE, Bethard S, Carrell DC, et al. Normalization and standardization of electronic health records for high-throughput phenotyping: the SHARPN consortium. *J Am Med Inform Assoc* 20: e341-348.
- [6] Gicquel Q, Kergourlay I, Gerbier-Colomban S, Chariout S, Bittar A, et al. (2014) Annotation methods to develop and evaluate a medical expert system based on natural language processing in electronic medical records. MIE. Istanbul, Turkey.
- [7] JBoss Drools Business Logic Integration Platform. <http://www.jboss.org/drools>
- [8] Proux D, Hagège C, Gicquel Q, Kergourlay I, Pereira S, et al. (2012) ALADIN : Développement d'un outil sémantique d'analyse des documents textuels médicaux pour la détection d'infections associées aux soins IRBM, *Ingenierie et Recherche BioMedicale* 33: 137-142.
- [9] Metzger M, Gicquel Q, Kergourlay I, Cluze C, Grandbastien B, et al. (2011) Codage standardisé de données médicales textuelles à l'aide d'un serveur multi-terminologique de santé : exemple d'application en épidémiologie hospitalière In: Degoulet P, Fieschi M, editors. *Systèmes d'information pour l'amélioration de la qualité en santé : comptes rendus des quatorzièmes Journées francophones d'informatique médicale*, Tunis, 23 et 24 septembre 2011. Paris: Springer-Verlag France. pp. 109-120.
- [10] Hripcsak G, Albers DJ (2013) Next-generation phenotyping of electronic health records. *J Am Med Inform Assoc* 20: 117-121.

Adresse de correspondance

Marie-Hélène Metzger, UCBL-CNRS UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 43 boulevard du 11 novembre 1918, Villeurbanne F-69622, France.

E-mail : marie-helene.metzger@chu-lyon.fr

URL : <http://www.synodos..fr>

Protection des données de santé lors de leur réutilisation à des fins de recherche au sein des entrepôts de données biomédicaux (EDBM)

Health data protection for secondary use in clinical data warehouse

Christine Riou¹ Marc Cuggia²

¹ *Département d'Information Médicale, CHU de Rennes, France*

² *INSERM LTSI - Equipe Projet Données massives en santé
Faculté de Médecine de Rennes, France*

Résumé

Introduction : La mise en œuvre des entrepôts de données biomédicaux (EDBM) pose des questions d'ordre juridique, déontologique et éthique. L'objectif de cet article est d'effectuer un état des lieux sur ces points et de formuler des propositions organisationnelles pour l'exploitation des EDBM. Méthodes : les auteurs ont effectué une revue de la littérature dans Pubmed, Google Scholar, Google, Science Direct. Elle a concerné des articles scientifiques, des textes de loi ou réglementaires ainsi que des procédures de mise en œuvre des systèmes publiés. Résultats : l'utilisation des données à caractère personnel pour la recherche en santé est conditionnée à l'accord des patients et à l'autorisation d'autorités de contrôle. Le patient dispose d'un droit d'information et d'un droit d'opposition. Les solutions proposées pour assurer la protection des données dans les EDBM s'appuient essentiellement sur la dé-identification des données et l'organisation des accès. Conclusion : l'information du patient reste délicate. Les procédures de dé-identification sont à améliorer et à certifier. Le chaînage des données s'avère nécessaire pour le croisement avec d'autres sources externes. L'organisation proposée par les auteurs pour l'exploitation des EDBM repose sur la mise en place de structures dédiées ainsi que sur la rédaction de documents cadre et sur la définition de procédures.

Abstract

Background : Implementing clinical data warehouse (CDW) brings out legal and ethical issues. Methods : the authors make a review of recent litterature on these points in Pubmed, Google scholar, Google, Science direct. It includes scientific papers, texts of laws, and operational procedures. Then they propose an organisation for CDW implementation. Results : use of health personal data required patient's consent and an authorization of competent authorithies. Rights of information and opposition are due to the patient. To assure data protection in CDW measures rely on data deidentification and hierarchical management of access. Conclusion : patient information remains difficult. Deidentification procedures are to be improved and

certified for french language. Crossing data needs chaining them. The organisation proposed by the authors lies on specific structures and documents (policy for the use of CDW, patient information note, charter for good use).

Mots clés : Confidentialité, dé-identification, entrepôt de données biomédicales, dossier patient informatisé, recherche en santé

Keywords : Confidentiality, deidentification, data warehouse, electronic patient record, secondary use, medical research

1 Introduction

Avec l'informatisation des établissements de santé, la dématérialisation des données patient modifie le rapport à l'information médicale. Ces données sensibles, hétérogènes et multidomaines deviennent à présent massives (notion de bigdata) et potentiellement accessibles de façon ubiquitaire. Ces données constituent un gisement d'informations majeur qui peut être réutilisé pour la recherche. L'exploitation concomitante des données cliniques et des données issues du génome (données OMICS) et plus généralement la recherche de biomarqueurs in silico répond aux besoins actuels de la médecine translationnelle et personnalisée.

Pour cela, les établissements se dotent de nouveaux outils appelés entrepôts de données biomédicaux (EDBM) permettant la fouille et l'exploitation de ces données. Cependant, leur mise en oeuvre pose un certain nombre de questions, tant sur le plan de la sécurité des données, que sur les plans réglementaire, déontologique et éthique. Ainsi, selon la loi Informatique et libertés [1], l'accès aux données à caractère personnel suppose que le patient ait donné son accord¹ pour une étude précise mobilisant des données bien définies². Or, les EDBM contiennent par définition l'exhaustivité des données patient rétrospectives d'un établissement. Il est donc impossible de connaître à l'avance précisément l'objet des études qui seront réalisées grâce à l'entrepôt et la nature des données qui seront utilisées. La loi définit un principe de parcimonie consistant à n'utiliser strictement que les données nécessaires et suffisantes à l'étude. Or dans un EDBM, le principe de fouille et de recherche d'information suppose l'accès et le traitement d'informations beaucoup plus larges. La loi actuelle est facilitatrice pour les professionnels de santé qui exploitent les données des patients dont ils ont eu la charge³. Or, un EDBM permet une exploitation transversale des données par des utilisateurs qui n'ont pas été nécessairement impliqués dans la prise en charge du patient. Ce cas de figure peut être sujet à interprétation, simple déclaration ou autorisation « recherche médicale »⁴, plus contraignante à mettre en oeuvre et longue à obtenir.

D'autres questions se posent : Quel est le rôle du patient dans ce dispositif ? Comment assurer la protection de sa vie privée ? Comment organiser les accès à l'entrepôt ?

L'objectif de ce travail est d'effectuer un état des lieux des solutions adoptées pour assurer la

¹ sauf dérogation spécifique de la Commission Nationale Informatique et Libertés (CNIL)

² Article 30 de la loi informatique et libertés

³ Article 53 de la loi informatique et libertés

⁴ Chapitre IX de la loi Informatique et libertés

protection des données de santé lors de leur réutilisation pour la recherche dans les entrepôts de données biomédicales.

A partir de cette analyse les auteurs formulent des propositions opérationnelles dans le respect de la réglementation française et adaptées au cadre organisationnel des établissements de santé permettant l'exploitation des documents structurés et textuels dans les EDBM tout en garantissant la protection des données des patients.

2 Etat de l'art

Nous commencerons par définir les concepts clefs abordés dans cet article.

2.1 Les entrepôts de données biomédicales

Les entrepôts de données biomédicales ont vu le jour aux Etats-Unis au milieu des années 2000 [2-3], ils sont en cours d'implémentation dans les autres pays notamment en Europe [4]. Un entrepôt de données biomédical (EDBM) regroupe les données patient, produites lors de leurs prise en charge, stockées dans les différentes bases de données du système d'information de l'établissement (dossier patient informatisé, système de gestion des laboratoires, système d'information de radiologie, système de dispensation des médicaments, système de recueil d'activité des blocs opératoires et plateaux médico-techniques). Il peut-être étendu aux données médico-économiques mais aussi aux données OMICS.

Il permet une indépendance par rapport au système de production et une centralisation des données. Il intègre des outils de fouille de données, autorisant des requêtes croisées sur des données structurées ou textuelles, sur des critères biologiques et cliniques par exemple, ainsi que des outils de visualisation (timeline, cluster, regroupement géographique) [4].

Il peut être utilisé dans différents domaines :

- Recherche clinique [5] : étude de faisabilité, recherche de patients éligibles (Pre-screening) à un protocole de recherche, alimentation des cahiers d'observation électronique (e-CRF)
- Epidémiologie : constitution de cohortes épidémiologiques, alimentation de registres [6]
- Evaluation des pratiques professionnelles
- Vigilances : la détection d'effets indésirables [7] ou la détection des infections nosocomiales [8]
- Etude médico-économique [9]
- Recherche translationnelle : [10]

Leur exploitation est centrée dans une approche populationnelle. A titre d'exemple nous présentons 3 outils actuellement opérationnels :

- Informatics for Integrating Biology and the Bedside (i2b2) [3]

Il s'agit d'un système développé à la faculté de médecine de Harvard et implanté dans plusieurs centres hospitalo-universitaires américains et européens [11-12]. I2B2 est un outil de recherche translationnelle permettant l'analyse combinée de données cliniques et génomiques. Cette plateforme Open Source connaît une large adoption par le monde

académique et industriel.

- Stanford Translational Research Integrated Database Environment (STRIDE) [2] développé et mis en œuvre à l'université de Stanford.

Dans l'entrepôt STRIDE il est mis à disposition des chercheurs de l'université de Stanford trois outils [13] :

- Outil pour étude de faisabilité avec production de résultats agrégés (cohort discovery tool),
 - Outil pour la constitution de cohorte (cohort data review tool)
 - Outil d'extraction de données (data extraction tool).
- Roogle entrepôt déployé en France dans différents établissements de santé [14].

Le système Roogle est un moteur de recherche développé par l'université de Rennes 1 et le CHU de Rennes. Il permet la réexploitation combinée des données structurées et textuelles, la réalisation de requêtes centrées sur une population, et également l'interrogation au sein d'un dossier patient individuel.

2.2 Les données à caractère personnel

Un entrepôt de données biomédical contient des données de santé à caractère personnel. On entend par donnée à caractère personnel « toute information relative à une personne physique identifiée ou qui peut être identifiée directement ou indirectement par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres ». [1].

Les conditions de protection de ces données lors de leur traitement sont fixées par la loi, comme la loi Informatique et libertés en France [1], l'Health Insurance Portability and Accountability Act (HIPAA) [15] ou encore la Common Rule aux Etats-Unis, the Data Protection Act au Royaume Uni [16].

Dans tous les pays, l'utilisation de données de santé à caractère personnel pour un projet de recherche est conditionnée à l'accord des patients et à l'autorisation d'autorités de contrôle garantissant ainsi la protection des personnes, la confidentialité, le bien fondé et la sécurité du traitement. L'autorisation ainsi que l'accord du patient sont donnés pour chaque traitement. En France les comités de protection des personnes (CPP) jouent un rôle éthique, le Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) apporte un avis méthodologique et la Commission nationale informatique et libertés (CNIL) s'assure que le traitement est effectué conformément à la législation [17]. Aux Etats Unis les Institutional Review Boards (IRB) mis en place dans les différentes institutions autorisent les traitements après s'être assurés que le projet ne pose pas de problème éthique et que les droits des personnes participant au projet de recherche sont respectés [18]. Il en est de même au Canada pour les Research Ethics Boards (REB) et au Royaume Uni pour les Research Ethics Committees (RECs). Au Royaume Uni chaque institution nomme une personne, le Caldicott Guardian, responsable de la protection des données, garant du respect de la législation en matière de confidentialité notamment lors du partage d'informations [19]. Celle-ci est consultée lors de la réutilisation de données de santé.

Les traitements concernant des données anonymes ne sont pas concernés par les lois.

2.3 Les données anonymes

En France au sens de la CNIL [20], les données sont considérées anonymes si elles ne permettent pas d'identifier, directement ou indirectement par regroupement, une personne physique. Au niveau européen, les principes édictés dans la directive 95/46/CE relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel [21] ne s'appliquent pas « aux données rendues anonymes d'une manière telle que la personne concernée n'est plus identifiable ».

Il n'y a pas de procédé permettant d'assurer un anonymat strict des données individuelles. Le terme de dé-identification est apparu aux Etats Unis après définition dans la loi (HIPAA) des catégories de données identifiantes dénommées protected health information (PHI).

2.4 Données identifiantes

Aux Etats-Unis, la loi, l'HIPAA, a défini 18 catégories de données identifiantes, protected health information (PHI) [15]. Elles comportent les noms propres (patients, médecins, établissements de santé...), les adresses et codes de résidence précis, les dates, les numéros attachés au patient (numéro sécurité sociale, numéro assurance, numéro de dossier médical, ...), numéros de téléphone, fax, adresses email, données biométriques, photographies d'identité, les caractéristiques uniques du patient.

L'HIPAA définit trois contextes de dé-identification [22].

- Safe Harbor Standard où les 18 catégories de données identifiantes sont occultées
- Limited Data Set standard (LDS) où les dates et les codes géographiques sont autorisés.
- Statistical standard consiste à certifier que la probabilité d'identifier le patient est très faible.

Au Royaume Uni, dans l'objectif de développer un service pour la réutilisation des données de soins [23], des préconisations ont été définies pour la dé-identification des données à travers le Pseudonymisation Implementation Project (PIP) [24]. Sont considérées comme données identifiantes les noms, adresses, codes postaux, dates de naissance et de décès, numéros d'identification patient (N° NHS, N° local, N° hospitalier, ...) et catégorie ethnique [25].

En France la CNIL n'a pas à ce jour établi de liste de données identifiantes. Dans le cadre du Programme de Médicalisation du système d'Information [26], un certain nombre d'informations sont considérées comme sensibles, les dates, le numéro de séjour, l'unité médicale, le code postal.

2.5 Dé-identification des données.

Dé-identification n'est pas synonyme d'anonymat. Il s'agit d'un procédé par lequel les données identifiantes sont supprimées ou occultées. La probabilité de retour à l'identité du patient est alors extrêmement faible. Il persiste cependant un risque de réidentification [27]. Le croisement de fichiers augmente ce risque.

Les préconisations de la CNIL en matière d'anonymisation de fichier s'apparentent à une technique de dé-identification, outre la suppression des données nominatives et des numéros identifiants, remplacement des dates par l'année, des codes géographiques par le département

[28]. Une réflexion est en cours dans le cadre de l'open data sur de nouvelles solutions techniques d'anonymisation.

Nous préférons utiliser par la suite le terme de dé-identification.

2.6 Droits des patients lors d'un traitement de données à caractère personnel

Un traitement de données à caractère personnel doit se faire dans le respect des droits des patients. Conformément à la loi informatique et libertés le patient dispose d'un droit d'information et d'un droit d'opposition. Les informations contenues dans son dossier médical sont couvertes par le secret. Cependant une dérogation au secret médical pour la recherche est prévue dans la loi conditionnellement aux dispositions inscrites dans le chapitre IX Traitement de données à caractère personnel ayant pour fin la recherche en santé.

Les informations fournies au patient selon l'article 57 de la loi concernent la nature des informations transmises, la finalité du traitement de données, les personnes physiques ou morales destinataires des données, le droit d'accès et de rectification, le droit d'opposition.

L'utilisation de données génétiques est soumise au consentement écrit du patient.

La réutilisation des données de soins implique le plus souvent l'accès aux données nominatives. Selon la loi l'accord du patient est donc requis. Devant la difficulté à retrouver les patients, une dérogation à l'information individuelle⁵ peut cependant être obtenue auprès de la CNIL pour une étude donnée.

Le dossier médical contient des informations relevant de l'intimité du patient. Le code de santé publique [29] stipule que « toute personne prise en charge dans un établissement de santé a droit au respect de sa vie privée et au secret des informations le concernant ».

3 Matériel et méthodes

Une revue systématique de la littérature a été effectuée. La recherche documentaire s'est basée sur l'interrogation de Pubmed, Google Scholar, Google, Science direct. Elle concernait des articles scientifiques d'intérêt, et des textes de loi ou réglementaires, ainsi que des procédures de mise en œuvre des systèmes publiés.

Les requêtes ont utilisé les mots clefs suivants : electronic health record (EHR), electronic medical record (EMR), electronic patient record, medical record system, data warehouse, secondary use, reuse, de-identification, pseudonymisation, patient identifier, anonymisation, anonymized data, privacy, confidentiality, informed consent, consent, clinical research, translationnal medicine.

Nous avons également inclus des publications référencées dans les papiers sélectionnés.

Enfin nous nous sommes appuyés sur le guide de bonnes pratiques pour assurer la confidentialité des données de soins lors de leur réutilisation pour des traitements biostatistiques rédigé par un groupe de travail constitué de chercheurs, d'enseignants de biostatistiques, de médecins de département d'information médicale et de juristes [30].

⁵ Article 57 loi Informatique et libertés

4 Résultats

Au total, 66 références ont été sélectionnées et utilisées dans notre étude. Elles concernent 40 publications scientifiques, 7 références de texte de loi ou réglementaire, 12 rapports ou guides de référence, 5 site web de projets.

Nous dresserons dans un premier temps un panorama des méthodes déployées pour garantir la confidentialité des informations de santé dans les EDBM. Nous traiterons ensuite un sujet qui lui est lié, celui de l'information du patient.

4.1 Protection des données

Les solutions proposées pour assurer la protection des données dans les entrepôts biomédicaux mettent en avant deux points essentiels, les procédures de dé-identification des données issues des dossiers médicaux et l'organisation des accès.

Nous dresserons dans un premier temps un panorama des méthodes déployées pour garantir la confidentialité des informations de santé dans les EDBM. Nous traiterons ensuite un sujet qui lui est lié, celui de l'information du patient.

4.1.1 Les procédures de dé-identification

Les procédures de dé-identification comportent deux étapes : le repérage puis le traitement des données identifiantes (suppression ou remplacement).

Le repérage

Il est relativement aisé lorsqu'il s'agit de données structurées puisque les algorithmes peuvent s'appuyer sur les éléments et les types de données (par exemple les informations structurées issues du Système d'Information Hospitalier (SIH) comme le nom, l'adresse ou le numéro de séjour).

Il n'en va pas de même pour les documents textuels. Pour ce type de document les algorithmes reposent sur deux types de méthodes [31,32]. Les méthodes symboliques (pattern matching) utilisent des listes de termes (listes de noms, de prénoms, de villes, etc.), des listes de déclencheurs (déclencheurs de personnes : « M., Mme, Melle, Dr, Pr », etc.) et des dictionnaires (langue générale ou langue de spécialité). Les méthodes d'apprentissage automatique nécessitent auparavant l'annotation de textes. Des méthodes hybrides combinent les deux types d'approche.

Pour le traitement de données structurées, il faut citer une autre méthode, la méthode du k anonymat qui permet de repérer les cas identifiables du fait de leur caractère d'unicité et de les affecter de caractéristiques plus génériques [33].

Même si ces méthodes ont démontré de bonnes performances [34, 35], elles présentent cependant des limites [34,36]. L'ambiguïté de certains termes, l'absence de termes dans les dictionnaires ou des fautes d'orthographe peuvent conduire à un mauvais filtrage des données. Ces algorithmes sont souvent développés à partir d'échantillons de compte rendus spécifiques à une institution ou à un domaine, la performance peut ne pas être reproductible. Enfin, la lisibilité et la compréhensibilité de textes dé-identifiés peuvent également être affectées.

Le traitement

Une fois repérées, les données identifiantes sont traitées pour être transformées ou supprimées. Dans ce dernier cas il n'est pas possible de chaîner les données relatives à un même patient.

Certaines données peuvent être floutées, comme les dates remplacées par l'année ou un intervalle de valeurs pour l'âge, ou les codes postaux regroupés pour atteindre un nombre d'habitants suffisant ou remplacés par le département. Elles peuvent aussi être remplacées par des valeurs non significatives ne permettant pas de déduire l'identité des patients, on parle alors de pseudonymisation [24, 32, 37]. Les pseudonymes relatifs aux numéros de séjour ou de patient sont générés de telle sorte que toutes les données propres à un séjour ou un patient puissent être regroupées et chaînées entre elles. Un pseudonyme doit être unique et répliquable pour une même donnée à dé-identifier. Un pseudonyme peut être généré soit par une fonction de hachage irréversible, soit par une méthode de cryptage réversible. Le pseudonyme généré n'est alors pas intelligible par un humain. Cette approche est utilisée l'entrepôt de données I2B2 de l'HEGP [11]. En cas de cryptage réversible, le déchiffrement permet de restituer les données initiales en clair. La réversibilité de la pseudonymisation peut également être assurée par une table de correspondance. Pour une meilleure lisibilité des données textuelles, les données identifiantes (comme le nom et le prénom) peuvent aussi être substituées par des informations de même nature puisées aléatoirement dans des tables disponibles. Les dates sont remplacées de telle sorte que la temporalité des événements soit conservée [32, 33].

Les outils de repérage et de traitement

Plusieurs algorithmes ont été publiés pour la dé-identification de textes en langue anglaise [38, 39]. Ils s'adaptent mal à la langue française. Des outils propres au français ont été développés et évalués comme FASDIM par E Chazard [40] et Medina par C. Grouin ou encore Wapiti [32, 41].

Concernant la pseudonymisation, on peut citer trois services utilisés à l'échelle nationale pour assurer à la fois la pseudonymisation et le chaînage des données patients au delà d'un établissement.

Au Royaume Uni dans le cadre du Pseudonymisation Implementation Project, des outils de dé-identification ont été publiés [24]. Ils sont mis en œuvre dans des projets d'institutions affiliées au NHS [42]. Par ailleurs, l'Agence sanitaire nationale, Health and social care information center (HSIC) [43], met en place un service sécurisé d'appariement des données pour permettre le chaînage des données de santé hospitalières et ambulatoires [44].

En Allemagne un service de pseudonymisation est mis à disposition des organismes de recherche sur une plate-forme gérée par une structure indépendante centrale jouant le rôle de tiers de confiance, the german technology and method platform for networked medical research (TMF) [37]. Il permet d'obtenir un numéro patient unique non significatif. Il est mis en œuvre dans le projet d'implémentation de l'entrepôt i2b2 à l'université d'Erlangen [12].

En France des fonctions d'occultation des informations nominatives (FOIN) sont utilisées pour l'anonymisation et le chaînage des séjours des patients des fichiers PMSI (Programme de Médicalisation des Systèmes d'Information) ainsi que pour l'anonymisation (plus exactement la pseudonymisation) des données de la base SNIIRAM (Système National Inter régime de l'Assurance Maladie) qui comprend les données de remboursement de l'ensemble des assurés [45].

4.1.2 Les procédures d'accès et d'exploitation des entrepôts de données biomédicaux

L'exploitation sécurisée d'un entrepôt de données suppose la mise en œuvre d'une procédure d'accès basée sur la définition de droits utilisateur et la traçabilité des accès.

En fonction des cas d'utilisation, par exemple étude de faisabilité, pré-screening ou monitoring, les utilisateurs accèdent respectivement à des données agrégées, dé-identifié ou nominatives.

Ainsi, à chaque niveau de dé-identification correspondent un profil d'utilisateur et un contexte d'utilisation.

Nous présentons ici trois exemples de procédures d'accès et d'exploitation d'entrepôts de données biomédicaux.

I2B2 :

Les droits d'accès à l'entrepôt sont basés sur le niveau de confidentialité des données et le degré de confiance envers l'utilisateur. Le niveau de confidentialité est lié au niveau de dé-identification. L'algorithme de dé-identification des données dans l'entrepôt doit être validé par l'IRB. L'exploitation de l'entrepôt I2B2 aux Etats Unis fait transparaître cinq niveaux d'accès correspondant à des cas d'utilisation différents [46].

Les niveaux 1 et 2 correspondent à l'accès à des données agrégées anonymes. Ces niveaux sont affectés à des utilisateurs non formés à la sécurité informatique. L'entrepôt est utilisable à partir de postes de travail dont la sécurisation n'est pas optimale.

Concernant le niveau 1, les résultats sont floutés. Ils sont produits à partir d'un entrepôt de données non dé-identifié. Des croisements d'informations sont possibles mais limités et prédéfinis.

Dans le niveau 2, ce sont les effectifs réels qui sont présentés mais calculés à partir d'un entrepôt de données dé-identifié. La fonction de floutage n'est pas dans cas considérée nécessaire.

Ces deux premiers niveaux sont typiquement utilisés pour la réalisation d'études de faisabilité.

Le niveau 3 permet l'accès à des données individuelles structurées de type LDS, l'accès est réservé à des utilisateurs de confiance.

Le niveau 4 autorise le chercheur à visualiser les données individuelles textuelles dé-identifiées de type LDS. L'IRB doit se prononcer sur la confidentialité des données textuelles ainsi dé-identifiées. Là-aussi, l'accès n'est possible que pour des utilisateurs de confiance.

Ces deux niveaux sont utilisés pour la réalisation d'études observationnelles rétrospectives.

Le niveau 5 donne l'accès aux données identifiantes notamment lorsqu'il est nécessaire de recontacter le patient pour par exemple, participer à un essai clinique.

Il est restreint à certains utilisateurs qualifiés et autorisés par l'IRB. Lorsque dans l'entrepôt les données identifiantes (PHI) sont cryptées, seuls les utilisateurs autorisés sont en mesure de décrypter les données.

Pour les accès de niveaux 3 à 5, pour chaque projet de recherche un sous-entrepôt est créé. Il s'agit d'une partition de l'entrepôt global réalisé à partir d'une requête traduisant les critères de sélection. Il ne contient que les dossiers des patients répondant aux critères de l'étude.

L'accès aux données individuelles ne concerne que le sous-entrepôt. Auparavant le projet doit

avoir reçu l'approbation de l'IRB, et en cas d'accès aux données identifiantes une dérogation à l'obtention du consentement du patient doit avoir été délivrée.

Par ailleurs, les chercheurs reçoivent une formation à la sécurité informatique dans le domaine de la santé. Une validation de cette formation est nécessaire pour les accès du cinquième niveau.

STRIDE :

A Stanford l'entrepôt STRIDE [2] est sous la responsabilité d'une structure universitaire (Stanford Center for Clinical Informatics SCCI). Cette dernière détermine et contrôle les accès. Une attestation de formation au droit numérique est nécessaire à l'utilisation de l'entrepôt par les chercheurs.

Aux trois outils d'exploitation des données de l'entrepôt Stride correspondent des niveaux de confidentialité et des conditions d'accès graduées :

L'outil pour étude de faisabilité produit des résultats agrégés sans identification possible des patients (cohort discovery tool). Il est accessible pour les chercheurs de Stanford sur simple demande auprès de la structure en charge de l'entrepôt STRIDE. Les requêtes réalisées avec le « cohort discovery tool » sont tracées et auditées.

Pour utiliser l'outil de constitution de cohortes (cohort data review tool), le chercheur doit disposer de l'autorisation d'un responsable universitaire chargé de la protection des données (school of medicine privacy officer). Les informations consultées sont celles strictement nécessaires pour le projet de recherche. Un premier cadre d'utilisation correspond à un travail préparatoire à la mise en place d'un protocole de recherche, il ne nécessite pas l'accord d'un IRB, seules des données dé-identifiées sont accessibles. Un deuxième cadre d'utilisation est celui d'un projet de recherche déjà validé par un IRB. Ce dernier délivre une autorisation à l'accès aux données individuelles dé-identifiées ou non en fonction du type de projet.

Il en est de même concernant l'outil d'extraction de données (data extraction tool) [13]. Dans l'éventualité où des données identifiantes sont communiquées, le projet doit avoir reçu de l'IRB une dérogation à l'obtention du consentement du patient.

Roogle :

Dans le cadre de l'entrepôt Roogle [14] mis en oeuvre au CHU de Rennes, des règles pour la gestion des accès ont été établies en fonction du type d'étude réalisée sur l'entrepôt [47]. Il est utilisé par les professionnels de santé du CHU ainsi que par les acteurs étant amené à traiter l'information (techniciens d'information médicale, Assistants de recherche clinique, Techniciens de recherche clinique) sous la responsabilité d'un médecin de l'établissement.

Pour la réalisation d'études de faisabilité ou la production de tableaux de bord, l'exploitation des données est opérée par l'unité hospitalo-universitaire responsable de l'exploitation de l'entrepôt (Unité fouille de données - UFD) qui transmet au demandeur des décomptes et des agrégats floutés en cas de petits effectifs.

Pour une étude nécessitant le recueil de données individuelles, par exemple constitution d'une cohorte ou pré-recrutement dans un essai clinique, l'unité effectue une première requête sur l'entrepôt global à partir des critères d'éligibilité (par exemple une pathologie, une tranche d'âge, un ensemble d'unités fonctionnelles). Un sous entrepôt (Datamart) est constitué avec les documents résultats de la requête. Il est mis à disposition de l'investigateur de l'étude afin qu'il

repère sa cohorte et puisse recueillir et exploiter les données. A ce stade et parce que l'étude ne nécessite pas de connaître l'identité des patients, les données demeurent dé-identifiées.

Pour une étude où l'investigateur a besoin de recontacter les patients (par exemple pour proposer leur participation à un essai clinique), l'identité des patients sélectionnés est rétablie par l'UFD. Les patients qui se sont opposés à être recontactés sont alors exclus.

4.2 L'information des patients et leur accord à la réutilisation de leurs données de santé

Aux Etats unis, dans les centres hospitaliers où l'entrepôt I2b2 a été implanté, une note d'information est remise au patient [40]. Elle stipule que les données de son dossier médical seront exploitées pour la recherche sous forme dé-identifiée.

Il est à préciser que le patient signe un simple accusé de réception mais pas une autorisation. En effet, dans le cas où le traitement concerne des données dé-identifiées, l'accord du patient n'est pas requis pour la réutilisation de ses données de santé.

A Stanford, l'entrepôt STRIDE a été autorisé par l'IRB avec une dérogation à l'obtention du consentement du patient pour l'intégration de ses données de soins dans l'entrepôt [2].

Sur les deux sites, l'accès à des données identifiantes nécessite soit l'accord du patient, soit une dérogation au consentement du patient obtenu auprès d'un IRB.

Au CHU de Rennes il est prévu de joindre une note d'information avec le livret d'accueil selon le modèle du guide de bonnes pratiques pour assurer la confidentialité des données de soins lors de leur réutilisation pour des traitements biostatistiques [30].

Au Canada, un rapport a été commandité en 2009 par les autorités sur l'utilisation des données du dossier patient pour la recherche en santé [48]. Des propositions d'évolution du cadre réglementaire relatif à l'information du patient sont avancées. L'information serait collective et diffusée par les établissements ou les diverses structures de santé fréquentées par les patients, sous forme de brochures, affiches ou vidéo.

Le type d'accord du patient pourrait être conditionné par le domaine d'utilisation des données.

Par exemple, l'accord du patient est non requis pour des besoins de santé publique concernant la surveillance sanitaire, la pharmacovigilance et la matériovigilance.

Le recueil de l'opposition éventuelle du patient est requis pour les études relatives à la qualité des soins. Enfin le recueil de l'accord du patient à être recontacté est nécessaire pour les essais cliniques ou les études épidémiologiques comportant des inclusions.

Pour les biobanques, l'accord formalisé du patient est nécessaire dès lors qu'un prélèvement est réalisé et que les données le concernant sont recueillies.

Le recueil du choix du patient pourrait se faire au travers d'un portail web, dans l'esprit du Dossier Médical Personnel. Il serait également utilisé pour le retour d'informations au patient (indication des projets de recherche où ses données sont utilisées, résultats des études).

S'il est nécessaire de contacter le patient, cela se passe en deux temps. Dans un premier temps le patient est contacté par une personne habilitée à accéder à son dossier afin de savoir s'il est d'accord pour participer au projet de recherche. Dans l'affirmative, un membre de l'équipe de recherche le contacte pour recueillir son consentement éclairé.

5 Discussion

5.1 Une procédure d'information et de recueil de l'opposition difficile à mettre en oeuvre

La réutilisation des données de santé ne peut se faire sans en avoir informé le patient.

Or, la finalité précise d'un traitement, les données en faisant l'objet, les destinataires ne sont pas connus au moment du recueil initial. L'information individuelle des patients est difficile à mettre en oeuvre a posteriori.

Cette information ne peut être que générale et collective. Elle devrait figurer dans le livret d'accueil remis à tout patient hospitalisé. Au delà, à l'instar des propositions canadiennes, une campagne publique d'information, par exemple par voie de presse, pourrait être mise en place.

Sur les plans éthique et réglementaire, pour les études non interventionnelles, il est également nécessaire de recueillir auprès du patient son éventuelle opposition à réutiliser ses données ou à être recontacté [49]. L'exploitation des EDBM rentre dans ce cadre.

Pour en tenir compte au moment du traitement des données, cette opposition devrait être enregistrée dans le Système d'information de l'établissement.

En cas de recontact du patient, il serait préférable qu'il se fasse par un praticien intervenant dans sa prise en charge.

Le Dossier médical Personnel (DMP) pourrait être le lieu d'information et d'enregistrement de l'opposition du patient. L'opposition serait ensuite être transmise aux Système d'Information des établissements de santé. Les dernières orientations limitant le périmètre du DMP aux affections chroniques rendent inopérante cette solution. Par ailleurs, le recueil de l'opposition peut être source de biais dans l'exploitation des données [50]. La représentativité de la population n'est pas assurée. En effet, les patients qui s'opposent à la réutilisation de leurs données peuvent avoir un profil particulier.

5.2 Problématiques liées à la dé-identification

5.2.1 *Des méthodes de dé-identification à améliorer*

La réutilisation des données de santé serait possible sans formalité particulière, dans la mesure où elles sont anonymes. Or, une anonymisation automatique stricte est quasiment impossible à garantir. Cependant, les techniques de dé-identification permettent de s'en approcher. Elles devraient être systématiquement mises en oeuvre pour permettre la réutilisation des données patient pour la recherche. Le traitement simple qui consiste à occulter les traits d'identité à partir des informations administratives du SIH n'est pas suffisant. En effet, Grouin [32] dans son travail mené sur l'anonymisation de compte-rendus de cardiologie retrouve 25% de documents avec persistance de données personnelles en utilisant cette technique. Les méthodes de dé-identification des données textuelles apportent un gain supplémentaire. Les algorithmes existants ont démontré de bonnes performances [34, 35]. Ils restent toutefois à améliorer, voire à combiner, pour la langue française [32]. Des travaux récents effectués à partir de MEDINA tendent à privilégier les méthodes d'apprentissage [51].

Il existe aussi un risque dégradation de l'exploitabilité de l'EDBM en raison de la suppression

possible de données d'intérêt en cas de surfiltrage, d'une moins bonne lisibilité et compréhensibilité des textes du fait du cryptage ou encore en l'absence de conservation de la temporalité des événements [33].

5.2.2 Un besoin de certification des outils de dé-identification

A partir de quand peut-on considérer que la probabilité de remonter à l'identité du patient est suffisamment faible ? Cette question n'a pas reçu de réponse univoque. Il est proposé de demander au chercheur un engagement à ne pas essayer de ré-identifier le patient. A l'heure actuelle, concernant la dé-identification de données textuelles, aucune méthode n'a été validée ni certifiée à un échelon national. Par exemple, aux Etats Unis, l'algorithme de dé-identification dans l'entrepôt I2B2 est validé par l'IRB local. En France, dans la mesure où la dé-identification des dossiers médicaux serait la condition à leur utilisation pour la recherche, des directives plus précises sur les outils à utiliser devraient être émises.

5.2.3 Le risque de réidentification

En cas d'études longitudinales où sont analysées les trajectoires de soins du patient, le risque de réidentification est accru [33]. La combinaison d'informations non identifiantes prises isolément peut conduire à une description unique reliée à un patient si on a connaissance par ailleurs de certaines informations le concernant. Cette question a été évoquée dans le rapport de J. L Bras sur la gouvernance et l'utilisation des données de santé, lorsque sont examinées les conditions d'ouverture de la base SNIIRAM [52]. Cette problématique existe dans les EDBM puisque ceux-ci contiennent tous les épisodes de soins hospitaliers. Elle est d'autant plus prégnante pour les EDBM mettant en relation les données cliniques et les données génomiques. Certaines séquences sont d'une part liées à des maladies et sont d'autre part considérées comme identifiantes car caractérisant un seul patient, des informations médico-administratives pourraient permettre de remonter à l'identité du patient. Dans un article publié par *Science* en 2013 [53] les auteurs ont démontré qu'il était possible de ré-identifier un patient à partir d'informations contenues dans des bases publiques, une base de séquençage du génome et une base de généalogie. Avec l'avènement du big data et le croisement de données multi-sources peut-on encore aujourd'hui garantir l'anonymat lors des traitements des données ?

5.2.4 Un chaînage des données patient souvent nécessaire

Pour permettre des croisements avec d'autres sources de données externes (Base SNIIRAM par exemple) un numéro unique anonyme ou pseudonyme doit être généré. Des solutions ont déjà été proposées pour les études épidémiologiques multicentriques ou le recoupement de plusieurs études [54, 55]. Elles mettent en œuvre des fonctions de hachage et impliquent un tiers de confiance. La question de l'utilisation du NIR⁶ et de son encadrement par un tiers de confiance est actuellement débattue au sein de la commission sur l'ouverture de la base SNIIRAM. La CNIL y serait favorable et pourrait par ailleurs émettre des recommandations de bonnes pratiques sur les techniques d'anonymisation des données [56]. Contrairement à l'Allemagne

⁶ Toute personne née en France métropolitaine et dans les départements d'outre-mer (DOM) est inscrite au répertoire national d'identification des personnes physiques (RNIPP). L'inscription à ce répertoire entraîne l'attribution du numéro d'inscription au répertoire (NIR)

[37] ou le Royaume Unis [23-44], la France ne dispose pas à ce jour de service de pseudo-anonymisation accessible pour les chercheurs. Une procédure simple pour la création du numéro anonyme serait utile. Ceci permettrait de croiser les données avec d'autres sources comme des bases de données de recherche clinique à des fins de médecine translationnelle [57] ou de pharmacovigilance ou encore avec des bases de données épidémiologiques ou les registres.

5.2.5 Répondre au besoin de réidentification

Lorsqu'il est nécessaire de recontacter le patient, par exemple pour une étude prospective ou une inclusion dans un essai clinique, un lien doit être conservé avec les données d'identité. Dans un entrepôt, il est possible également de crypter les données identifiantes. Les clefs sont conservées à part. La gestion des liens permettant de remonter à l'identité ou des clefs de cryptage pourra être assurée par un tiers de confiance interne à l'établissement.

5.3 Aspects éthiques et réglementaires

Plusieurs questions se posent. L'accès à des données à caractère personnel nécessite-t-il l'accord du patient ? Le patient peut-il s'opposer à la réutilisation de ses données recueillies initialement pour soins ? L'intérêt individuel primerait-il sur l'intérêt collectif ?

K.W Goodman [58] estime que si les médecins sont investis d'une mission de santé publique (ex : déclaration de maladies obligatoires, causes de décès), les patients doivent aussi y contribuer.

Il semble cependant licite, qu'en vertu du respect de la vie privée [29] le patient ait un droit de regard sur l'accès à ses données personnelles.

Par contre la réutilisation de données ne permettant pas d'identifier le patient ne devrait pas être soumise à condition. On rappelle que pour la base SNIIRAM, s'agissant de données non identifiantes, l'accord du patient n'est pas requis pour son exploitation.

La directive européenne 95/46/CE est en cours de révision, et devrait être remplacée par un règlement sur la protection des données [59]. La mise en œuvre d'une pseudonymisation lors de traitement à caractère personnel notamment dans le cadre de la recherche médicale serait la règle.

Il faudra mieux définir la place de l'exploitation des EDBM dans ces dispositifs juridiques.

Concernant les démarches auprès de la CNIL, la mise en œuvre d'un EDBM au sein d'un établissement de santé nécessite d'effectuer une déclaration normale.

Ces aspects sont actuellement abordés dans la communauté d'informatique médicale à l'échelle internationale. Au niveau européen, trois principes sont retenus : la nécessité de définir une réglementation et un encadrement et d'informer pleinement les patients sur la réutilisation de leurs données de santé [61]. L'implication des citoyens et des patients est considérée comme obligatoire. La question d'un tiers de confiance pour le contrôle des accès aux données et la mise à disposition de celles-ci est posée.

5.3.1 Aspects déontologiques

Sur le plan déontologique, quel est le droit de regard des praticiens ayant produit les informations sur leur réutilisation ?

Le dossier médical est en fait une copropriété entre l'établissement, le patient et le médecin qui a produit les informations. Le médecin et l'établissement en sont dépositaires, le patient dispose

d'un droit d'accès [60].

Il est de bonne pratique d'informer et de s'assurer de l'accord des praticiens avant réutilisation des informations dont ils sont à l'origine.

Une évolution à l'échelle européenne du rôle du patient dans la gestion de ses données de santé pourrait avoir lieu, il en deviendrait le propriétaire et en contrôlerait ainsi l'accès [61].

5.3.2 Positionnement des patients

Les études menées sur ce sujet tendent à montrer que les patients sont réticents à la communication des données les concernant, même dé-identifiées.

Cette réserve concerne particulièrement l'accès ou la transmission de leurs données aux compagnies d'assurances ou à l'industrie pharmaceutique et de manière générale, chaque fois qu'il y a un intérêt commercial en jeu [62, 63].

Une majorité d'entre eux sont cependant favorables à la mise à disposition de leurs données aux équipes de recherche académiques et considèrent que le bénéfice l'emporte sur le risque d'atteinte à la vie privée. Leur principale crainte concerne les failles de sécurité des systèmes conduisant à des accès illicites [62, 63].

Dans une étude récente le choix des patients s'oriente vers le recueil de l'opposition et non du consentement à la réutilisation de leurs données de santé [63].

En France, le Collectif Inter associatif Sur la Santé (CISS) s'est prononcé pour le libre accès aux données publiques de santé anonymes constituées essentiellement par la base SNIIRAM [64].

D'un autre côté, les sites internet à destination des patients se développent comme le site américain, PatientLikeMe, les sites français Doctissimo ou Carenity. Ces sites emportent l'adhésion des patients et les positionnent de leur avis même comme partenaire dans l'évolution des connaissances médicales [66]

A leur initiative, les patients échangent des informations sur leur maladie et leur prise en charge. Les données ainsi collectées sont ensuite analysées (par exemple pour la recherche d'effets secondaires [67]) et peuvent être réutilisées pour la recherche clinique à titre payant. Les patients peuvent alors être sollicités, s'ils y consentent, pour répondre à des questionnaires ou participer à des études.

5.4 Sécurité des données dans les EDBM

A notre sens, la dé-identification des données dans les EDBM apporte de la confiance pour les patients et pour les utilisateurs lors de l'exploitation au sein des établissements en garantissant au mieux la confidentialité des données de santé.

Un risque de réidentification subsiste cependant. Certaines fonctionnalités renforcent ce risque, notamment les interrogations sur une courte période, la visualisation du parcours de soins du patient, les interrogations multi-critères sur des sous entrepôts.

On peut restreindre l'accès à des personnes que l'on considère de confiance, formées à la sécurité des informations de santé. Afin d'exploiter ces nouveaux outils dans toutes leurs fonctionnalités, un compromis devra être trouvé entre confidentialité et intérêt pour la recherche médicale.

L'avis du patient devra être pris en compte dès que le risque de réidentification est trop important.

La traçabilité des accès est assurée par étude. Elle concerne les utilisateurs et les requêtes effectuées.

Elle est renforcée lorsque les interrogations sont réalisées par les utilisateurs eux-mêmes, la multiplication des requêtes et des critères augmentant le risque de réidentification.

Sur le plan déontologique, la transparence est la règle, les études réalisées à partir de l'EDBM sont portées à la connaissance des médecins de l'établissement.

5.5 Propositions d'une organisation et d'une procédure pour l'exploitation des EDBM

Suite à l'analyse de la littérature des propositions peuvent être émises pour l'exploitation des EDBM dans le respect de la réglementation française et adaptés à l'organisation des établissements.

5.5.1 Les structures parties prenantes existantes

La responsabilité du traitement revient à la Direction de l'établissement de santé qui mettra en place (via la Direction des Systèmes d'Information) les mesures pour assurer la sécurité du système et effectuera la déclaration de l'EDBM pour une exploitation interne à l'établissement.

La Direction de la recherche porte à la connaissance des médecins et des soignants de l'établissement les modalités d'accès à l'EDBM.

La Direction des relations avec les usagers est l'interlocuteur des patients en cas d'opposition de leur part à la réutilisation de leurs données de santé.

La cellule communication est impliquée dans l'information du patient.

Le Directoire inscrit la mise en œuvre de l'EDBM dans le projet d'établissement et valide la politique d'exploitation de l'EDBM.

La Commission Médicale d'Etablissement informe les médecins et les soignants de l'établissement de la mise en œuvre de l'EDBM.

Le Département d'Information médicale s'assure du respect de la confidentialité des traitements lors de l'exploitation de l'EDBM.

Le Comité d'éthique contribue à définir les modalités d'exploitation de l'entrepôt respectant la protection des patients. En particulier ces questions deviendront cruciales avec l'intégration et le traitement de données OMICS, ou de biomarqueurs individuels.

La Commission des relations avec les usagers et la qualité de la prise en charge (CRUQPC) pourrait être un relais avec les associations de patients afin de valider les procédures d'information et de recontact. La CRUQPC est d'ailleurs amenée à évoluer tant sur sa composition avec une plus forte représentation des usagers, que sur ses missions notamment sur la garantie des droits des usagers [65].

5.5.2 Des structures dédiées à mettre en place

Un Comité de régulation de l'Exploitation des Données Patient (CREDoP) équivalent IRB

Il comprendra des représentants du corps médical, du DIM, du comité d'éthique, des patients, des directions de la recherche, de la Direction Informatique.

Ses missions consistent à :

valider le niveau d'accès à l'entrepôt et le traitement des données par étude
s'assurer que les règles déontologiques sont bien suivies,
vérifier que les autorisations réglementaires nécessaires ont été obtenues,
informer la communauté médicale des projets en cours dans un but de transparence.

Une Structure opérationnelle : le Centre de Données Cliniques (CDC)

Cette structure rassemble des acteurs compétents dans les domaines du traitement de l'information, en particulier médecins spécialistes en informatique médicale, data managers, techniciens de l'information, informaticiens, statisticiens.

La mission des CDC se déclinent ainsi:

Il développe et/ou maintient l'entrepôt de données, les outils d'exploitation et les bases de données patient en lien avec la recherche

Il instruit les demandes pour le CREDoP

Il réalise les extractions et le traitement des données et en fournit les résultats

Il fournit l'accès aux données patient individuelles, dé-identifiées ou non en fonction des demandes.

Il produit les indicateurs relatifs aux traitements et aux accès (reporting)

Il contribue avec le CREDoP à la rédaction de la politique d'exploitation de l'EDBM

5.5.3 Documents cadres

Politique d'exploitation et droit d'accès

Elle indiquera les modalités d'organisation des accès et d'affectation des droits d'accès.

Il faudra définir par qui sont réalisées les interrogations de l'EDBM, personnel d'une structure dédiée (par exemple de type CDC) ou médecins praticiens eux-mêmes.

Les cas d'usage sont définis : étude de faisabilité, pre-screening, constitution de cohortes. Ils déterminent le type de données visibles pour l'utilisateur, agrégées dans le premier cas, individuelles dans les autres cas, structurées ou textuelles selon le type de requête, nominatives ou dé-identifiées. La demande d'accès à des données nominatives doit être justifiée.

Des principes sont édictés concernant le périmètre des données interrogeables et consultables.

Les accès sont établis pour les professionnels de santé senior de l'établissement, les autres utilisateurs (interne, ARC, TEC) interviennent sous la responsabilité d'un professionnel de santé senior habilité à accéder au dossier du patient de par sa relation de soins avec les patients ou de par une mission confiée par l'établissement. Ce dernier est le plus souvent le responsable de l'étude ou l'investigateur principal.

Les accès à l'entrepôt se font sous son contrôle dans son service ou au sein de la structure chargée de l'exploitation de l'entrepôt.

Les utilisateurs n'accèdent par défaut qu'à des données dé-identifiées.

Pour les interrogations multi-services ou transversales sur l'établissement, la transparence est la règle avec une information des responsables médicaux concernés ou de la communauté médicale ou soignante de l'établissement.

Formulaire de demande d'accès

Il comprend les renseignements suivants : Nom du responsable de l'étude, type d'usage (faisabilité, pre-screening, repérage population, extraction données), type d'étude et objectif, justification pour la demande d'informations nominatives (recontact du patient, accès au dossier patient), accord des responsables médicaux pour les interrogations multi-services, autorisations des autorités compétentes jointes (études interventionnelles ou multicentriques), critères d'interrogation, professionnels de l'établissement utilisateurs de l'entrepôt pour cette étude (nom et fonction), période de l'étude, type de données, dans le cas d'extraction de données les informations souhaitées.

Charte d'utilisation

Tout utilisateur de l'EDBM s'engage à la respecter. Par là même il s'engage à respecter la confidentialité des informations auxquelles il a accès, à protéger les fichiers qu'il constitue et à les déclarer auprès du Correspondant Informatique et Libertés de l'établissement en cas de traitement de données à caractère personnel. Il s'engage également à ne pas essayer de réidentifier le patient.

Note d'information patient

La note d'information au patient explicite les besoins de réutilisation des données du dossier médical pour la recherche. Elle précise que le patient peut s'y opposer et que cela en l'occurrence ne modifie pas la qualité de sa prise en charge. Enfin, que les analyses sont réalisées de façon confidentielle dans le respect du secret médical.

5.5.4 Procédures

Procédure auprès des patients

La note d'information du patient est insérée dans le livret d'accueil.

L'opposition éventuelle du patient est enregistrée dans le système d'information pour être prise en compte dans l'exploitation de l'entrepôt. Le retour aux données identifiantes n'est alors pas possible.

Procédure auprès des utilisateurs

Il est souhaitable que les professionnels, utilisateurs de l'entrepôt soient sensibilisés à la sécurité des informations de santé et formés à l'utilisation des outils d'exploitation. Ces aspects devraient faire l'objet d'une formation initiale universitaire (par exemple dans le cadre de la certification C2I - 2 métiers de la santé) ou continue.

Circuit d'exploitation

Les demandes d'exploitation de l'entrepôt sont adressées au CDC qui joue le rôle de guichet unique. Elles y sont instruites sur le plan de leur faisabilité puis sont soumises à l'approbation du CREDoP. A partir des critères de sélection des patients, une requête est réalisée par le CDC. Un

sous entrepôt est constitué. Il est mis à disposition de l'équipe demandeuse sur une période déterminée. Le CDC peut apporter son appui méthodologique et statistique sur l'exploitation du sous entrepôt.

Avis CREDoP

L'avis du CREDoP concerne essentiellement les études transversales multiservices, ou multicentriques ou issues d'une demande extérieure à l'établissement.

L'avis est donné sur une période d'exploitation précise. L'accès est délivré sous la responsabilité d'un professionnel de santé interne à l'établissement.

Le CREDoP peut ne pas valider une demande si tous les prérequis ne sont pas réunis. Il doit informer le demandeur sur la raison du refus.

Procédure de reporting

La traçabilité des demandes ainsi que des interrogations et des consultations des sous-entrepôts doit être assurée et fait l'objet de rapports d'activité électroniques à destination du CREDoP, des services qui participent à l'étude, et des utilisateurs.

Chaque service possède un tableau de bord recensant l'ensemble des études (internes et externes au service), les utilisateurs et les données patients consultées du service.

5.5.5 Protection des données au sein de l'entrepôt

La consultation de l'ensemble des données de l'EDBM n'est pas possible, l'utilisateur ne peut voir que les informations d'un sous-ensemble de l'entrepôt résultat d'une requête. Une partition de l'entrepôt par service peut être prévue. Elle permet aux professionnels de santé appartenant à ce service d'exploiter les données des patients dont ils ont eu la charge.

La désidentification concerne à minima les traits d'identité des patients. Elle sera étendue de préférence à l'ensemble des noms et prénoms (professionnels de santé, famille), aux noms des établissements et services, dates, adresses (postales, électroniques), numéros (patient, séjour, établissement, service, téléphone). Les dates de naissance seront remplacées par l'âge, les autres dates par des dates fictives avec conservation de la temporalité. Les autres valeurs pourront être remplacées par des pseudonymes.

En fonction des cas d'usage, le retour aux données d'identification doit être possible. Un numéro patient sous forme de pseudonyme peut être créé pour permettre ensuite des retours aux données sources, par exemple pour des compléments d'informations sur un cas.

D'autre part des restrictions sont apportées lors des interrogations et la présentation des résultats. Il n'est pas possible d'interroger sur les traits d'identité d'un patient ou d'un professionnel (sauf à la demande de ce dernier). Hors besoins spécifiques (exemple : étude un jour donné) les interrogations sur une date précise ou sur un code postal ou code commune ne sont pas autorisées. L'affichage des statistiques agrégées respectent la borne de 10 patients, en deçà l'effectif est flouté et l'accès aux données individuelles n'est pas possible hors justification (ex : maladies rares).

6 Conclusion

La réexploitation secondaire des données patient représente actuellement un enjeu majeur pour la recherche en santé. Les EDBM contribuent à répondre à ce besoin et sont en pleine évolution : intégration de données OMIC, fédération d'entrepôts de données dont l'exploitation multicentrique devient nationale voire internationale, intégration ou couplage avec des données issues de bases nationales (comme la SNIIRAM) ou des données non médicales (par exemple environnementales), intégration des données des essais cliniques.

Au total nous connaissons une évolution technologique sans précédent, qui permet d'exploiter des données de santé multidomaines et à caractères personnel de façon multi échelle. A notre sens, la mise en œuvre de ces technologies au sein des établissements de santé doit s'accompagner d'une démarche claire et pragmatique. Cette démarche doit permettre d'encadrer strictement l'usage de ces outils sans toutefois entraver la recherche. C'est le sens des propositions que nous formulons dans ce travail.

Nous avons vu que sur le plan national le cadre juridique et les règles de bonnes pratiques éthiques et déontologiques doivent évoluer pour s'adapter à ce nouveau paradigme.

Plusieurs rapports parlementaires soulignent le besoin d'évolution des règles de protection des données dans les grandes bases biomédicales. Là aussi, il sera nécessaire de décliner sur le plan opérationnel les grands principes qui auront été définis et de les confronter à la réalité du terrain.

Références

- [1] LOI n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel et modifiant la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
- [2] Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE--An integrated standards-based translational research informatics platform. *AMIA Annu Symp Proc AMIA Symp AMIA Symp*. 2009;2009:391-395.
- [3] Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). *J Am Med Inform Assoc*. 26 févr 2010;17(2):124-130.
- [4] Prokosch HU, Ganslandt T. Perspectives for medical informatics. Reusing the electronic medical record for clinical research. *Methods Inf Med*. 2009;48(1):38-44.
- [5] Soto-Rey I, Bache R, Dugas M, Fritz F. Query engine optimization for the EHR4CR protocol feasibility scenario. *Stud Health Technol Inform*. 2013;192:1080.
- [6] Cuggia M, Bayat S, Garcelon N, Sanders L, Rouget F, Coursin A, et al. A full-text information retrieval system for an epidemiological registry. *Stud Health Technol Inform*. 2010;160(Pt 1):491-495.
- [7] Osmont M-N, Cuggia M, Polard E, Riou C, Balusson F, Oger E. [Use of the PMSI for the detection of adverse drug reactions]. *Thérapie*. août 2013;68(4):285-295.
- [8] Campillo-Gimenez B, Garcelon N, Jarno P, Chaplain JM, Cuggia M. Full-text automated detection of surgical site infections secondary to neurosurgery in Rennes, France. *Stud Health Technol Inform*. 2013;192:572-575.
- [9] Bahl V, McCreddie SR, Stevenson JG. Developing dashboards to measure and manage inpatient pharmacy costs. *Am J Health-Syst Pharm AJHP Off J Am Soc Health-Syst Pharm*. 1 sept 2007;64(17):1859-1866.
- [10] Kohane IS, Churchill SE, Murphy SN. A translational engine at the national scale: informatics for integrating biology and the bedside. *J Am Med Inform Assoc JAMIA*. avr 2012;19(2):181-185.
- [11] Zapletal E, Rodon N, Grabar N, Degoulet P. Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. *Stud Health Technol Inform*. 2010;160(Pt 1):193-197.

- [12] Ganslandt T, Mate S, Helbing K, Sax U, Prokosch HU. Unlocking Data for Clinical Research – The German i2b2 Experience: *Appl Clin Inform.* 30 mars 2011; 2(1):116-127.
- [13] Stanford Center for Clinical Informatics - Stanford School of medicine. Access to Clinical Data - Services <https://clinicalinformatics.stanford.edu/services/clinicaldata.html> (consulté le 16 mars 2014).
- [14] Cuggia M, Garcelon N, Campillo-Gimenez B, Bernicot T, Laurent J-F, Garin E, et al. Roogle: an information retrieval engine for clinical data warehouse. *Stud Health Technol Inform.* 2011;169:584-588.
- [15] US department of health and human services National Institutes of health. HIPAA Privacy Rule and Its Impacts on Research http://privacyruleandresearch.nih.gov/pr_08.asp#8a (consulté le 16 mars 2014)
- [16] [legislation.gov.uk](http://www.legislation.gov.uk). Data Protection Act 1998 <http://www.legislation.gov.uk/ukpga/1998/29/contents> (consulté le 15 mars 2014)
- [17] Claudot F, Alla F, Fresson J, Calvez T, Coudane H, Bonaiti-Pellie C. Ethics and observational studies in medical research: various rules in a common framework. *Int J Epidemiol.* 1 août 2009;38(4):1104-1108.
- [18] Bankert EA. Institutional Review Board: Management and Function. Jones & Bartlett Learning; 2006. 568 p.
- [19] Health and Social Care Information Center . Caldicott Guardians. <http://systems.hscic.gov.uk/infogov/caldicott> (consulté le 18 janvier 2014)
- [20] Commission nationale informatique et libertés. Guide des professionnels de santé Edition 2011. http://www.cnil.fr/fileadmin/documents/Guides_pratiques/CNIL-Guide_professionnels_de_sante.pdf (consulté le 23 septembre 2013).
- [21] Official Journal of the European Communities. Directive 95/46/EC of the european parliament and of the council. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0011:FIN:FR:PDF> (consulté le 23 septembre 2013).
- [22] Malin B, Karp D, Scheuermann RH. Technical and policy approaches to balancing patient privacy and data sharing in clinical and translational research. *J Investig Med Off Publ Am Fed Clin Res.* janv 2010;58(1):11-18.
- [23] Powell J, Buchan I. Electronic Health Records Should Support Clinical Research. *J Med Internet Res.* 14 mars 2005;7(1):e4.

- [24] National Health Service Connecting for health. <http://www.connectingforhealth.nhs.uk/systemsandservices/pseudo/pipwhitepaper.pdf> (consulté le 18 janvier 2014).
- [25] National Health Service Connecting for health. <http://www.connectingforhealth.nhs.uk/systemsandservices/pseudo/ref3deident.pdf> (consulté le 18 janvier 2014)
- [26] Agence technique de l'information sur l'hospitalisation. Présentation Programme de médicalisation des systèmes d'information en médecine, chirurgie, obstétrique (PMSI MCO). <http://www.atih.sante.fr/mco/presentation?secteur=MCO> (consulté le 15 mars 2014)
- [27] El Emam K, Jonker E, Arbuckle L, Malin B. A Systematic Review of Re-Identification Attacks on Health Data. Scherer RW, éditeur. PLoS ONE. 2 déc 2011;6(12):e28071.
- [28] Commission nationale de l'informatique et des libertés. L'état des lieux en matière de procédés d'anonymisation. <http://www.cnil.fr/documentation/fiches-pratiques/fiche/article/letat-des-lieux-en-matiere-de-procedes-danonymisation/> (consulté le 23 septembre 2013)
- [29] Code de la santé publique - Article L1110-4.
- [30] EMOIS 2014. http://emois.org/images/EMOIS2014/soumission_resp.pdf (consulté le 13 avril 2014).
- [31] Meystre SM, Friedlin FJ, South BR, Shen S, Samore MH. Automatic de-identification of textual documents in the electronic health record: a review of recent research. BMC Med Res Methodol. 2010;10(1):70.
- [32] Grouin C. Thèse de doctorat de l'Université Pierre et Marie Curie. Spécialité Informatique Biomédicale. Anonymisation de documents cliniques : performances et limites des méthodes symboliques et par apprentissage statistique. http://tel.archives-ouvertes.fr/docs/00/84/86/72/PDF/these_grouin.pdf (consulté le 22 mars 2014)
- [33] Tamersoy A, Loukides G, Nergiz ME, Saygin Y, Malin B. Anonymization of Longitudinal Electronic Medical Records. IEEE Trans Inf Technol Biomed. mai 2012;16(3):413-423.
- [34] Uzuner O, Luo Y, Szolovits P. Evaluating the State-of-the-Art in Automatic De-identification. J Am Med Inform Assoc. 28 juin 2007;14(5):550-563.
- [35] Deleger L, Molnar K, Savova G, Xia F, Lingren T, Li Q, et al. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. J Am Med Inform Assoc. 2 août 2012;20(1):84-94.

- [36] Kushida CA, Nichols DA, Jadrnicek R, Miller R, Walsh JK, Griffin K. Strategies for de-identification and anonymization of electronic health record data for use in multicenter research studies. *Med Care*. juill 2012;50 Suppl:S82-101.
- [37] Faldum A, Pommerening K. An optimal code for patient identifiers. *Comput Methods Programs Biomed*. juill 2005;79(1):81-88.
- [38] Uzuner O, Sibanda TC, Luo Y, Szolovits P. A de-identifier for medical discharge summaries. *Artif Intell Med*. janv 2008;42(1):13-35.
- [39] Aberdeen J, Bayer S, Yeniterzi R, Wellner B, Clark C, Hanauer D, et al. The MITRE Identification Scrubber Toolkit: Design, training, and assessment. *Int J Med Inf*. déc 2010;79(12):849-859.
- [40] Chazard E, Mouret C, Ficheur G, Schaffar A, Beuscart J-B, Beuscart R. Proposal and evaluation of FASDIM, a Fast And Simple De-Identification Method for unstructured free-text clinical records. *Int J Med Inf*. 7 déc 2013;
- [41] Grouin C, Zweigenbaum P. Automatic de-identification of French clinical records: comparison of rule-based and machine-learning approaches. *Stud Health Technol Inform*. 2013;192:476-480.
- [42] Gateshead Health NHS Foundation Trust. Pseudonymisation Policy. <http://www.qegateshead.nhs.uk/sites/default/files/users/user1/IG08%20Pseudonymisation%20Policy.pdf> (consulté le 23 septembre 2013).
- [43] Health and Social Care Information Centre. Who we are and what we do. <http://www.hscic.gov.uk/whoweare> (consulté le 15 mars 2014).
- [44] Health and Social Care Information Centre. Data Linkage and Extract Service. Disponible sur: <http://www.hscic.gov.uk/dles> (consulté le 15 mars 2014).
- [45] Institut des données de santé. Etudes et recherche - Présentation du SNIIRAM. http://www.institut-des-donnees-de-sante.fr/upload/06_etudes_recherches/01_form/SNIIRAM_Presentation_20131014.pdf (consulté le 16 mars 2014).
- [46] Murphy SN, Gainer V, Mendis M, Churchill S, Kohane I. Strategies for maintaining patient privacy in i2b2. *J Am Med Inform Assoc*. 7 oct 2011;18(Suppl 1):i103-i108.
- [47] Riou C, Cuggia M, Garcelon N. Comment assurer la confidentialité dans les entrepôts de données biomédicaux ? *Colloq Adelf-Émois Système Inf Hosp Épidémiologie*. mars 2012;60, Supplement 1(0):S19-S20.

- [48] Willison DJ. Use of Data from the Electronic Health Record for Health Research - current governance challenges and potential approaches. http://www.priv.gc.ca/information/research-recherche/2009/ehr_200903_e.pdf (consulté le 20 septembre 2013).
- [49] Claudot F, Fresson J, Coudane H, Guillemin F, Demoré B, Alla F. Recherche en épidémiologie clinique : quelles règles appliquer ? *Rev DÉpidémiologie Santé Publique*. févr 2008;56(1):63-70.
- [50] Kho ME, Duffett M, Willison DJ, Cook DJ, Brouwers MC. Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ*. 2009;338:b866.
- [51] Grouin C, Névéal A. De-identification of clinical notes in French: towards a protocol for reference corpus development. *J Biomed Inform*. 2013 Dec 29. pii:S1532-0464(13)00205-0.
- [52] Ministère des Affaires sociales et de la Santé. Rapport sur la gouvernance et l'utilisation des données de santé. http://www.social-sante.gouv.fr/IMG/pdf/Rapport_donnees_de_sante_2013.pdf (consulté le 18 janvier 2014).
- [53] Identifying personal genomes by surname inference. Gymrek M1, McGuire AL, Golan D, Halperin E, Erlich Y. *Science*. 2013 Jan 18;339(6117):321-4.
- [54] Lo Iacono L. Multi-centric universal pseudonymisation for secondary use of the EHR. *Stud Health Technol Inform*. 2007;126:239-247.
- [55] Quantin C, Fassa M, Coatrieux G, Riandey B, Trouessin G, Allaert FA. Linking anonymous databases for national and international multicenter epidemiological studies: A cryptographic algorithm. *Rev DÉpidémiologie Santé Publique*. févr 2009;57(1):e1-e6.
- [56] Dépêche APM du 4 décembre 2014, éditeur. Open data : la CNIL souscrit à un grand nombre de propositions du rapport Bras. 2014.
- [57] Szalma S, Koka V, Khasanova T, Perakslis ED. Effective knowledge management in translational medicine. *J Transl Med*. 2010;8:68.
- [58] Goodman KW. Ethics, information technology, and public health: new challenges for the clinician-patient relationship. *J Law Med Ethics J Am Soc Law Med Ethics*. 2010;38(1):58-63.
- [59] Proposal for a regulation of the european parliament and of the council on the protection of individuals with regard yo the processong of personal data and on the free movement of such data (General Data Protection Regulation). <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0011:FIN:EN:PDF> (consulté le 23 septembre 2013)

- [60] Haute Autorité de Santé. Evaluation des pratiques professionnelles dans les établissements de santé Dossier du patient Réglementation et recommandations. Disponible sur: http://www.has-sante.fr/portail/upload/docs/application/pdf/2009-08/dossier_du_patient_-_fascicule_1_reglementation_et_recommandations_-_2003.pdf (consulté le 18 janvier 2014).
- [61] Geissbuhler A, Safran C, Buchan I, Bellazzi R, Labkoff S, Eilenberg K, et al. Trustworthy reuse of health data: A transnational perspective. *Int J Med Inf.* janv 2013;82(1):1-9.
- [62] Perera G, Holbrook A, Thabane L, Foster G, Willison DJ. Views on health information sharing and privacy from primary care practices using electronic medical records. *Int J Med Inf.* févr 2011;80(2):94-101.
- [63] Hill EM, Turner EL, Martin RM, Donovan JL. « Let's get the best quality research we can »: public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. *BMC Med Res Methodol.* 2013;13:72.
- [64] Dépêche APM du 28 janvier 2013, éditeur. Le CISS réclame un libre accès aux données de santé. 2013.
- [65] Ministère des Affaires sociales et de la Santé. Rapport à la ministre des Affaires sociales et de la Santé : Pour l'an II de la Démocratie sanitaire. [cité 21 mars 2014]. Disponible sur: http://www.sante.gouv.fr/IMG/pdf/Rapport_DEF-version17-02-14.pdf (consulté le 21 mars 2014).
- [66] deBronkart D. How the e-patient community helped save my life: an essay by Dave deBronkart. *BMJ.* 2013 Apr 2;346:f1990.
- [67] Nikfarjam A, Gonzalez GH. Pattern mining for extraction of mentions of Adverse Drug Reactions from user comments. *AMIA Annu Symp Proc.* 2011;2011:1019-26.

Adresse de correspondance

Christine Riou

Département d'Information Médicale CHU de Rennes

35033 Rennes Cedex, France

E-mail : christine.riou@chu-rennes.fr

Place de la technologie Smartphone en santé mentale

Smartphone technology in mental health

Dalal Ben Loubir^{1,2}, Samy Housbane¹, Zeineb Serhier¹, Mohammed Bennani Othmani¹

¹*Laboratoire d'Informatique Médicale, Faculté de Médecine et de Pharmacie, Université Hassan II, Casablanca, Maroc*

²*Laboratoire de Neurosciences et Santé Mentale, Faculté de Médecine et de Pharmacie, Université Hassan II, Casablanca, Maroc*

Résumé

Introduction: Actuellement, on assiste à une croissance considérable dans le nombre d'applications mobiles pour smartphone, conçues pour l'amélioration de la pratique clinique en santé mentale. Cependant, l'application et l'apport de la santé mobile en santé mentale ne sont pas encore évalués. L'objectif de l'étude est de dresser un état des lieux des applications médicales mobiles pour smartphone dédiées à la santé mentale. Méthodologie: Trois ressources d'applications mobiles ont été fouillées, ainsi qu'une revue de la littérature a été établie afin d'en extraire les applications médicales mobiles pour smartphone, qui sont dédiées à la santé mentale. Résultats: La recherche sur les ressources des applications a permis d'extraire 112 applications, dont 13 % applicables en pratique clinique servant plutôt au diagnostic, 5 % applicables en pratique académique et 77 % ont été jugées non appropriées dans l'environnement clinique. Tandis que la revue de la littérature a permis d'extraire 8 articles, dont 5 présentaient des applications médicales mobiles pour smartphone destinées au patient. L'implication médicale dans le développement des applications médicales mobiles était très faible, alors que l'évaluation des applications développées a mis en évidence l'efficacité et l'utilité de celles-ci. Conclusion: La technologie de la santé mobile est nécessaire aujourd'hui plus que jamais pour faciliter et améliorer l'activité professionnelle du médecin.

Abstract

Introduction: Currently, there is a significant growth in the number of smartphone mobile

Articles courts des 15^{es} Journées francophones d'informatique médicale, JFIM 2014, pages 188–197
Fès, Maroc, 12–13 juin 2014

applications, designed to improve clinical practice in mental health. However, the application and the contribution of mobile technology in mental health are not yet evaluated. The objective of the study is to provide an inventory of smartphone mobile medical applications dedicated to mental health. Methodology: Three resources of mobile applications have been searched, and a review of the literature was made to extract smartphone medical applications dedicated to mental health. Results: Research on mobile applications resources has extracted 112 applications, of which 13% were used in clinical practice, 5% were used in academic practice and 77% were considered inappropriate in the clinical environment. In the other hand, the literature review was enabled to extract 8 articles, of which 5 articles showed mobile medical applications dedicated for patient use. In addition, medical involvement in the development of mobile medical applications was very low, while the evaluation of the applications developed demonstrated their effectiveness and usefulness. Conclusion: The technology of mobile health is needed now more than ever to facilitate and improve doctor's professional activity.

Mots-clés : Technologie smartphone ; Santé mobile ; Santé mentale

Keywords: Smartphone technology; Mobile health; Mental health

1 Introduction

L'émergence de la technologie smartphone et son apport sophistiqué en terme de fonctionnement a grandement amélioré le style de vie de ses utilisateurs. De plus, sa nature interactive et sa capacité d'accéder à internet permettent l'accès à des quantités de données illimitées en un temps très court.

La santé mobile ou mSanté est l'utilisation de la technologie mobile, principalement les smartphones et les tablettes, pour accéder à des solutions de santé diverses. Cette technologie s'est rapidement intégrée en domaine médical [1] et a pu améliorer divers aspects de la pratique clinique et de l'éducation en médecine, en servant ainsi en tant qu'outil pour le suivi à distance et le diagnostic des patients, pour la collecte des données en santé, pour la réception d'alertes et de rappels, pour l'éducation et la communication même dans les milieux à ressources limitées, pour l'acquisition d'informations nécessaires en soins de santé et pour l'accompagnement dans la prescription [2–7]. De ce fait, la mSanté est devenue très sollicitée par les professionnels de santé pour l'utilisation professionnelle et le partage.

Bien que la santé mobile intéresse en premier lieu les professionnels en santé, la sphère des patients souhaitent aussi jouir d'applications médicales mobiles. Dans une étude menée en France, près de 94 % des médecins utilisent leurs smartphones à des fins professionnelles et pour marquer la proximité avec leurs patients (les médecins communiquent leur numéro de téléphone) [7], aussi 46 % des mobinautes santé atteints d'une maladie chronique souhaitent que leur médecin leur conseille des applications mobiles de santé [8]. Dans certains pays, notamment en grande Bretagne, on incite les médecins praticiens à recommander des applications médicales pour leurs patients, leur fournissant ainsi un moyen de s'enrichir d'informations nécessaires, concernant le diagnostic et le traitement, et dans le but d'encourager la responsabilité individuelle en santé [9].

2 Etat de l'art

La disponibilité des logiciels open source pour le développement des applications mobiles a non seulement favorisé une prolifération explosive des applications mobiles, mais encore a

permis une croissance considérable dans le nombre des applications médicales mobiles qui visent à promouvoir la qualité de vie des individus, en particulier la santé mentale. Toutefois, les études qui évaluent l'application de la santé mobile en santé mentale sont rares. Parmi elles, certaines avaient conclu cette technologie procure un outil efficace pour l'évaluation, la gestion, la prévention et le traitement des problèmes de santé mentale [10,11], et qu'elle présente un outil facilement manipulable et convenable pour le diagnostic de la santé comportementale [12]. Ainsi la mSanté influence l'évolution des soins en santé mentale. Le but de notre étude est de recenser les applications médicales mobiles disponibles pour smartphone et spécifiques à la santé mentale, en particulier les applications liées aux troubles mentaux les plus réponsus, et qui sont le plus souvent coexistants : la dépression, l'anxiété et le stress [13], et de faire une description des applications médicales impliquées dans la pratique clinique.

3 Matériels et méthodes

La recherche des applications médicales mobiles pour smartphone en rapport avec la santé mentale a été effectuée dans quelques ressources des applications mobiles. Une revue de la littérature a été établie dans la base de données Medline, en utilisant PubMed, pour chercher les articles présentant les applications mobiles pour smartphone.

3.1 Recherche dans les plateformes des applications mobiles

Une recherche préliminaire des applications smartphone liées à la santé mentale, en particulier le stress, la dépression et l'anxiété, a été élaborée dans les ressources des applications mobiles de « app store », « google play store » et « windows store » des systèmes les plus utilisés actuellement « Apple, Android et Windows » durant la première semaine de Mars 2014.

La recherche des applications smartphones a été réalisée en utilisant les termes suivant «mental health», «stress», « depression » et « anxiety ». Les applications, dont le titre comportait un terme ou plus indicatif lié à l'un ou à l'ensemble des termes précédents de la recherche, étaient retenues. Ont été éliminées les applications qui n'appartenaient pas à la catégorie « médicale » et les applications qui ne figuraient pas en anglais ou en français. Les informations à-propos des applications ont été extraites à partir des pages web des applications concernées, avant le téléchargement. Concernant la recherche dans la base « app store » seules les applications iPhone ont été retenues dans notre étude.

Pour toute application retenue, les informations liées à : la langue, la gratuité, la date de création ou de mise à jour, système d'exploitation utilisé, type de l'application et implication des médecins dans le développement des applications, ont été recherchées. Les applications ont été regroupées suivant leurs types d'utilisation et d'utilisateur, en application académique, clinique, destinée au patient et autres.

- Les applications destinées à **la pratique académique**: sont celles dont le contenu est à but éducatif. Ces application sont pour but d'améliorer les connaissances et les performances de leurs utilisateurs qui peuvent être professionnels de santé ou étudiants.
- Les applications destinées à **la pratique clinique** : sont conçues pour les cliniciens afin de leur permettre d'établir une évaluation ou un diagnostic, à l'aide d'échelles de la santé mentale. Ces applications peuvent fournir un moyen de gestion des problèmes de santé mentale, de suivi et de traitement des patients. Les applications qui

implémentent des échelles validées de mesure de la symptomatologie psychologique et qui sont destinées au large public, sont aussi incluses dans cette catégorie.

- Les applications **destinées au patient** : sont des applications d’accompagnement pour celui-ci. Parmi leurs fonctionnalités; elles permettent au patient de faire un autodiagnostic et de suivre la progression de son état psychologique pour mieux comprendre sa situation et ainsi participer activement à son traitement.
- Les applications **non-applicables**: sont les applications destinées au large public, et dont le contenu utilisé est soit inadéquat pour une application en médecine, ou il est d’une source inconnue (utilisation de méthodes de dépistage non reconnues : quizz, questions simples), ces applications ne peuvent être administrées ni en domaine médical clinique ni en domaine académique.

3.2 Revue de la littérature

PubMed a été utilisé afin d’extraire les articles publiés en anglais et en français, en utilisant la combinaison des mots-clés présentés dans le **Tableau 1**. Les critères d’inclusions des articles étaient : 1) identification d’applications mobiles pour smartphone et 2) utilisation de ces applications en santé mentale. La liste des références des articles inclus dans la recension a été consultée afin d’inclure d’éventuels articles.

Tableau 1: Combinaison des variables mots-clés anglais utilisées pour la recherche des articles

Cellular Phone OR Mobile Applications	AND (Mental Health OR Mental Health Services OR Community Mental Health Services OR Community Mental Health Centers OR Mental Disorders OR Diagnostic and Statistical Manual of Mental Disorders OR Psychology OR Psychology, Medical OR Psychology, Applied OR Psychiatry OR Psychiatric Aides OR Psychiatric Department, Hospital OR Emergency Services, Psychiatric OR Brief Psychiatric Rating Scale OR Psychological Techniques OR Psychotherapy, Brief)
---	---

4 Résultats

4.1 Résultats de la recherche dans les plateformes des applications

Un total de 1460 applications était identifié dans un premier temps à travers les recherches dans les trois plateformes d’applications mobiles pour smartphone. Après avoir éliminé les redondances et les applications qui ne répondaient pas aux critères d’inclusions, 112 applications ont été retenues, dont 39 % étaient payantes, 97 % existaient en anglais et 3 % en français. La majorité des applications (77 %) n’étaient destinées ni à la pratique académique ni à la pratique clinique. Parmi les quatre applications destinées au patient, une seule permettait le transfert des résultats du patient vers le médecin traitant. Les caractéristiques des applications sont présentées dans le **Tableau 2**.

L’implication médicale dans le développement des applications a été notée dans 7 % des cas. Alors que dans 57 % des applications, aucune indication sur l’implication d’un médecin ou sur la source du contenu, n’a été mentionnée. (**Tableau 3**).

Tableau 2: Répartition des applications selon leurs caractéristiques

Catégorie	Sous-catégorie	Effectif	Pourcentage(%)
Type d'application	Pratique académique	6	5,4
	Pratique clinique	15	13,4
	Patient	4	3,6
	Non-applicable	87	77,6
Thème traité	Dépression	38	33,9
	Anxiété	19	17,0
	Stress	31	27,7
	Autre thème ou plus d'un thème	24	21,4
Système d'exploitation de l'application	iOS	42	37,5
	Android	44	39,3
	Windows	25	22,3
	iOS & Android	1	0,9

Tableau 3: Implication médicale dans le développement des applications médicales mobiles

Type de l'implication médicale	Nombre (%)
Application développée avec accompagnement des médecins	5 (4,5)
Application développée par des médecins	3 (2,7)
Contenu médical : manuels, articles scientifiques, échelle de mesures validées, protocoles, guides de diagnostic	26 (23,2)
Application qui propose une thérapie (relaxation, visualisation, hypnose, massage régime alimentaire, etc.)	14 (12,5)
Applications dont la source du contenu n'est pas mentionnée	64 (57,1)

Concernant l'information de la date de mise à jour ou de création des applications, 19 (17 %) applications ont été créées ou mises à jours durant les trois derniers mois, 13 (12 %) durant les six derniers mois, 12 (11 %) dans moins d'une année, 33 (29 %) dans moins de deux ans et 24 (21%) applications durant les quatre dernières années. Alors que 11 (10 %) applications n'ont pas mentionnées leur date de création ou de mise à jour.

4.2 Résultats de la revue de la littérature

Au total, 902 articles ont été obtenus. La **Figure 1** présente Les étapes suivies durant la sélection des articles. Huit articles répondant aux critères d'inclusions ont été retenus. La vérification de la liste des références des articles retenus n'a abouti à aucun article supplémentaire.

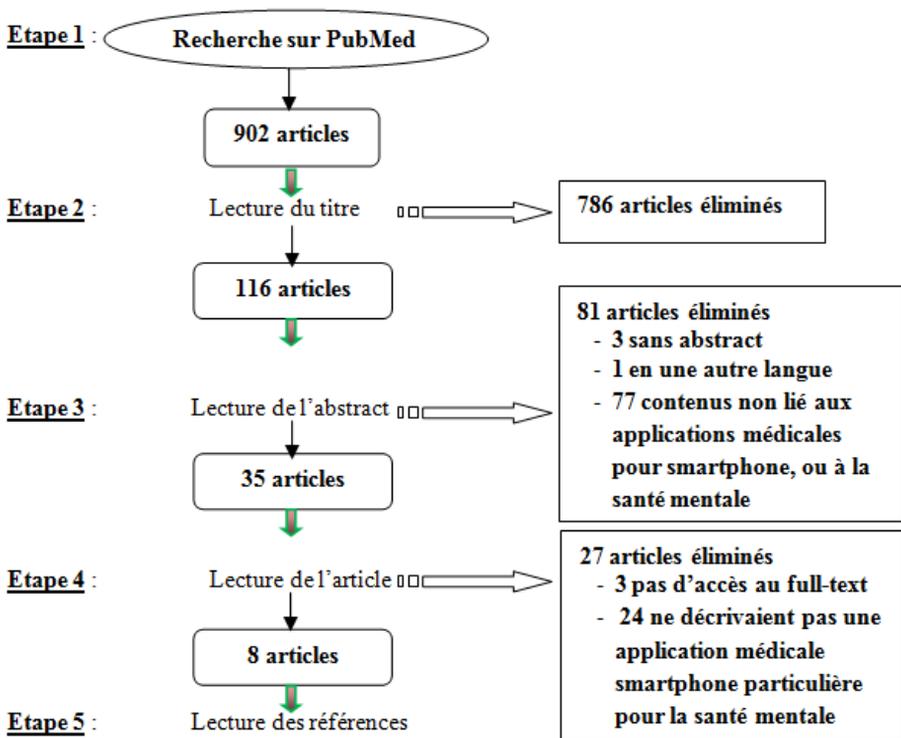


Figure 1: Les étapes élaborées pour la sélection des articles

Le **Tableau 4** résume certaines caractéristiques des articles sélectionnés. La majorité des applications obtenues concernaient la pathologie de la schizophrénie. Cinquante pour cent des applications étaient développées pour Android, 25 % pour Windows et 25 % pour iOS. Les applications étaient conçues en différentes langues incluant l'Anglais, le Suédois et l'Italien et ont été toutes développées durant les 3 dernières années. L'implication médicale des médecins dans le développement de ces applications a été inexistante. Or, l'intégration du contenu médical (échelles de mesure validées pour la dépression et l'anxiété) dans ces applications a été estimée à 25 % [16,17]. Dans les études qui ont évalué l'efficacité de l'application mobile, 50 % avaient une durée de suivie d'un mois, chez un échantillon de la population qui ne dépassait pas 8 individus [14,16]. La majorité des applications (63 %) était destinée aux patients [17–21], et une seule application pouvait être destinée à la pratique clinique [16].

La majorité des applications qui ont été évaluées, se sont montrées efficaces (ex : gain du temps) et ont été bien appréciées par les utilisateurs. De même, les utilisateurs des applications arrivaient facilement à appliquer le contenu thérapeutique applicatif qu'ils ont appris [16]. Il a été aussi montré que l'utilisation des applications est associée significativement à la diminution du niveau de l'anxiété et à l'amélioration des capacités d'adaptation chez les utilisateurs [14].

Tableau 4 : Caractéristiques des études incluses

Pathologie mentale	Auteurs, date	Objectif l'étude	de	Population cible	Durée de suivie
Stress	Villani D. et al. , 2012 [14]	Evaluation de l'efficacité de la thérapie proposée dans l'outil développé		infirmiers	4 semaines
	Pallavicini F. et al. , 2013 [15]	Présentation de l'outil développé		Population générale (âge entre 25 et 60 ans)	5 semaines
Anxiété	Morris M.E. et al. , 2010 [16]	Evaluation de l'efficacité de la thérapie proposée dans l'outil développé		Population générale	4 semaines
Dépression	Ly K.H. et al. , 2012 [17]	Présentation de l'outil développé		patients dépressifs	8 semaines
Autre (Démence, schizophrénie)	Ainsworth J. et al., 2013 [18]	Evaluation de l'outil développé dans le recensement psychologique		patients schizophréniques	2 semaines
	Skillen K. et al., 2012 [19]	Présentation de l'outil développé		patients schizophréniques	Une seule fois
	Brouillette R.M. et al., 2013 [20]	Evaluation de l'outil développé dans le recensement psychologique		population gériatrique	2 fois seulement
	Sposaro E. et al., 2010 [21]	Présentation de l'outil développé		-	-

5 Discussion

Les résultats de la présente étude montrent qu'il existe un manque concernant les applications destinées à la pratique clinique en santé mentale, et qui sont relatives essentiellement au diagnostic et au suivi. La même chose a été notée à propos des applications qui concernent la pratique académique, dont la proportion était très faible. Aussi, la majorité des applications mobiles issues des ressources des applications, était jugée inappropriée dans l'environnement clinique.

Les applications médicales mobiles destinées à la pratique clinique offrent une aide potentielle

au clinicien dans l'amélioration de la précision du diagnostic et dans la pratique médicale, d'autant plus, ces applications facilitent l'accès aux ressources de l'information médicale (revues et livres médicaux) [22]. D'autre part, la mise en place d'applications médicales éducatives portables et facilement accessibles à l'environnement clinique va largement contribuer à l'amélioration de l'apprentissage clinique chez les jeunes médecins [23], d'où l'intérêt d'encourager le développement de telles applications.

D'autre part on a noté que l'intégration du contenu médical, dans les applications, était faible, et que l'implication des médecins dans le développement de celles-ci était presque absente. Les médecins et les professionnels de santé doivent se familiariser avec la technologie de la santé mobile, et doivent être conscient de l'utilité et des risques de l'utilisation des applications médicales avec contenu non validé. Ainsi, il est nécessaire que les médecins et les cliniciens prennent part dans le processus de développement des applications médicales afin d'assurer la qualité de l'outil fourni. Dans l'ensemble des applications médicales étudiées, la majorité n'avait pas mentionné la source des données utilisées, et une grande partie n'a pas été mise à jour pendant des années, ce qui remet en question la qualité et la validité du contenu de ces applications.

L'évaluation, de chaque application médicale mobile développée, est une étape primordiale pour tester son acceptabilité chez l'utilisateur, et surtout pour s'assurer de son utilité et de son efficacité chez un échantillon de la population cible. Or, la revue de la littérature réalisée a montré que seulement la moitié des applications ont fait l'objet d'une évaluation et la majorité des applications évaluées étaient destinées aux patients. Le résultat des différentes évaluations a mis en évidence l'acceptabilité, l'efficacité et l'utilité des applications médicales mobiles, notamment le gain du temps [18], la diminution de l'anxiété et l'amélioration des capacités d'adaptation aux situations [14], de plus l'implication des utilisateurs dans l'utilisations des thérapies mobiles et application de celles-ci dans le changement personnel [16].

La recherche dans les ressources des applications mobiles a montré des proportions des applications développées sous système Android et iOS qui sont proches. Cependant, la littérature concernait beaucoup plus des applications qui étaient développées pour Andoid.

Parmi les limites de l'étude présente, la recherche des applications s'est basée uniquement sur trois ressources d'applications mobiles. D'autre part, la revue de la littérature a porté uniquement sur PubMed, ce qui pouvait exclure d'éventuels articles traitant le même sujet.

6 Conclusion

La promotion de la technologie mSanté en santé mentale est nécessaire aujourd'hui plus que jamais pour faciliter l'activité professionnelle du médecin et pour impliquer le patient dans son traitement.

Références

1. Visser BJ, Bouman J, There's a medical app for that. Site du BMJ Careers. [En ligne]. <http://careers.bmj.com/careers/advice/view-article.html?id=20007104>. [Consulté le 15/04/2014].
2. Klonoff DC. The current status of mHealth for diabetes: will it be the next big thing? *J Diabetes Sci Technol*. 2013; 7(3):749- 58.

3. Braun R, Catalani C, Wimbuch J, Israelski D. Community Health Workers and Mobile Technology: A Systematic Review of the Literature. *PLOS ONE*. 2013 ; 8(6) : e65772.
4. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. SMART MOVE - a smartphone-based intervention to promote physical activity in primary care: study protocol for a randomized controlled trial. *Trials*. 2013; 14: 157.
5. Chang AY, Ghose S, Littman-Quinn R, Anolik RB, Kyer A, Mazhani L, et al. Use of mobile learning by resident physicians in Botswana. *Telemed J E-Health Off J Am Telemed Assoc*. 2012; 18 (1):11- 3.
6. Workman AD, Gupta SC. A plastic surgeon's guide to applying smartphone technology in patient care. *Aesthetic Surg J Am Soc Aesthetic Plast Surg*. 2013; 33(2):275- 80.
7. VIDAL, 2ème baromètre sur les médecins utilisateurs des smartphones en France. Site du conseil national du médecin. [En ligne]. http://www.conseil-national.medecin.fr/sites/default/files/CP_Barometre_VIDAL_CNOM_2.pdf. [consulté le 15/04/2014]
8. Cerisey C, Klein E, Marsico G, Mignon L. A la recherche du ePatient. Site de Patients & Web. [En ligne]. <http://www.patientsandweb.com/wp-content/uploads/2013/04/A-la-recherche-du-ePatient-externe.pdf> . [Consulté le 15/04/2014].
9. Department of Health, Apps to be 'prescribed' for patients by GPs - News stories - GOV.UK. Site du gouvernement du Royaume Uni. [En ligne]. <https://www.gov.uk/government/news/gps-to-prescribe-apps-for-patients--2>. [Consulté le 15/04/2014].
10. Reid SC, Kauer SD, Khor AS, Hearps SJC, Sanci LA, Kennedy AD, et al. Using a mobile phone application in youth mental health - an evaluation study. *Aust Fam Physician*. 2012; 41(9):711- 4.
11. Dennis TA, O'Toole LJ. Mental Health on the Go Effects of a Gamified Attention-Bias Modification Mobile Application in Trait-Anxious Adults. *Clin Psychol Sci*. 2014; 2167702614522228.
12. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. *J Nerv Ment Dis*. 2013;201(11):991- 5.
13. Devane CL, Chiao E, Franklin M, Kruep EJ. Anxiety disorders in the 21st century: status, challenges, opportunities, and comorbidity with depression. *Am J Manag Care*. 2005; 11(12 Suppl):S344- 53.
14. Villani D, Grassi A, Cognetta C, Cipresso P, Toniolo D, Riva G. The effects of a mobile stress management protocol on nurses working with cancer patients: a preliminary controlled study. *Stud Health Technol Inform*. 2012;173:524- 8.
15. Pallavicini F, Gaggioli A, Raspelli S, Cipresso P, Serino S, Vigna C, et al. Interreality for the management and training of psychological stress: study protocol for a randomized controlled trial. *Trials*. 2013; 14:191.
16. Morris ME, Kathawala Q, Leen TK, Gorenstein EE, Guilak F, Labhard M, et al. Mobile therapy: case study evaluations of a cell phone application for emotional self-awareness. *J Med Internet Res*. 2010; 12(2):e10.
17. Ly KH, Carlbring P, Andersson G. Behavioral activation-based guided self-help treatment administered through a smartphone application: study protocol for a randomized controlled trial. *Trials*. 2012; 13: 62.

18. Ainsworth J, Palmier-Claus JE, Machin M, Barrowclough C, Dunn G, Rogers A, et al. A comparison of two delivery modalities of a mobile phone-based assessment for serious mental illness: native smartphone application vs text-messaging only implementations. *J Med Internet Res*. 2013; 15(4):e60.
19. Skillen K-L, Chen L, Nugent CD, Donnelly MP, Solheim I. A user profile ontology based approach for assisting people with dementia in mobile environments. *Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Conf*. 2012; 2012:6390- 3.
20. Brouillette RM, Foil H, Fontenot S, Correro A, Allen R, Martin CK, et al. Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. *PloS One*. 2013; 8(6):e65925.
21. Sposaro F, Danielson J, Tyson G. iWander: An Android application for dementia patients. *Conf Proc Annu Int Conf IEEE Eng Med Biol Soc IEEE Eng Med Biol Soc Conf*. 2010; 2010: 3875-8.
22. Payne KFB, Wharrad H, Watts K .Smartphone and medical related App use among medical students and junior doctors in the UnitedKingdom (UK): a regional survey. *BMC Med Inform Decis Mak*. 2012 30; 12:121.
23. Vinay KV, Vishal K. Smartphone applications for medical students and professionals. *NUJHS*. 2013;1(3): 2249-7110.

Adresse de correspondance:

Dalal Ben Loubir, dallouber@gmail.com

Un outil de visualisation de classifications et d'intégration de données phénotypiques et génétiques pour faciliter le codage des maladies rares

Coding rare diseases in health information systems: a tool for visualizing classifications and integrating phenotypic and genetic data

Rémy Choquet^{1,2}, Yannick Fonjallaz¹, Albane de Carrara¹,
Meriem Maaroufi^{1,2}, Pierre-Yves Vandenbussche², Ferdinand Dhombres^{2,3*},
Paul Landais^{1,4*}

¹Banque Nationale de Données Maladies Rares, Hôpital Necker Enfants Malades, Assistance Publique Hôpitaux de Paris, Paris, France. ²INSERM, U1142, LIMICS, Sorbonne Universités, UPMC Paris 6, et Université Paris 13, Villetaneuse, France. ³Unité de Diagnostic Prénatal et d'Échographie et Centre Pluridisciplinaire de l'Est Parisien, hôpital Armand Trousseau, UPMC et AP-HP, Paris, France. ⁴Montpellier University, EA2415 & BESPIM, Carêmeau University Hospital, Nîmes, France. *Ces deux auteurs ont contribué de manière analogue à ce travail.

Résumé

Pour diagnostiquer et coder une maladie rare (MR), il est nécessaire de caractériser des informations multiples, y compris génotypique et phénotypique. Aujourd'hui, un obstacle à l'activité de codage des diagnostics MR est le manque de consolidation de ces informations dispersées dans différentes bases de connaissances telles qu'Orphanet, OMIM ou HPO. Nous avons développé l'application web LORD (Linking Opendata for Rare Diseases), en offrant une vue intégrée de 8.336 maladies rares et groupes de maladies liés à plus de 12.500 signes et 3.000 gènes. LORD offre une fonctionnalité de navigation contextuelle (médicale) dans les relations entre les groupes de maladies, les maladies, les signes et les gènes. Il est développé sur un ensemble d'interfaces de programmation d'application (APIs) permettant son intégration au sein de systèmes d'informations propriétaires. Il est dédié aux 131 centres français de référence MR et aux 501 centres de compétences, mais aussi aux départements d'information médicale pour coder les diagnostics MR dans les systèmes d'information de santé.

Abstract

Establishing the diagnosis and coding for a rare disease (RD) needs to characterize multiple information including patient's phenotype and genotype. A major barrier to coding is a lack of consolidation of such information, scattered in several resources such as Orphanet, OMIM or HPO. We developed a web portal, Linking Open data for RD (LORD), offering an integrated view of 8,336 RDs linked to more than 12,500 signs and 3,000 genes. It allows navigating through the relationships between diseases, signs and genes, and provides Application Programming Interfaces for its integration in information systems (IS). LORD is dedicated to the 131 French RD reference centers and 501 competence centers, for coding RD diagnoses in the health IS.

Mots-clés : Ontologies des maladies rares ; Big Data ; Codage diagnostique ; Web sémantique ; Informatique médicale ; Services web

Keywords: *Rare Diseases Ontologies ; Big data ; Coding support ; Semantic web ; Medical Informatics ; Web services*

1 Introduction et état de l'art

Les systèmes de codage (thésaurus, classifications, terminologies, vocabulaires contrôlés) sont utilisés pour permettre l'entrée de données contrôlées dans les systèmes d'information de santé afin d'en faciliter l'analyse pour l'amélioration de la qualité des soins [1], l'élaboration d'études épidémiologiques [2] ou la comparabilité des protocoles de recherche [3]. Historiquement, la classification internationale des maladies a été utilisée pour mesurer la mortalité dans les populations. Son utilisation a été progressivement adaptée pour mesurer la morbidité puis généralisée aux systèmes d'information hospitaliers non sans quelques difficultés [4]. Les causes en sont nombreuses comme notamment l'utilisation de structures de classification à base d'arbres simples qui ne permettent pas de modéliser les différents usages possibles de ces classifications. L'utilisation d'arbres simples peut en effet nuire à la capacité expressive des éléments d'un domaine et n'est pas adaptée à la formalisation du sens médical [5] de l'information codée.

Près de 80% des maladies rares (MR) ont une origine génétique. Actuellement, 3.000 MR parmi plus de 8.000 ont un gène identifié. Coder un cas de maladie rare est une tâche qui peut être complexe, et qui nécessite un niveau de connaissances médicales spécifiques provenant de diverses sources d'information. Par exemple, la Cystinose est une maladie métabolique accompagnée d'une atteinte cornéenne et rétinienne, mais aussi d'une néphropathie. Si la première information est informative pour un ophtalmologiste, la deuxième l'est pour un néphrologue. La cystinose peut atteindre d'autres organes et présenter des signes cliniques complexes : photophobie, hypothyroïdie, troubles de la croissance, atteinte neurologique,.. Actuellement, une ressource de représentation des maladies rares est désignée pour coder les diagnostics de MR : Orphanet. Elle est actuellement utilisée par 44% des sites nationaux de prise en charge MR (réseau des centres de référence MR labellisés par le ministère de la santé) [6]. Cependant, les diagnostics « non confirmés » représentent 1 cas sur 4 (17% indéterminés, 9% non encore classables). Pour coder les patients atteints de MR dans les systèmes d'information de santé (dossiers médicaux, registres ou cohortes), plusieurs questions ont été examinées par des spécialistes et des cliniciens de l'information en santé : i) le manque d'information sur les maladies rares dans la CIM-10 (moins de 400 maladies rares représentées), ii) le recours souvent nécessaire à plusieurs ressources pour coder un diagnostic MR, comme par exemple de l'information phénotypique et/ou génotypique, iii) la navigation dans 8.336 MR et groupes de MR qui ne peut se faire à travers une liste à plat ou à travers un simple algorithme de recherche, iv) enfin, pour des maladies souvent multi-systémiques, la limite des représentations sous forme d'arbres monoparentaux.

En termes de ressources d'aide au codage pertinentes pour les MR, Diseasecard [7] vise à relier chaque MR aux bases de connaissances connues. Cependant, elle recueille des données à partir de toutes ses sources au sein d'une même interface utilisateur, et ne propose pas de navigation contextuelle dans sa base de connaissances. HeTOP [8] relie de nombreuses bases de connaissances médicales, y compris les principales pour les MR. Cependant, les données d'une seule source peuvent être affichées à la fois et l'application ne propose pas de navigation contextualisée dans l'arborescence des MR. Nous avons donc défini et construit une nouvelle application web pour aider les cliniciens et les spécialistes de l'information de santé à naviguer dans des graphes représentant les diagnostics MR proposés par Orphanet.

2 Matériel et méthodes

2.1 Les sources de données

Au cours de la dernière décennie, de nombreuses ressources terminologiques ou ontologiques (RTO) ont été développées. Peu d'outils sont proposés pour utiliser leur complète représentativité de l'information médicale dans un contexte de codage. Dans le contexte des MR, nous avons identifié quatre ressources d'intérêt avec l'aide d'experts. (i) Orphanet [9] a été notre principale ressource pour naviguer dans les maladies. Cette unité de service INSERM diffuse, via son site OrphaData (www.orphadata.org), 30 classifications MR, ainsi que les synonymes, les signes cliniques, les gènes et les liens à d'autres ressources externes. (ii) OMIM [10] donne accès à une information textuelle importante sur les maladies, les gènes ou les mutations fournis dans des fichiers plats. (iii) HPO [11] propose un grand nombre de phénotypes qui peuvent être utilisés pour le diagnostic clinique des maladies génétiques. Son ontologie est disponible à partir de leur site web dans différents formats. (iv) HRDO [10] enfin, une ontologie dédiée aux MR, propose un méta-modèle pour les données Orphanet. HPO, Orphanet et OMIM mettent à jour leurs données en continu et de manière asynchrone. Leur utilisation comme source primaire pour le codage est un enjeu majeur pour établir des registres de données de patients homogènes et de qualité.

2.2 Architecture

Afin de réunir les informations sur les maladies, les classifications, les gènes et les signes cliniques dans un même espace d'information pour les cliniciens, nous avons tout d'abord effectué une intégration des formats des sources (XML, CSV, OWL) en RDF dans un triplestore Virtuoso. Habituellement, les applications basées sur des données RDF sémantiques sont assez lentes, en raison de la quantité de requêtes SPARQL nécessaires. Donc nous avons choisi de préparer les données à l'affichage en générant chaque maladie et ses données associées en objet JSON que nous stockons dans une collection MongoDB (NoSQL). Enfin, nous avons construit une application web pour aider l'utilisateur à naviguer dans la base de connaissances (figure 1).

2.3 Intégration de données médicales et génétiques dans un entrepôt RDF et définition des objets JSON

Pour construire une base de connaissances cohérente intégrant diverses terminologies et classifications, l'intégration des données dans un espace commun d'information RDF (Virtuoso triplestore) était nécessaire. La source d'information sur laquelle nous avons intégré d'autres données externes (HPO et OMIM) était Orphanet. L'utilisation de SPARQL pour obtenir des triplets de maladies n'a pas été efficace dans un cadre de navigation web. Nous avons donc utilisé un modèle d'application web moderne qui échange des données avec le serveur via des services web, par opposition au rechargement de page classique. Le format de données commun utilisé pour ces échanges était JSON.

Pour la conception de l'interface utilisateur, nous avons d'abord conçu une interface web maquette en collaboration avec des experts MR afin d'identifier les données appropriées à des fins de codage de diagnostic. Les informations et les caractéristiques suivantes ont été retenues : (i) le nom et OrphaCode de la maladie, (ii) les liens externes par source (1-n relations possibles), (iii) les classifications auxquelles la maladie est liée, (iv) un graphique pour naviguer dans la classification actuelle (en raison de la nature multi-parentale du graphique, seuls 3 niveaux de navigation - les parents, la maladie, les enfants - sont conservés) ; (v) les signes ou groupes de signes liés à la

maladie actuelle ; (vi) des informations générales sur la maladie (synonymes, la prévalence, l'héritage, l'âge de début et de décès), (vii) la liste des gènes liés à la maladie à partir de différentes sources le cas échéant (OMIM et Orphanet), (viii) le synopsis clinique ; (ix) les autres informations textuelles connexes lorsqu'elles sont disponibles avec la possibilité de choisir l'affichage de chaque section.

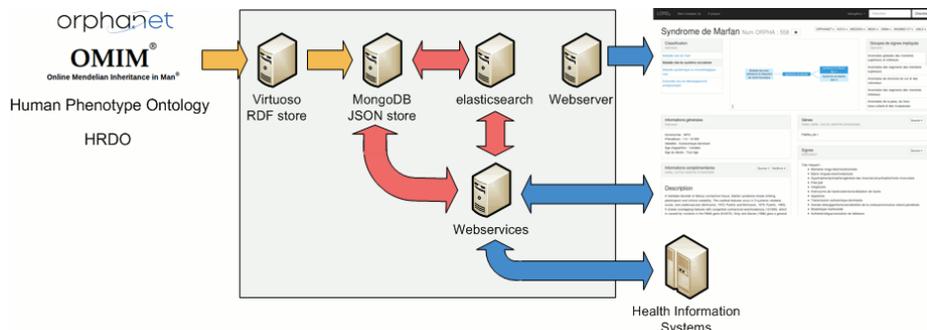


Figure 1 : Architecture de l'application LORD

Cette maquette nous a également permis de définir la structure de l'objet JSON que nous avons utilisé pour transmettre les données des services web. Nous avons créé un script Ruby qui a permis d'intégrer des objets JSON dans une base de données MongoDB.

Catégories	Information	Orphanet	OMIM	HPO
Disorders	PrefLabels	8336	7666	0
	Synonyms	9750	8650	0
	Description	472	2604	0
	Diagnosis	0	412	0
	Clinical features	0	4783	0
Signs	PrefLabels	1360	0	10152
	Synonyms	0	0	6424
	Descriptions	0	0	6844
	Broaders	1310	0	13371
Genes	PrefLabels	3045	14419	0
	Synonyms	7168	21275	0
	Symbols	3045	23429	0
	Methods	0	20639	0

Tableau 1 : Nombre de triplets intégrés dans l'entrepôt RDF Virtuoso ; PrefLabels : nom le plus répandu d'une entité ; Synonyms : noms alternatifs d'une entité ; Description : définition générale ; Diagnosis : description textuelle détaillée du diagnostic ; Clinical features : description textuelle des signes rencontrés ; Broaders : lien vers un élément parent ; Symbols : nom abrégé d'un gène ; Methods : caractérisation de la mutation génétique.

3 Résultats

L'application web est accessible à l'adresse suivante : <http://lord.bndmr.fr>. Nous avons intégré des ensembles de données d'OMIM, HPO, HRDO et d'Orphanet au 1^{er} Janvier 2014. Nous intégrons

des données annuelles dans notre base de connaissances. Le volume et la nature des données intégrées sont décrits dans le tableau suivant.

Un serveur elasticsearch, relié à la collection MongoDB, fournit le moteur de recherche rapide et paramétrable adapté à une recherche efficace. La taille de l'ensemble de données RDF est de 126 Mo, alors que la taille de l'ensemble de données JSON est 63,55 Mo. Nous avons effectué des tests de performance sur le même matériel (machine virtuelle avec 2 cores Opteron 6132 et 8Go de RAM). Obtenir les données nécessaires à l'affichage depuis l'entrepôt RDF prend 2 secondes. Ce délai a été réduit à 1 seconde en utilisant des technologies de mise en cache. Obtenir les mêmes informations depuis MongoDB prend 0,03 seconde. Nous avons également défini un ensemble de services web qui sont utilisés par notre application web. Ces services sont accessibles au public. Ils permettent d'obtenir : la description d'une maladie, la liste des classifications et la recherche de maladies.

4 Conclusion

L'identification des patients atteints de MR est un objectif du plan national maladies rares 2. L'utilisation d'une classification telle qu'Orphanet, qui vise à représenter les maladies rares pour diverses utilisations (codage, diagnostic clinique, registres, cohortes) peut aider à répondre à cet objectif mais elle pose des problèmes de représentativité pour les codeurs. Les cliniciens ont exprimé un besoin spécifique pour avoir à la fois des données phénotypiques et génotypiques, au même niveau, afin d'attribuer le code plus approprié au diagnostic retenu. A titre d'exemple, la *néphronophytose*, une maladie rénale kystique de la médullaire, dispose de 11 formes génétiques distinctes (<http://lord.bndmr.fr/#disorders/655>). La *Cystinose*, une maladie de surcharge lysosomale, peut être considérée comme une maladie rénale rare par le néphrologue ou une MR de l'œil par l'ophtalmologiste (<http://lord.bndmr.fr/#disorders/213>). Il pourrait toutefois être intéressant de pouvoir sélectionner plus d'une spécialité médicale. Les bases de connaissances sur les MR étant en constante évolution, l'arborescence des maladies change chaque jour, ce qui peut compliquer la navigation et les habitudes des médecins. C'est pourquoi une version est créée chaque année, qui peut dès lors être utilisée comme référence. Un processus mensuel d'intégration des nouvelles maladies est en cours d'étude. L'application est étudiée pour permettre l'ajout de nouvelles sources de données au cours de l'année. Permettre la navigation sur des grands graphes sémantiques pose des problèmes de performances sur triplestores RDF classiques. Nous avons utilisé des technologies Big Data pour la persistance des données et permettre une navigation temps réel sur notre ressource intégrant sémantiquement des données génétiques et phénotypiques. Cette application a été requise dans le cadre du 2^e Plan National Maladies Rares afin d'aider les centres de référence et de compétences MR à coder avec un code Orphanet leurs patients avec le même système national de codage dans les différents systèmes d'information de santé. L'application a d'ores et déjà été traduite en anglais pour satisfaire à des opportunités de collaboration européennes sur le sujet.

Remerciements

Nous remercions chaleureusement C. Messiaen, J.-P. Necker et C. Angin pour leur aide. Ce travail a été financé par le ministère de la Santé.

Références

- [1] JM. Januel et al., “ICD-10 adaptation of 15 Agency for Healthcare Research and Quality patient safety indicators.” *Rev Epidemiol Sante Publique*, vol. 59, no. 5, pp. 341–50, 2011.
- [2] E. R. Dubberke, A. M. Butler, H. A. Nyazee, K. A. Reske, D. S. Yokoe, J. Mayer, J. E. Mangino, Y. M. Khan, and V. J. Fraser, “The impact of ICD-9-CM code rank order on the estimated prevalence of *Clostridium difficile* infections.,” *Clin. Infect. Dis.*, vol. 53, no. 1, pp. 20–5, 2011.
- [3] P. Bruland, B. Breil, F. Fritz, and M. Dugas, “Interoperability in clinical research: from metadata registries to semantically annotated CDISC ODM.,” *Stud Health Technol. Inform.*, vol. 180, pp. 564–8, 2012.
- [4] E. J. Picardi and J. B. Peoples, “Mesenteric venous thrombosis: ten year record review and evaluation of difficulties with the ICD coding system.,” *S D J Med*, vol. 44, no. 2, pp. 33–7, 1991.
- [5] A. Sollie et al., “A new coding system for metabolic disorders demonstrates gaps in the international disease classifications ICD-10 and SNOMED-CT, which can be barriers to genotype-phenotype data sharing.,” *Hum. Mutat.*, vol. 34, no. 7, pp. 967–73, 2013.
- [6] P. Landais et al., “CEMARA an information system for rare diseases.,” *Stud Health Technol Inform*, vol. 160, no. Pt 1, pp. 481–5, 2010.
- [7] P. Lopes and J. L. Oliveira, “An innovative portal for rare genetic diseases research: The semantic Diseasecard,” *J Biomed Inform*, vol. 46, no 6, pp.1108-15, 2013.
- [8] J. Grosjean, T. Merabti, N. Griffon, B. Dahamna, and S. J. Darmoni, “Teaching medicine with a terminology/ontology portal.,” *Stud Health Technol Inform*, vol. 180, pp. 949–53, 2012.
- [9] H. Nabarette, D. Oziel, B. Urbero, N. Maxime, and S. Aymé, “Utilisation d’un annuaire des services spécialisés et orientation dans le système de soins: l’exemple d’Orphanet dans les maladies rares,” *Rev Epidemiol Sante Publique*, vol. 54, no. 1, pp. 41–53, 2006.
- [10] A. Hamosh, A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, “Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders.,” *Nucleic Acids Res*, vol. 33, pp. D514–7, 2005.
- [11] S. Köhler et al., “The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data.,” *Nucleic Acids Res.*, vol. 42, no. 1, pp. D966–74, 2014.

Adresse de correspondance

Rémy Choquet, PhD, Hôpital Necker Enfants Malades, Banque Nationale de Données Maladies Rares, Bâtiment Imagine, 149 rue de Sèvres, 75743 Paris. <http://www.bndmr.fr>

Les technologies de l'information et de la communication (TIC) et la promotion de la santé sexuelle auprès des adolescents et jeunes adultes: au-delà de l'efficacité.

Information and communication technologies (ICT) and the promotion of sexual health among adolescents and young adults : beyond the effectiveness

Serge DJOSSA ADOUN ^{1,2}

Marie-Pierre GAGNON ^{1,3}

1- Faculté des sciences infirmières de l'Université Laval

2 - Ph.D. (candidat) santé communautaire

3- Ph.D. ; Professeure agrégée

Résumé

Contexte : La santé sexuelle est un aspect important de la santé générale de l'humain. Malgré plusieurs décennies de promotion de saines pratiques sexuelles, la santé sexuelle des adolescents et jeunes adultes demeure préoccupante. La grande familiarité des jeunes générations avec les technologies de l'information et de la communication (TIC) amène à envisager l'utilisation des TIC pour faire de la promotion de la santé sexuelle comme une avenue prometteuse. Toute fois, pour optimiser les effets positifs des technologies, il est important de savoir si les interventions sont efficaces et si elles seront adoptées au sein des adolescents et jeunes adultes.

Méthode : Une synthèse des connaissances a été réalisée sur l'efficacité des interventions utilisant les TIC visant la promotion de la santé sexuelle auprès des adolescents et jeunes adultes. Au cours de ce processus, les études ayant rapporté les points de vue des bénéficiaires des interventions ont fait l'objet d'une attention particulière afin de ressortir les facteurs qui influencent l'adoption de ce type d'intervention.

Résultats : Les participants aux études inclues dans notre synthèse de connaissances sont à l'aise avec les interventions utilisant les TIC, et ils les trouvent plus intéressantes que les interventions standard. Conclusion : Ceci est à prendre en compte par les différents acteurs qui travaillent à l'amélioration de la santé sexuelle des adolescents et jeunes adultes.

Abstract

Background: Sexual and Reproductive Health (SRH) is an important aspect of human health. Despite several decades of sexual education and promotion of safer sexual practices, adolescents and young adults' SRH remains poor. Information and communication technologies (ICT) are especially popular with young people, thus proposing a promising avenue for SRH promotion. However, before rolling out ICT interventions, we need to know whether ICT interventions are effective for the promotion of sexual & reproductive health in adolescents and young adults.

Methods: A knowledge synthesis was carried out on the effectiveness of ICT intervention for the promotion of sexual health among adolescents and young adults. During this process, the studies that reported the views of the participants in interventions have been gathered to highlight the factors that influence the adoption of such intervention.

Results : Adolescents and young adults were comfortable with ICT interventions and found them useful and interesting.

Conclusion : This should be taken into account by the actors working to improve the sexual health of adolescents and young adults.

Mots Clés : Santé sexuelle ; VIH ; Adolescence ; Promotion de la santé ; Technologies de l'information; meta-analyse, Adoption

Keywords : *Sexual Health ; HIV ; Adolescence; Health Promotion; Information Technologies ; meta-analysis, Adoption*

1. Introduction

La Conférence internationale pour la population et le développement (CIPD), tenue au Caire en 1994, en mettant en évidence la situation particulière des adolescents, a stipulé que ces derniers devraient notamment pouvoir accéder à l'information et à des services qui les aident à comprendre leur sexualité et à se protéger contre les grossesses non désirées, les maladies sexuellement transmissibles et les risques de stérilité en résultant [1]. Toutefois, malgré les multiples initiatives mises en œuvre dans ce domaine, la situation des adolescents et jeunes adultes en matière de santé sexuelle demeure préoccupante globalement. Ceci met en évidence la nécessité de moyens novateurs pour la promotion de pratiques sexuelles plus sécuritaires au sein de cette tranche de la population particulièrement à risque.

2. État de l'Art

2.1 Santé sexuelle des adolescents et jeunes adultes

La santé sexuelle est un aspect important de la santé générale de l'humain. Selon l'Organisation mondiale de la santé (OMS), la santé sexuelle et reproductive (SSR) signifie que les gens sont en mesure d'avoir une vie sexuelle responsable, satisfaisante et sécuritaire, qu'ils ont la capacité de se reproduire et la liberté de décider si, quand et combien de fois le faire [2]. Le cadre conceptuel proposé par l'OMS, soulignant les liens entre la santé sexuelle et reproductive et le VIH, démontre bien que ces deux concepts sont étroitement liés [3].

La situation des adolescents et jeunes adultes en matière de sexualité demeure préoccupante malgré les multiples interventions menées à l'égard de cette tranche de la population [4]. En effet, de plus en plus de personnes ont des relations sexuelles avant le mariage et l'initiation à l'activité sexuelle est de plus en plus précoce [5]. Les adolescents sont définis comme des personnes âgées entre 10 et 19 ans [5, 6], et les jeunes, des personnes de 15 à 24 ans [5]. Aussi, la plupart des statistiques dans le domaine de la santé sexuelle et reproductive des adolescents mentionnent-elles les 15 – 24 ans [7, 8]. Cette tranche de la population est fortement à risque de contracter des infections transmissibles sexuellement et par le sang (ITSS) [9]. Ainsi, 333 millions de nouveaux cas d'ITSS curables surviennent chaque année dans le monde, avec les taux les plus élevés parmi les 20-24 ans, suivi des 15-19 ans [7, 10]. De même, jusqu'à 60% des nouvelles infections ont été enregistrées chez des personnes ayant entre 15 et 24 ans et la moitié des personnes vivant avec le VIH dans le monde se trouve dans ce même groupe d'âge [10].

2.2 Limites des interventions existantes

Différentes stratégies ont alors été utilisées : marketing social, éducation par les pairs, jeu de rôle, interventions basées dans la communauté, etc. Cependant, la plupart de ces interventions visant la promotion de la SSR des adolescents et jeunes adultes ont montré des limites, tant dans la couverture de leur public-cible [11] que dans l'atteinte de leurs objectifs [12]. En conséquence, tant dans les pays en développement le taux de grossesses avant le mariage et celui de certaines ITSS sont en augmentation, tant dans les pays développés le taux des ITSS est en augmentation [13]. La prévention des ITSS chez les adolescents demeure une question cruciale et la réduction de l'incidence du VIH est encore l'un des objectifs principaux de plusieurs ministères de la santé [14]. Les limites des différentes initiatives mises en œuvre dans le domaine de la SSR des adolescents et jeunes adultes montrent l'importance de trouver des modalités innovatrices d'intervention pour la promotion de la SSR au sein de la population des jeunes. Dans ce contexte, les technologies de l'information et de la communication (TIC) offrent une avenue intéressante à explorer.

2.3 Utilisation des TIC pour intervenir auprès des adolescents et jeunes adultes

Les adolescents et jeunes adultes ont davantage accès à l'ordinateur ainsi qu'aux outils technologiques et ce, à l'échelle mondiale [15, 16]. L'Internet, le téléphone cellulaire et les autres technologies sont plus que jamais au cœur du quotidien des jeunes [17]. La vie d'une proportion de plus en plus importante d'adolescents et de jeunes est aujourd'hui fortement tributaire des TIC. Les TIC apparaissent donc comme un moyen incontournable d'information

et de communication auprès de cette tranche de la population. En effet, les jeunes sont des utilisateurs actifs d'Internet [15, 16, 18]. Une étude réalisée auprès d'adolescents américains (âgés entre 13 et 17 ans) recrutés dans le service d'urgences pédiatriques du Rhode Island Hospital indique que la moitié des participants préfère les formats d'interventions basés sur la technologie [19]. Au sujet de l'Internet, O'Neill affirme que, pour les institutions de santé publique, n'y être que peu présentes ou peu s'en préoccuper équivaudrait à se couper d'un lieu central où la vie humaine se situe et se situera de plus en plus [20]. D'autres auteurs argumentent que, parmi les adolescents et jeunes adultes, l'Internet facilite l'accès à la pornographie et à la rencontre de partenaires sexuels, favorisant ainsi la prise de risques accrus en matière de santé sexuelle, et qu'il est essentiel que les éducateurs dans le domaine de la santé adoptent les mêmes moyens utilisés par cette population pour la promotion de la santé sexuelle [21].

Toutefois, pour optimiser les effets positifs des technologies, le développement d'interventions et de stratégies efficaces et durables utilisant les TIC pour la prévention du VIH et autres ITSS chez les adolescents et les jeunes adultes requiert des preuves scientifiques pertinentes. Des auteurs rapportent que les interventions utilisant les TIC ont plusieurs avantages dans le domaine de la santé sexuelle. Ces interventions offrirait, entre autres, la possibilité de délivrer une intervention moins complexe et fondée sur des données probantes en vue de la prévention du VIH [22] ; ainsi qu'une meilleure standardisation, précision, facilité de mise à jour de contenu et une plus grande flexibilité en termes de canal de dissémination et favorisent l'implantation avec fidélité [15, 22-26], en plus de fournir l'anonymat. Selon Bailey (2010), l'utilisateur a le contrôle sur les interventions utilisant les TIC et peut construire sa propre expérience. Ceci démontre l'importance de recueillir l'opinion des bénéficiaires des interventions utilisant les TIC afin de documenter les facteurs reliés à leur adoption au sein des jeunes.

Nous avons réalisé une revue systématique de la littérature suivant un protocole approuvé et publié par la Collaboration Cochrane [27] afin d'évaluer l'efficacité des interventions utilisant les TIC pour la promotion de la santé sexuelle et reproductive incluant la prévention du VIH/sida chez les adolescents et jeunes adultes. L'objectif secondaire de cette revue systématique était d'identifier les facteurs d'adoption qui feront l'objet de cet article.

3. Méthode

Notre revue systématique a été réalisée suivant la méthode de synthèse de connaissances suggérée par la Collaboration Cochrane. Une revue systématique Cochrane est une démarche rigoureuse qui permettra de prendre des décisions éclairées concernant lesdites interventions. En tant que telle, sa réalisation est subséquente à l'évaluation de son protocole afin de s'assurer du respect des critères de scientificité. Pour ce faire, un titre a été enregistré auprès du Groupe Cochrane HIV/AIDS et le protocole de la revue a été évalué par les pairs-réviseurs dudit groupe et publié au début de l'année 2011 [27].

Pour le compte de la revue systématique Cochrane, une stratégie de recherche documentaire a été élaborée par un spécialiste de recherche documentaire. La recherche a été effectuée, entre autres, dans les bases de données suivantes : Cochrane Central Register for Controlled Trials, PubMed, CINAHL, PsychINFO, Computer and Information Systems Abstracts, Communication and Mass Media Complete, ERIC ainsi que certaines des bases de données recommandées pour le domaine de la santé publique et de la promotion de la santé en lien avec les problématiques de la santé sexuelle. Les types d'études recherchées étaient ceux qui ont un des designs suivants : essais cliniques randomisé (randomized controlled trials -RCT- ou quasi-randomized trials - q-RCT – cluster randomized controlled trials – c-RCT) et les essais contrôlés avant-après (controlled before-after - CBA).

Les études incluses devraient porter sur la promotion de la santé sexuelle ou la prévention des ITSS, incluant le VIH/SIDA, auprès des 15-24 ans, et sur des interventions utilisant les TIC. Les TIC incluent l'ordinateur, l'Internet, le téléphone cellulaire, le PDA (Personal Digital Assistant), le CD-Rom, les vidéos interactives, etc. Étaient exclus les articles ne portant pas sur la promotion de la santé sexuelle ou sur la prévention des ITSS/VIH, ceux qui ne rapportaient pas d'interventions utilisant les TIC, ceux qui ne portaient pas sur la tranche d'âge de 15-24 ans, ceux

qui ne présentaient pas le bon design et ceux qui n'étaient pas publiés dans une des langues suivantes: anglais, français ou espagnol. La sélection des études et l'extraction des données ont été faites par deux réviseurs indépendants, un troisième réviseur jouant le rôle d'arbitre pour trancher en cas de divergence. En marge des données statistiques recueillies pour la méta-analyse, les données qualitatives ont permis de réaliser une synthèse narrative concernant, entre autres, les facteurs d'adoption des interventions utilisant les TIC pour la promotion de la santé sexuelle au sein des adolescents et jeunes adultes.

4. Résultats et discussion

La recherche documentaire a généré 15 224 citations, parmi lesquelles 4083 doublons ont été identifiés et rejetés. Un total de 11 127 citations a été examiné. Cinquante-huit (58) références ont été exclues à l'étape d'évaluation des textes intégraux et 14 autres articles répondaient à nos critères d'inclusion.

Nous nous attarderons essentiellement sur des résultats partiels, se rapportant à l'objectif de notre communication. En effet, la revue systématique Cochrane réalisée a été révisée par les pair-réviseurs du Groupe HIV/AIDS de la Collaboration Cochrane et sa publication est imminente. Nous ne saurons donc aborder ici, en détail, les résultats de notre méta-analyse. D'après cette méta-analyse, les interventions utilisant les TIC semblent prometteurs, mais davantage de recherches évaluatives sont requises pour déterminer leur efficacité dans la prévention des infections sexuellement transmissibles, du VIH/SIDA et des grossesses non désirées chez les adolescents et les jeunes adultes.

Les facteurs d'adoption considérés dans notre revue incluaient les considérations économiques reliées à la mise en œuvre des interventions utilisant les TIC ainsi que des aspects affectant l'acceptabilité de ces interventions (barrières et facilitateurs), et leur pertinence du point de vue des bénéficiaires.

Aucune des études incluses n'a réalisé l'évaluation cout-efficacité des interventions utilisant les TIC pour la promotion de la santé sexuelle ou la prévention des ITSS/VIH au sein de la population à l'étude. Quatre études ont toutefois rapporté des données relatives aux facteurs d'adoption.

Parmi les auteurs qui ont étudié ces facteurs, différents aspects ont été abordés :

- Evans et ces collaborateurs (2000) ont utilisé un questionnaire qualitatif pour recueillir les commentaires des participants du groupe d'intervention sur leurs expériences avec le programme d'enseignement assisté par ordinateur (Computer-Assisted Instruction Program). Les participants ont bien apprécié les capsules (vignettes) vidéo incluses dans le programme parce que la vidéo montrait des personnes (réelles) vivant avec le sida, ce qui faisait que le programme était moins monotone [28].
- L'étude de Ito (2008) a évalué l'acceptabilité et la faisabilité d'une intervention interactive menée via un CD-ROM. Il a été demandé aux participants d'indiquer à quel point ils avaient aimé chaque composante de l'intervention sur une échelle de type Likert à 4 points, et dans quelle mesure ils recommanderaient le programme à des amis. Par ailleurs, la faisabilité a été estimée en observant leur degré de difficulté à interagir avec le CD-ROM. Les bénéficiaires ont aimé l'intervention et aucun d'entre eux n'a déclaré ou montré (à l'observation) de la difficulté à l'utilisation du CD-ROM [29].
- L'étude de Marsh (2011) s'est penchée sur l'utilité de l'intervention délivrée via l'ordinateur, selon la perspective des participants, en comparaison à une intervention traditionnelle. Les participants ont perçu l'intervention assistée par ordinateur significativement plus utile comparée à l'intervention standard [22].
- Les participants à l'étude de Roberto (2007) ont trouvé que, dans l'ensemble, le programme (utilisant les TIC) était informatif, clair, utile et intéressant. Ils ont estimé que le contenu ne présentait aucune confusion, mais ils avaient tendance à être neutres sur les aspects « ennuyeux » et « moralisateurs » [30].

En somme, les études rapportant les points de vue des bénéficiaires des interventions utilisant les TIC pour la promotion de la santé sexuelle et reproductive suggèrent que les adolescents et les jeunes adultes sont à l'aise avec lesdites interventions qu'ils trouvent utiles et intéressantes en comparaison à d'autres types d'interventions en santé sexuelle. D'autres auteurs ont également rapporté la préférence des jeunes pour les interventions menées par des voies technologiques comme l'Internet [31] qu'ils perçoivent comme étant un bon moyen pour réaliser l'éducation sexuelle [32]. Enfin, les interventions utilisant les TIC peuvent être propices même dans les pays à ressources limitées vu la grande flexibilité qu'elles offrent en matière de canal de dissémination [16, 18]. Un exemple en a été donné dans l'étude de Tian et collaborateurs (2007) qui rapporte la mise en œuvre d'interventions soutenues par les TIC dans une zone rurale en Chine [33].

5. Conclusion

En marge des données probantes sur l'efficacité des interventions utilisant les TIC pour la promotion de la santé sexuelle incluant la prévention des ITSS/VIH, notre revue systématique a documenté le point des adolescents et jeunes adultes qui semblent pencher plus pour ce type d'interventions. Nos résultats seront utiles aux décideurs et autres utilisateurs de connaissances pour la définition de politiques et stratégies qui contribueront à l'amélioration de la santé sexuelle au sein des adolescents et jeunes adultes.

Références

1. *Report of the International Conference on Population and Development (Cairo, 5-13 September 1994)*. United-Nations, 1994.
<http://www.un.org/popin/icpd/conference/offeng/poa.html>, (consulté le 11 novembre 2010).
2. *Sexual and reproductive Health : overview*. World Health Organization - Regional Office for Africa. <http://www.afro.who.int/en/divisions-a-programmes/drh/sexual-and-reproductive-health/overview.html>. (consulté le 11 novembre 2010).
3. *La santé sexuelle et reproductive et le VIH/SIDA. Un cadre de liens prioritaires*. Organisation mondiale de la santé – Département VIH/SIDA.
http://whqlibdoc.who.int/hq/2005/WHO_HIV_2005.05_fre.pdf. (consulté le 11 novembre 2010).
4. *Rapport sur les infections transmissibles sexuellement au Canada*. Jayaraman, G., Totten, S., Perrin, M., Fang, L., Remes, O. eds. Ottawa. Agence de la santé publique du Canada. 100, promenade Eglantine, Immeuble de Santé Canada n°6.
5. Lou CH, Zhao Q, Gao ES. Effects of high middle school students' learning sex and reproductive health knowledge by Internet. *Fudan University Journal of Medical Sciences*. 2005, 32(3) : 289-94
6. *The second decade: improving adolescent health and development. Programme brochure*. Department of Child and Adolescent Health and Development, Family and Community Health, World Health Organization, eds. Geneva 1998, pp 2.
7. World Health Organization, *Press release*, in *WHO/641995*, WHO: Geneva
8. *Sexually transmitted infections among adolescents: the need for adequate health services*. Dehne L Karl, Riender Gabrielle, eds. Geneva Department of Child and Adolescent Health, World Health Organization, 20 Avenue Appia 1211 Geneva 27, Switzerland. 2005, pp IX.

9. Ito, KE, Kalyanaraman, S, Ford, CA, Brown, JD, & Miller, WC., "Let's Talk about Sex": Pilot Study of an Interactive CD-ROM to Prevent HIV/STIS in Female Adolescents. *Aids Education and Prevention*. 2008 20(1): 78-89.
10. Sexually transmitted infections among adolescents: the need for adequate health services. Dehne L Karl, Riender Gabrielle, eds. Geneva Department of Child and Adolescent Health, World Health Organization, 20 Avenue Appia 1211 Geneva 27, Switzerland. 2005, pp IX.
11. Djossa Adoun Serge, Évaluation des effets du programme adolescents et jeunes de l'Association Béninoise pour la Promotion de la Famille (ABPF) à travers les activités d'un Centre socio-éducatif pour adolescents et jeunes (CSEAJ). Mémoire de maîtrise, Université Laval, Québec. 2009 : pp. 83-4.
12. Magnussen L, Ehiri JE, Ejere HO, Jolly PE. Interventions to prevent HIV/AIDS among adolescents in less developed countries: are they effective? *Int J Adolesc Med Health*. 2004. 16(4): 303-23.
13. Kirby DB, Laris BA, Rolleri LA. Sex and HIV Education Programs: Their Impact on Sexual Behaviors of Young People Throughout the World. *Journal of Adolescent Health*. 2007, 40(3): 206-217.
14. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL. The effects of a computer-based pregnancy, STD, and HIV prevention intervention: a nine-school trial. *Health Commun*. 2007, 21(2): 115-24.
15. Bull S, Pratte K, Whitesell N, Rietmeijer C, McFarlane M. Effects of an Internet-based intervention for HIV prevention: the Youthnet trials. *AIDS Behav*. 2009, 13(3): 474-87.
16. Halpern CT, Mitchell EMH, Farhat T, Bardsley, P. Effectiveness of web-based education on Kenyan and Brazilian adolescents' knowledge about HIV/AIDS, abortion law, and emergency contraception: Findings from TeenWeb. *Social Science & Medicine*. 2008, 67(4): 628-637.
17. *Les jeunes et les technologies de l'information et de la communication*, in *Les jeunes d'aujourd'hui, quelle société pour demain?* Hoibian Sandra, ed. Paris : Centre de recherche pour l'étude et l'observation des conditions de vie. 2012. pp 117-133.
18. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, Cassell J, Nazareth I. Interactive computer-based interventions for sexual health promotion. *Cochrane Database Syst Rev*. 2010(9): CD006483.
19. Ranney ML, Choo EK, Spirito A, Mello MJ. Adolescents' preference for technology-based emergency department behavioral interventions: does it depend on risky behaviors? *Pediatric Emergency Care*. 2013, 29(4): 475-81.
20. O'Neill M. L'internet comme lieu d'intervention en santé publique : pourquoi? Pour qui? Par qui? *Éditorial - Santé publique*. 2009, 21(Hors Série): 2-4
21. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, Cassell J, Nazareth I. Interactive computer-based interventions for sexual health promotion. *Cochrane Database Syst Rev*. 2010(9): CD006483.
22. Marsch LA, Grabinski MJ, Bickel WK, Desrosiers A, Guarino H, Muehlbach B, Solhkhah R, Taufique S, Acosta M. Computer-assisted HIV prevention for youth with substance use disorders. *Subst Use Misuse*. 2011, 46(1): 46-56.
23. Lightfoot M, Comulada WS, and Stover G. Computerized HIV preventive intervention for adolescents: Indications of efficacy. *American Journal of Public Health*. 2007, 97(6): 1027-1030.

24. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: A meta-analysis. *AIDS*. 2009, 23(1): 107-115.
25. Noell J, Ary D, Duncan T. Development and evaluation of a sexual decision-making and social skills program: "the choice is yours--preventing HIV/STDs". *Health Educ Behav*. 1997, 24(1): 87-101.
26. Yom YH, Eun LK. Effects of a CD-ROM educational program on sexual knowledge and attitude. *Comput Inform Nurs*. 2005, 23(4): 214-9.
27. Djossa-Adoun MAS, Gagnon MP, Godin G, Tremblay N, Njoya MM, Ratté S, Gagnon H, Côté J, Miranda J, Ly BA. Information and communication technologies (ICT) for promoting sexual and reproductive health (SRH) and preventing HIV infection in adolescents and young adults (Protocol). *Cochrane Database of Systematic Reviews*. 2011, (2).
28. Evans AE, Edmundson-Drane EW, Harris KK. Computer-assisted instruction: an effective instructional method for HIV prevention education? *J Adolesc Health*. 2000, 26(4): 244-51.
29. Ito, KE, Kalyanaraman, S, Ford, CA, Brown, JD, & Miller, WC., "Let's Talk about Sex": Pilot Study of an Interactive CD-ROM to Prevent HIV/STIS in Female Adolescents. *Aids Education and Prevention*. 2008 20(1): 78-89.
30. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL. The effects of a computer-based pregnancy, STD, and HIV prevention intervention: a nine-school trial. *Health Commun*. 2007, 21(2): 115-24.
31. Miranda, J., *Participants' preferences for HIV prevention interventions, 2009, Thesis*, University of Toronto: Toronto.
32. Lou CH, Zhao Q, Gao ES. Effects of high middle school students' learning sex and reproductive health knowledge by Internet. *Fudan University Journal of Medical Sciences*. 2005, 32(3) : 289-94
33. Tian L, Tang S, Cao W, Zhang K, Li V, Detels R. Evaluation of a web-based intervention for improving HIV/AIDS knowledge in rural Yunnan, China. *AIDS*. 2007, 21 Suppl 8: S137-42.

Adresse de correspondance

Serge DJOSSA ADOUN Ph. D. (candidat) en santé communautaire
 Faculté des sciences infirmières de l'Université Laval
 245, rue de Lanaudière
 Québec, QC, G1L 4H2
 Cel. 1418 805-0629, Courriel : medamas2@hotmail.com

Transition vers la mobilité dans le monde médical : n'oubliez pas l'utilisabilité

*Switching to handheld tools in healthcare:
don't forget usability*

**Frederic Ehrler¹, Evelyne Sarrey², Magali Walesa³, Rolf Wipfli¹, Christian
Lovis^{1,3}**

¹*Service des sciences de l'information médicale, Hôpitaux universitaires de Genève, Genève, Suisse*

²*Direction des soins, Hôpitaux universitaires de Genève, Genève, Suisse*

³*Faculté de Médecine, Université de Genève, Genève, Suisse*

Résumé

A l'heure où un accès instantané et ubiquitaire à l'information devient une demande grandissante de la part des chacun, l'intégration d'outils ultramobiles dans le monde des soins devient un défi prioritaire. Cependant, il est utile de s'interroger sur la meilleure façon de réaliser la transition entre la station fixe, puis mobile et enfin ultra-mobile. Trop souvent, les équipes informatiques, sous la pression du temps et des ressources, se contentent de recycler directement les interfaces existantes sur les nouveaux équipements. Malheureusement, cette transition n'est pas si simple. Le paradigme d'interaction tactile des ces outils nécessite de repenser entièrement la manière de présenter l'information et son interaction afin de garantir un haut niveau d'utilisabilité. Dans cet article, nous cherchons à démontrer cette dépendance à travers les résultats obtenus lors d'une grande étude visant à évaluer les performances utilisateurs lors de l'enregistrement de valeurs simples sur un dispositif tactile à travers différentes interfaces. Trois interfaces ont été présentés aux utilisateurs qui ont du rentrer un certain nombre de signes vitaux. Les résultats montrent clairement que le type de solution choisi influence significativement les performances des utilisateurs ainsi que leur satisfaction. Ceci est un message fort aux décideurs qui devraient clairement prendre conscience de ces problèmes avant de mettre en place des solutions mobiles.

Abstract

There is an increasing pressure from caregivers to get ubiquitous access to clinical information. The use of ultra-mobile devices, such as smartphone, is getting increasingly high. However, there is a need for in-depth thinking regarding the best strategy to realize this transition. Too often, IT teams go for the most straightforward solution and may think that presenting the existing interfaces on the new devices will be an efficient solution. Unfortunately, this transition is not that simple. The tactile interaction paradigm makes the usability strongly dependant of the way

Articles courts des 15^{es} Journées francophones d'informatique médicale, JFIM 2014, pages 211–218

Fès, Maroc, 12–13 juin 2014

information is presented. In this article, we are looking at demonstrating this dependence through the results from a large study testing user performances when recording simple values through different interfaces on mobile devices. Users had to enter some vital signs measure through three different interfaces. The results showed clearly that the chosen solution influences significantly user performances as well as satisfaction. This is a clear message to the stakeholders that should clearly get aware of these issues before implementing mobile solutions.

Mots-clés : *Interface homme machine, Evaluation des systèmes, Utilisabilité,*

Keywords: *Human-computer interaction, Systems evaluation, Usability*

1. Introduction

Toutes les statistiques sont unanimes à reconnaître la forte progression des outils portables et tactiles dans toutes les couches de la population. Alors qu'un nombre croissant de personnes deviennent dépendantes de ces assistants dans plusieurs de leurs activités quotidiennes, il peut être frustrant de devoir abandonner son appareil lorsque l'on franchit la porte de son travail. Cette pression des utilisateurs, pour utiliser des appareils ultra-mobile dans le cadre de leur travail, se fait, sans exception, sentir dans le milieu des soins ou nombre de soignants aimeraient continuer à utiliser ces technologies pour faciliter leurs tâches quotidiennes. Cette pression est ressentie par les organisations qui poussent leurs équipes informatiques à mettre en place, le plus rapidement possible, ces nouveaux supports d'information. Il est nécessaire de bien réfléchir cette transition car elle n'est pas exempte de risques. Un petit examen historique nous rappelle que le monde médical n'est pas exempt d'exemples de transitions technologiques trop rapides qui ont amenés à des échecs [1]. La transition vers les outils mobile n'est donc pas à prendre à la légère. Certaines études montrent déjà que la transition de clavier comme moyen d'interaction vers un paradigme d'interaction tactile peut entraîner une dégradation de la qualité des données enregistrées. Dans cet article, nous présentons une étude menée auprès d'une centaine de soignants afin de tester différentes interfaces de saisie sur une tâche d'enregistrement de signes vitaux. Nous démontrons qu'un choix pertinent de l'interface de saisie est crucial pour assurer un niveau optimal d'utilisabilité, que cela soit à la qualité des données enregistrée ou la satisfaction de l'utilisateur.

2. État de l'art

Dans le monde des soins, la qualité de l'information enregistrée est de première importance. La vie d'un patient peut facilement être mise en danger par un report incorrect d'une prescription médicamenteuse ou par le mauvais report d'une valeur physiologique ou biologique [2]–[4]. Ces problèmes peuvent être atténués par le développement d'appareils médicaux qui sont ergonomiques et qui tiennent en compte des facteurs humains [5]. Concevoir des interfaces fiables et précises est un réel défi. Cela est particulièrement vrai pour les smartphones qui ont des écrans de taille réduite. Des chercheurs ont étudié l'influence de la limitation d'espace sur l'écran de l'utilisateur pour des tâches telles que la navigation, la recherche d'information [6]–[8]. Ils ont montré que la taille du clavier ainsi que d'autres facteurs, tels que la taille des doigts de l'utilisateur, influencent ses performances [9]. D'un autre côté, le paradigme d'interaction tactile ouvre un grand nombre de nouvelles possibilités pour la conception d'interfaces conviviales permettant un travail rapide en minimisant les erreurs.

3. Matériel et méthodes

La transition vers les outils mobiles induit un changement de paradigme d'interaction. Ce nouveau paradigme requiert de concevoir de nouvelles interfaces. Leur conception est cruciale, car elle risque d'influencer significativement l'utilisabilité de l'outil. Afin de tester cette hypothèse, nous présentons une étude comparant les performances et la satisfaction des utilisateurs sur plusieurs interfaces d'interaction tactile. Plus particulièrement, nous proposons trois modèles d'interaction sur lesquels des utilisateurs enregistrent des valeurs de signes vitaux. Pendant le processus, des mesures d'efficacité, d'efficacités et de satisfaction sont enregistrées afin de pouvoir évaluer le niveau d'utilisabilité.

3.1 Les différents modèles d'interaction

L'étude a demandé la programmation de différentes interfaces de saisie tactiles sur un support mobile. La plateforme sélectionnée pour nos tests est le « Samsung Galaxy Note » considérant que le ratio entre la taille disponible sur l'écran et la portabilité était particulièrement intéressant pour une utilisation dans un contexte médical. Ce smartphone possède un écran 5.3 pouces et fonctionne avec le système d'exploitation Android. Concernant les modèles d'interaction implémentés, nous avons identifié 3 modèles qui représentent trois paradigmes d'interaction différents et qui sont susceptibles de modifier l'expérience des utilisateurs lorsque ceux-ci réalisent leurs enregistrements de données.

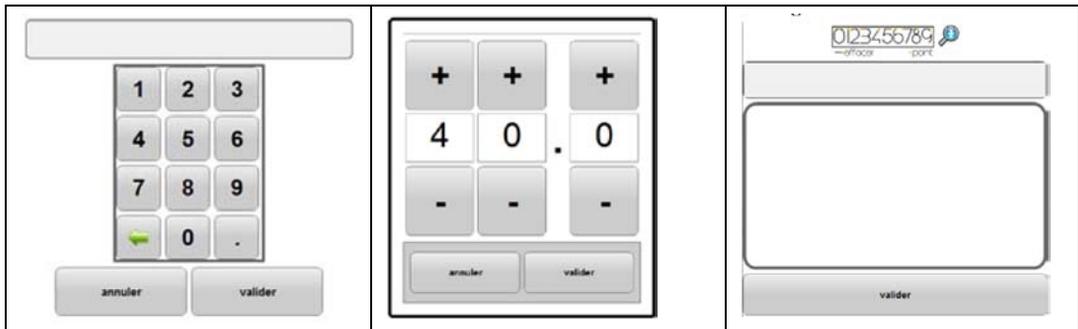


Figure 1 : Les trois modèles d'interaction, le clavier numérique, le stepper et la reconnaissance de caractères

- **Le clavier numérique** : Modèle le plus répandu pour rentrer de l'information numérique, le clavier numérique est utilisé sur de nombreux supports existants comme la calculatrice ou le téléphone.
- **Le stepper** : Ce modèle est aussi relativement commun sur un certain nombre d'outils, par exemple, pour choisir une heure sur un réveil. Le principe de cette interface est que le chiffre sélectionné est augmenté ou diminué d'une unité en appuyant sur le bouton «+» ou «-».
- **La reconnaissance de caractères** : Les interfaces de reconnaissance de caractères sont pratiquement aussi vieilles que les interfaces tactiles. Pour entrer un chiffre avec cet interface, il suffit à l'utilisateur de l'écrire sur la surface mise à disposition.

3.2 Les mesures

L'utilisabilité est définie par la norme ISO 9241-11 comme « le degré selon lequel un produit peut être utilisé, par des utilisateurs identifiés, pour atteindre des buts définis avec efficacité, efficacité et satisfaction, dans un contexte d'utilisation spécifié ». Afin de mesurer ces niveaux d'efficacité, d'efficacité et satisfaction nous avons enregistré un certain nombre d'indicateurs à l'usage de notre applications. Premièrement, l'efficacité est mesurée au moyen du temps pris par l'utilisateur pour réaliser la tâche demandée. Pour cela, nous avons intégré dans l'application un chronomètre mesurant le temps entre le début de la tâche et son accomplissement. Deuxièmement l'efficacité est mesurée en comparant les informations réelles et celles entrées par l'utilisateur. Finalement, la satisfaction est mesurée par un questionnaire de satisfaction demandant explicitement à l'utilisateur s'il a apprécié l'utilisation de l'interface de saisie.

3.3 Design expérimental

Afin de s'assurer de la significativité des résultats, nous avons recruté une centaine de participants. La réglementation en matière de protection des droits des collaborateurs et de la sphère privée a été respectés. La participation était volontaire.

Les utilisateurs ont participé au test suivant une procédure unifiée qui a été définie antérieurement dans un protocole expérimental.

Présentation de l'expérience : Un superviseur explique aux participants ce qu'ils vont devoir faire durant l'expérience. Le superviseur peut intervenir pour aider les participants lors de la présentation des modèles (phase d'apprentissage), mais pas pendant le test en temps que tel.

L'exécution des tâches : Une fois le test démarré, chaque utilisateur se voit proposer les modèles de façon aléatoire. Pour chaque modèle, il lui est demandé de rentrer une série de 3 valeurs successives grâce à l'interface proposée. La procédure est répétée pour trois types de signe vitaux (pouls, température, fréquence respiratoire). Cela signifie qu'au total, chaque utilisateur aura rentré 9 valeurs sur chacun des modèles.

L'évaluation : Après chaque utilisation d'une interface, l'utilisateur doit évaluer son niveau de satisfaction concernant le modèle d'interaction grâce à une échelle de Likert à 6 niveaux.

4. Résultats

Durant les trois mois de la phase expérimentale, au total 93 infirmiers ont participé. Après élimination des tests incomplets et abandons, il restait un total de 87 jeux de données.

4.1 Efficacité

Les résultats obtenus montrent que le temps nécessaire pour entrer des mesures au clavier numérique est significativement moins élevé que celui pris pour entrer des mesures avec le stepper. Le modèle de reconnaissance de caractère à été le moins rapide pour l'enregistrement des données. De façon relative, entrer un nombre avec la reconnaissance de caractères a nécessité trois fois plus de temps qu'avec le clavier numérique et deux fois plus qu'avec le stepper.



Figure 2 : Temps moyen, en millisecondes, nécessaire pour entrer une valeur selon les différents modèles d'interaction

4.2 L'efficacité

L'efficacité est une mesure très importante lorsque l'on travaille avec des données médicales, car on attend de celles-ci qu'elles soient de qualité irréprochable. On voit sur ce graphique que les deux interfaces de saisie que sont le clavier numérique et le stepper ont obtenu une efficacité de 98%. L'efficacité de la reconnaissance de caractère est bien en deçà avec seulement 86%.



Figure 3 : Précision lors de l'enregistrement des données par les utilisateurs

4.3 La satisfaction

La satisfaction des utilisateurs est obtenue par le vote des participants sur une échelle de Likert à 6 niveaux (1 étant la satisfaction maximale et 6 la satisfaction minimale). La figure 4 illustre que plus de 90% des utilisateurs sont extrêmement satisfaits du clavier numérique alors que moins de 50% ont le même niveau de satisfaction avec le stepper. Cependant, la satisfaction du modèle augmente rapidement pour rejoindre celle du clavier numérique. Pour la reconnaissance de caractères, l'évolution est différente, car moins de 50% des utilisateurs ont décrit l'interface de saisie comme satisfaisante (notes 1 à 3).

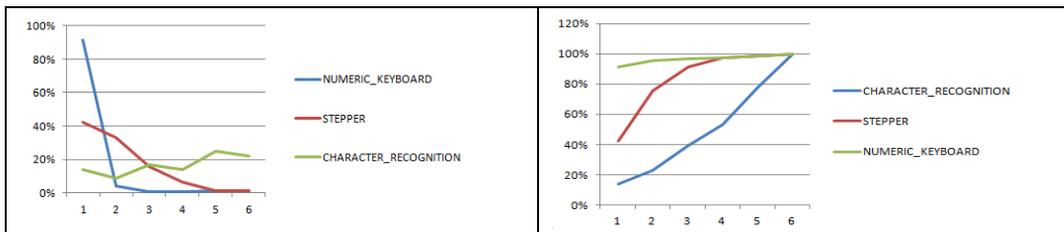


Figure 4 : (A gauche) Satisfaction de l'utilisateur pour les différents modèles. (A droite) satisfaction cumulée

5. Discussion

Les résultats de notre étude montrent que le choix de l'interface influence clairement le niveau d'utilisabilité. Même si les différences de temps pour entrer une valeur (de l'ordre de la seconde) peuvent paraître minimes, lorsque la même opération est répétée une centaine de fois par jour la perte de temps devient clairement quantifiable. Plus important encore, le choix d'une interface inappropriée peut amener une dégradation de la qualité des données enregistrées. De plus, on peut supposer que le nombre d'erreurs s'accroît dans un contexte réel de soins lorsque les soignants subissent une pression constante et ont un risque supérieur de distraction comparé aux conditions de l'expérience où les soignants étaient exclusivement concentrés à saisir les mesures demandées. Finalement, le graphique de la satisfaction montre clairement la préférence des utilisateurs pour le modèle de clavier numérique.

Un point cependant qui reste à éclaircir est l'influence de l'entraînement sur les performances des utilisateurs. On peut en effet supposer, qu'avec le temps, un gain de performance des utilisateurs puisse favoriser un modèle ou un autre. Cela est spécialement possible pour la reconnaissance de caractères qui, d'après nos expériences empiriques, a montré avoir besoin d'un plus grand temps d'adaptation. Ceci fera l'objet d'études ultérieures.

6. Conclusion

La transition vers l'ultra-mobilité est bien en marche. Poussé par la pression toujours plus grande des utilisateurs, les organisations essaient de fournir des solutions le plus rapidement possible, parfois en négligeant d'évaluer l'utilisabilité des outils proposés. Afin de démontrer l'influence de la conception de l'interface sur cette mesure, nous avons mené une étude, effectuée sur près de cent soignants, qui démontre que le niveau d'utilisabilité sur une tâche aussi basique que l'entrée de signe vitaux est fortement influencé par le choix de l'interface. Les choix de conception n'influencent pas seulement la satisfaction utilisateur mais peuvent également avoir un effet délétère sur la qualité des données enregistrées. Ceci est un message fort en faveur d'une réflexion de fond sur les enjeux des interfaces lors d'une transition vers la mobilité, aspects souvent éclipsés par les discussions sur la sécurité, les stratégies de type BYOD ou encore les aspects de coûts ou de maintenance.

Références

- [1] R. Heeks, "Health information systems: failure, success and improvisation.," *Int. J. Med. Inform.*, vol. 75, no. 2, pp. 125–37, Feb. 2006.
- [2] R. Koppel, J. P. Metlay, A. Cohen, B. Abaluck, a R. Localio, S. E. Kimmel, and B. L. Strom, "Role of computerized physician order entry systems in facilitating medication errors.," *JAMA*, vol. 293, no. 10, pp. 1197–203, Mar. 2005.
- [3] K. Henriksen, J. Battles, and E. Marks, "Usability testing and the relation of clinical information systems to patient safety," in *Advances in Patient Safety: From Research to Implementation*, 2005, pp. 365–378.
- [4] R. J. Fairbanks and S. Caplan, "Poor interface design and lack of usability testing facilitate medical error.," *Jt. Comm. J. Qual. Saf.*, vol. 30, no. 10, pp. 579–84, Oct. 2004.
- [5] L. Fennigkoh, "The complexities of the human-medical device interface.," *Biomed. Instrum. Technol.*, vol. 45, no. 1, pp. 39–43, 2011.
- [6] Y.-C. Lu, Y. Xiao, A. Sears, and J. a Jacko, "A review and a framework of handheld computer adoption in healthcare.," *Int. J. Med. Inform.*, vol. 74, no. 5, pp. 409–22, Jun. 2005.
- [7] S. Fischer, T. Stewart, and S. Mehta, "Handheld computing in medicine," *JAMIA*, vol. 10, no. 2, pp. 139–149, 2003.
- [8] G. Haller, D. M. Haller, D. S. Courvoisier, and C. Lovis, "Handheld vs. laptop computers for electronic data collection in clinical research: a crossover randomized trial.," *J. Am. Med. Inform. Assoc.*, vol. 16, no. 5, pp. 651–9, 2009.
- [9] P. Parhi, A. K. Karlson, and B. B. Bederson, "Target size study for one-handed thumb use on small touchscreen devices," *Proc. 8th Conf. Human-computer Interact. with Mob. devices Serv. - MobileHCI '06*, pp. 203–210, 2006.

Adresse de correspondance

Frédéric Ehrler, PhD

Hôpitaux universitaires de Genève

Service des Sciences de l'information Médicale

Rue Gabrielle-Perret-Gentil 4

CH-1211 Genève 14, Suisse

Tél +41 (0) 22 372-8695

Fax +41 (0) 22 372-6255

Mail frederic.ehrler@hcuge.ch

Web <http://www.hug-ge.ch/sciences-de-linformation-medicale>

Importance de la caractérisation des situations de travail dans l'évaluation de l'impact des Systèmes d'Information en Santé

Importance of the work system characterization for impact evaluation of Health Information Technology

**Ludivine Watbled¹, Sandra Guerlinger¹, Marie-Catherine Beuscart¹,
Stéphanie Bernonville¹, Régis Beuscart¹, Philippe Massari², Eric Lepage³,
Sylvia Pelayo¹**

¹ *CIC-IT 807, CHRU de Lille, France*

² *Service d'Informatique Biomédicale, CHU de Rouen & CISMef, TIBS, LITIS EA 4108,
Université de Rouen, France*

³ *LIMICS, INSERM U1142, Université Paris 13, Université Paris Descartes, France*

Résumé

Les Systèmes d'Information (SI) font partie intégrante des environnements de travail en santé. Il devient donc indispensable de pouvoir évaluer leur impact réel sur les pratiques, la performance et la sécurité. Or, de nombreuses études d'impact de l'informatisation ne sont pas satisfaisantes, les résultats étant très hétérogènes. Le problème est qu'elles ne considèrent pas ou non correctement le système socio technique. Peu d'auteurs décrivent de manière détaillée les tâches réalisées, informatisées ou non, le type de SI mis en place dans l'établissement, et l'organisation (la distribution des tâches entre les acteurs, et le SI). Or, au-delà des performances techniques des SI, la caractérisation détaillée des environnements de travail est essentielle pour assurer une interprétation correcte des résultats des études d'impacts, et ainsi identifier toutes les variables de confusion. Cette étude se focalise sur l'impact d'un SI dans le processus de génération – transmission du courrier de sortie. Le présent papier (i) décrit la méthodologie générale adoptée qui sera mise en place pour identifier les variables de confusion liées aux facteurs humains et (ii) présente quelques résultats préliminaires : une cartographie des Systèmes d'Informations comparant les niveaux d'informatisation selon les établissements, des premiers indicateurs mesurant l'impact du SI, et des éléments de l'analyse de l'activité caractérisant le système socio technique. Les premiers résultats encouragent à poursuivre et à développer cette méthodologie pour l'intégrer dans les études de mesures d'impacts.

Abstract

Nowadays, Health Information Technology (IS) are closely integrated in healthcare work environments. It becomes therefore essential to be able to evaluate their real impact on practices, performance and safety. Unfortunately, a number of impact studies in the literature are unusable as their results are heterogeneous. The problem is that the socio-technical system is rarely properly considered. Few authors (if any) provide a detailed description of the underlying tasks (whether or not automated.), of the HIT system in use and of the organization of work

(distribution of tasks between humans and IT, etc.). Indeed beyond the technical performance of SI, the detailed characterization of work systems and work environment is mandatory to ensure a proper interpretation of the results of impact studies and make sure all confusing factors have been identified. This study focuses on the evaluation of the impact of Information Systems on the creation - transmission process of discharge summaries. The present paper (i) describes the methodology adopted to identify confusing variables of Human Factors (HF) nature and (ii) presents preliminary results, e.g. a first mapping of different HIT implementation comparing the level of task automation in different institutions, preliminary indicators measuring the impact of the HIT on the discharge summary process, and elements for the analysis of the corresponding socio-technical system. These early results are promising and encourage further developments of the methodology to be integrated in future impact studies.

Mots-clés : Evaluation ; Systèmes d'Information Hospitaliers ; Courrier de sortie ; Facteurs Humains.

Keywords: Evaluation ; Hospital Information Systems ; Discharge summary ; Human Factors.

1 Introduction

Depuis quelques années, les Systèmes d'Information (SI) font partie intégrante des environnements de travail en santé avec des portails Internet pour les professionnels en santé ou pour les patients, des Systèmes d'Information (SI) destinés à la gestion de l'information médicale (les courriers médicaux, la prescription connectée, le dossier de soins, etc.), etc. Ces SI sont destinés à améliorer la sécurité, l'efficacité des environnements de travail et plus globalement le fonctionnement des organisations de santé. Aujourd'hui, il devient indispensable de pouvoir évaluer l'impact réel de ces SI sur les pratiques des professionnels de santé et sur la qualité des soins délivrés aux patients et sur la performance globale des systèmes de santé eux-mêmes, à travers divers indicateurs. Ainsi, les études évaluant l'impact de l'informatisation de tout ou partie de la génération / transmission du courrier de sortie (« Discharge summary ») s'intéressent le plus souvent à la satisfaction des bénéficiaires [1, 2, 3, 4, 5, 6], au délai de génération / réception du courrier par rapport au jour de sortie du patient [2, 4], à la qualité du contenu du courrier au regard d'un certain nombre de critères [7]. Le problème est que les résultats de ces études sont pour le moins très hétérogènes. Si la plupart des études observent un effet positif de l'informatisation sur les délais de génération / transmission des « discharge summaries », avec des délais parfois raccourcis de façon drastique (de 80 jours à 0), les analyses de qualité des courriers sont plus contradictoires. Par exemple, Maslowe *et al.* (2009) ne montre aucun impact du SI sur la qualité ou satisfaction des usagers [1] alors que O'Leary *et al.* (2009) observent un impact significatif sur la satisfaction quant aux courriers de sortie électroniques fréquemment plus complets. [8]. Ces études emploient le plus souvent une méthodologie de type qualitative à base d'enquêtes (questionnaires et/ou interviews) ou d'audit d'experts, sur la base de grilles pré établies d'évaluations de la qualité de contenu des courriers. Mais très peu d'entre elles décrivent le système socio technique considéré. Par exemple, elles ne mentionnent pas le type de logiciel, ni son mode d'intégration au dossier patient électronique, ni le support technique utilisé (ex. : dictaphone filaire, ou mobile). Elles s'attachent encore moins à décrire les tâches, informatisées ou non, à leur distribution entre les acteurs et le système, et plus généralement l'organisation dans laquelle interagissent les acteurs avec le système. Or, l'introduction des SI dans les systèmes de travail a un impact important sur les activités professionnelles : il peut modifier, parfois de façon radicale, les pratiques et les organisations de travail. Les changements suscités vont porter non seulement sur les individus et leurs tâches mais aussi sur la répartition de ces tâches au sein d'une

équipe, modifiant parfois les rôles et les fonctions des uns et des autres. A un niveau plus large, les organisations sont aussi impactées. Par exemple, l'introduction de la reconnaissance vocale permettant la génération automatique du courrier sur un logiciel de traitement de texte par la dictée du courrier par le médecin, va profondément modifier le travail du médecin : celui-ci doit obligatoirement se positionner face à un écran d'ordinateur pour réaliser son courrier, contrairement aux pratiques habituelles de dictée qui peuvent se faire en tout lieu (couloir, chambre du patient, etc.). Le travail de la secrétaire médicale est également transformé par la réduction de la tâche de retranscription. C'est pourquoi, la description fine du système socio technique est indispensable dans la mesure où les SI ne peuvent être dissociés de leurs utilisateurs, des activités pour lesquelles ils sont conçus et de leurs contextes d'utilisation, au-delà de leurs performances techniques et informatiques [9]. Ainsi il est nécessaire de décrire et d'analyser le système socio technique dont chaque élément peut être une variable explicative des nombreux résultats hétérogènes des études de mesure d'impact de l'informatisation.

Nous proposons ici de définir une démarche méthodologique pour évaluer l'impact de l'informatisation sur tout ou partie du processus de génération / transmission du courrier de sortie. Nous nous focalisons sur les méthodes de caractérisation du système socio technique qui permettront d'identifier les différences existantes selon les organisations, selon les choix techniques, selon les choix d'implémentation, etc. Cette caractérisation du système socio technique permettra (i) d'analyser finement toutes les variables de confusion cachées pouvant expliquer les résultats à première vue hétérogènes reportés dans la littérature et (ii) d'identifier les effets négatifs pouvant minorer les gains observés et pouvant être responsables d'une sous-utilisation ou d'une utilisation détournée des systèmes.

Cette recherche s'inscrit dans le cadre plus général d'un projet PREPS (Programme de recherche sur la Performance du Système des Soins), lancé par la Direction Générale de l'Offre des Soins (DGOS) dont l'objectif est l'évaluation de l'impact des Systèmes d'Information sur la qualité des soins, les conditions de travail et l'efficacité économique et opérationnelle du système de santé, à la fois dans les établissements de santé mais aussi aux interfaces avec le secteur ambulatoire. L'Agence Nationale d'Appui à la Performance des établissements de santé et médico-sociaux (ANAP) assure le pilotage scientifique du projet et assure le lien entre les besoins opérationnels des établissements de santé et cette production scientifique. L'objectif final du projet est l'élaboration d'un cadre commun de référence pour l'évaluation de la création de valeur des projets par les systèmes d'information de production de soins. Le consortium GRESI (Groupe de Recherche dans le domaine de l'Evaluation des Systèmes), en charge de l'exécution du projet UPRES (« Evaluation de la création de valeur par l'usage des systèmes d'information de production de soins »), est coordonné par l'Assistance Publique – Hôpitaux de Paris et comprend les CHU de Rouen, Nice et Lille (des unités de l'Institut national de la santé et de la recherche médicale (Inserm), le CIC-IT / Evalab du CHRU de Lille), le Centre de Recherche pour l'Etude et l'Observation des Conditions de vie (Credoc), le centre de gestion scientifique de l'École des Mines de Paris. Le projet se focalise sur trois domaines d'application pour élaborer et tester les méthodes : (axe 1) le circuit des examens de biologie et d'imagerie, (axe 2) la génération – transmission des courriers de sortie, et (axe 3) la mise en place du Dossier Patient Informatisé (DPI).

La présente étude décrit la démarche méthodologique prenant en compte les situations de travail et plus largement le système socio technique pour étudier l'impact quantitatif et qualitatif d'un SI sur la génération - transmission des courriers de sortie.

2 Méthodologie

2.1 Méthodologie générale du projet

Pour les trois domaines étudiés (circuit de biologie et d'imagerie, la génération – transmission des courriers de sortie et la mise en place du Dossier Patient Informatisé), le projet UPRES se décompose en quatre phases : une phase initiale d'étude des processus et de l'informatisation des circuits étudiés, deux phases de mesures des indicateurs et une phase d'évaluation des impacts de l'informatisation à partir des mesures des indicateurs et de la description des processus des établissements étudiés.

Dans ce premier travail qui a débuté en juillet 2013, pour l'étude des circuits de biologie et d'imagerie menée par le CHU de Rouen et pour l'étude du processus de génération – transmission des courriers de sortie menée par le CHRU de Lille, nous nous sommes focalisés sur la première phase du projet, à savoir l'analyse des circuits des établissements de santé, qui permettra de déterminer l'enchaînement des activités et les flux qui les relient.

2.1.1 *Description du processus*

Pour décrire le processus de génération – transmission des courriers de sortie, plusieurs méthodes issues de la psychologie cognitive ergonomique ont été mises en place :

- Des entretiens semi dirigés auprès des acteurs concernés : médecins, secrétaires médicales ;
- Des observations des tâches impliquées dans la génération et la transmission du courrier pour obtenir les données détaillées des activités des professionnels en interaction avec les systèmes qu'ils utilisent.

Ces méthodes ont été appliquées dans 14 services du CHRU de Lille (neurologie, neurochirurgie, médecine interne, traumatologie, cardiologie, médecine gériatrie, psychiatrie, rééducation, médecine et chirurgie pédiatrique, réanimations, endocrinologie, urologie, oncologie), avec une généralisation auprès des CHU de Rouen, Paris, et Nice.

2.1.2 *Cartographie des SI*

A partir de la description du processus, une cartographie des systèmes d'information dont le formaliste a été créé par l'équipe du CHU de Rouen, est utilisée pour l'appliquer au processus de génération – transmission du courrier de sortie. Cette cartographie permet d'établir les différences de niveaux d'information et ainsi de réaliser des comparaisons inter – services et inter établissements. Les éléments décrits dans la cartographie permettent de spécifier le moyen par lequel sera mesuré l'indicateur retenu.

2.1.3 *Analyse de la littérature*

Afin de déterminer les indicateurs qui seront mesurés, une analyse de la littérature et de la réglementation ont été réalisées. L'analyse bibliographique a porté sur les études s'intéressant à l'impact de l'informatisation sur tout ou une partie du processus de production, communication des courriers d'hospitalisation. A partir d'une grille d'analyse, ont été recherchés : le contexte de l'étude, le système de travail, les composants du workflow automatisés, la ou les méthodes utilisées, les variables dépendantes (indicateurs / données collectées) et les résultats de l'étude.

L'analyse de la réglementation a essentiellement porté sur les éléments que doit contenir tout courrier de sortie. Ces analyses ont permis de dégager quelques indicateurs qui pourront être utilisés dans la mesure de l'impact du SI.

2.2 Méthodologie spécifique mise en place pour caractériser le système socio technique : l'analyse de l'activité

Pour être en mesure de caractériser la situation de travail et plus largement le système socio technique, la méthodologie d'analyse de l'activité du domaine de la psychologie cognitive ergonomique a été utilisée. Ce type d'analyse s'intéresse au professionnel (qui effectue la tâche), à la tâche (ce qui est à faire, un but à atteindre dans des conditions déterminées), à l'activité (ce qui est effectivement réalisé par l'opérateur pour accomplir sa tâche), et au contexte de travail dans lequel le professionnel, son activité et sa tâche vont s'insérer [9, 10]. Analyser une situation de travail c'est aussi décrire les relations entre les professionnels, leurs tâches et les activités mises en œuvre pour les accomplir, le contexte de travail concernant le périmètre, les conditions propres de la situation, et ainsi considérer l'ensemble du contexte organisationnel et socio technique. Les méthodes mises en œuvre pour cette analyse sont l'observation, et l'entretien semi dirigé.

Actuellement, un ensemble de travaux montrent l'ampleur que prend la modélisation des processus métier notamment pour les domaines complexes tels que la santé [11, 12, 13]. Au vu de ce constat, nous avons opté pour un travail autour de la modélisation des processus métier afin de faciliter la compréhension et l'analyse de l'activité de génération - transmission des courriers de sortie. Nous avons choisi l'outil BonitaSoft pour mettre en œuvre notre travail. Cet outil est basé sur la notation BPMN (Business Process Modeling Notation) de plus en plus répandue dans le monde industriel. Il est libre d'accès et simple d'utilisation. La représentation obtenue est relativement facile à comprendre pour les non experts. Par ailleurs, il est possible de procéder à des simulations de flux avec l'outil BonitaSoft, ce qui permet de reproduire et d'analyser le comportement d'un système réel ou cible.

3 Résultats

3.1 Description du processus et cartographie des SI

A partir des descriptions des processus (cf. Figure 1), il a été effectué entre mai et décembre 2013 une cartographie de l'informatisation des circuits des établissements participant au projet (les quatre CHU de Lille, Nice, Paris et Rouen ; voir Figure 1) qui ont précisé pour chaque étape :

- L'informatisation ou la non informatisation ;
- Le mode d'informatisation ;
- Les flux de données inter-étapes ;
- Les données disponibles susceptibles d'être utilisées pour le calcul des indicateurs (dates et heures, connexions, codes des analyses...).

Cette cartographie permet d'identifier certaines phases des processus dont le mode ou le niveau d'informatisation est différent dans les établissements et qui par comparaison d'indicateurs pourront permettre de mesurer l'impact de l'informatisation.

		CHRU Lille : un exemple de circuit avec logiciel de dictée numérique, non intégré au SI	CHU Rouen : un exemple de circuit pour des Comptes-rendus de séjour "courts"
Etape 1	Identification patient	X-Identité scannée depuis une étiquette patient	I-Identité récupérée par le dictateur en sélectionnant le patient depuis le dossier patient
	Commentaires libres		
Etape 2	Recueil d'informations nécessaires à la dictée	I-Depuis le dossier électronique du patient et le dossier papier et la visite	I-Depuis le dossier électronique du patient et le dossier papier et la visite
	Commentaires libres		
Etape 3	Réalisation du courrier	I-Dans le logiciel de dictée numérique distinct du dossier patient	I-Dans le dossier électronique patient
	Commentaires libres		
Etape 4	Transmission de la dictée au secrétariat	I-Transmis après avoir branché le dictaphone à un ordinateur ou socle dédié	A-Autre
	Commentaires libres		Pas de passage par le secrétariat
Etape 5	Lecture / correction par le dictateur	M-Lecture, correction sur impression papier	I-Lecture, correction sur écran
	Commentaires libres		
Etape 6	Validation / signature par le dictateur	M-Signature manuscrite sur exemplaire papier	X-Aucune validation
	Commentaires libres		
Etape 7	Envoi / partage du courrier	M-Envoi par la poste	A-Autre
	Commentaires libres		

Figure 1 : Cartographie (partielle) du circuit de génération et transmission des courriers de sortie aux CHU de Lille et de Rouen.

3.2 Analyse de la littérature

Pour les courriers de sortie, les variables d'évaluation les plus couramment retrouvées dans la littérature [1, 2, 3, 7, 8, 14, 15, 16, 17, 18, 19, 20, 21] appartiennent aux catégories suivantes : (a) les indicateurs temporels (délais d'envoi), (b) les indicateurs de qualité de courrier, (c) la satisfaction des professionnels et du patient. Ces premiers indicateurs ont été retenus pour mesurer l'impact du SI sur la génération et la transmission des courriers de sortie.

D'autres indicateurs viendront probablement enrichir cette liste, mais cela nous permet déjà de regarder en quoi les caractéristiques de la situation de travail ont un impact sur l'informatisation des circuits de génération et de transmission des courriers de sortie.

3.3 Analyse de l'activité

Pour l'indicateur retenu « délai d'envoi », outre le fait de regarder le niveau d'informatisation pouvant expliquer les mesures, en analysant le circuit réalisé entre la génération du courrier et l'envoi (cf. Figure 2), on observe que des composantes de la situation de travail peuvent avoir des impacts sur le délai. En effet, suivant les services, le nombre d'acteurs intervenant dans le processus peut être plus ou moins important. Par exemple, dans un service donné, pour un courrier de sortie, un seul médecin va relire, apporter des corrections et signer le courrier qui pourra alors être envoyé. Pour un autre service, toujours pour un courrier de sortie, plusieurs médecins (deux à quatre) vont relire, apporter des corrections et signer le courrier qui pourra alors être envoyé : dans ce cas de figure, la procédure s'allonge considérablement en temps et en nombre de tâches tant pour le rôle médecin que pour le rôle secrétaire qui devra prendre en compte toutes les corrections des médecins intervenant sur le courrier. Ces deux situations de travail sont des variables explicatives des délais d'envoi recueillis (informatisé partiellement ou totalement). Dans les études d'impact, ces variables ne sont ni prises en compte, ni contrôlées alors qu'elles peuvent expliquer un allongement du délai d'envoi dans les mesures pour la situation informatisée ou pour la situation non informatisée. Cet exemple de l'activité démontre la nécessité de décrire et d'analyser le système socio technique dont chaque élément peut être une variable explicative des mesures d'impacts. Cette analyse du système socio technique permettra d'identifier si l'impact est réellement induit par le SI ou par l'organisation. Ainsi, il s'agira d'instancier un modèle de l'activité. Cette méthode, largement utilisée en psychologie cognitive, permet de caractériser les différentes situations de travail selon quelques déterminants qui peuvent avoir une incidence sur les indicateurs, dont l'informatisation de tout ou une partie des tâches. Ceci permet d'analyser les variations dans les indicateurs non seulement de façon globale (ex. : situation informatisée, situation non informatisée) mais aussi en considérant chaque déterminant pris isolément et leurs combinaisons.

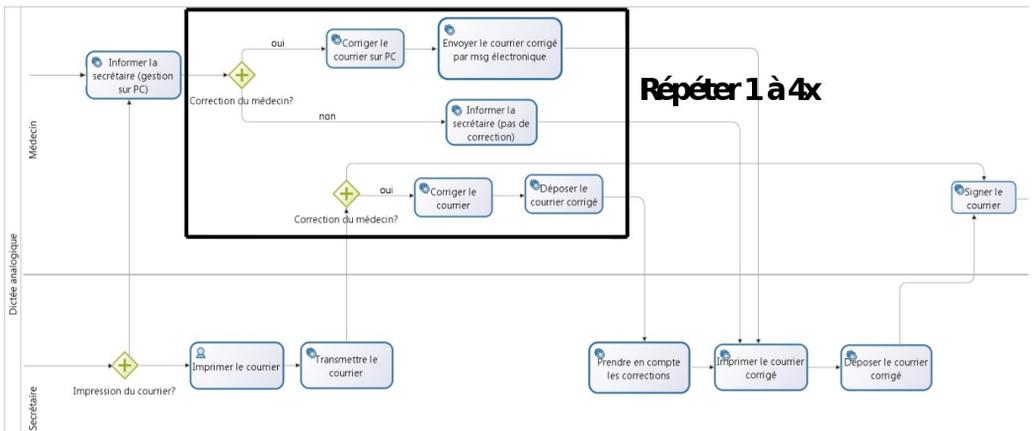


Figure 2 : Description du circuit entre la génération du courrier de sortie et l'envoi.

4 Discussion

Nous avons présenté dans ce papier la première phase du projet UPRES. Les premiers résultats montrent que la méthodologie utilisée permet la description du processus (avec l'exemple du

processus de génération – transmission des courriers de sortie), et la comparaison des niveaux et mode d’informatisation entre les établissements de santé. L’analyse de littérature permet de retenir des premiers indicateurs pour évaluer l’impact de l’informatisation : (a) les indicateurs temporels (b) les indicateurs de qualité, (c) la satisfaction des professionnels et du patient.

Les experts en facteurs humains ont proposé d’intégrer une méthodologie d’analyse de l’activité pour caractériser le système socio technique dont chaque élément qui le compose pourrait être une variable explicative des résultats de mesure d’impact de l’informatisation. Les premiers résultats mettent en évidence que suivant les services de soins et leur organisation, le nombre de tâches par acteur varient fortement, indépendamment de l’informatisation. Il semble alors nécessaire d’inclure cette démarche méthodologique dans l’évaluation de l’impact de l’informatisation. Dans les recherches à venir, il s’agira de prendre en compte le système socio technique en considérant également la politique de l’établissement de santé (ex. : mode de déploiement du SI), et ainsi considérer l’ensemble du contexte organisationnel et socio technique comme défini dans l’approche macro ergonomique [22].

La démarche globale, proposée dans le projet UPRES, devra fournir, aux établissements de santé un cadre commun permettant d’évaluer l’impact du SI. Pour des questions de faisabilité, il s’agira de trouver le moyen de recueillir les données quantitatives et qualitatives par différents types de profils de professionnels (les ingénieurs qualité, les DSI, les professionnels de santé) qui seront amenés à réaliser ce recueil. Ce point clé est un réel verrou méthodologique pour ce qui concerne la méthodologie d’analyse de l’activité qui devra permettre d’identifier les éléments du système socio technique à recueillir dans les établissements de santé. Il s’agira plus particulièrement de vérifier si ce recueil pourra être réalisé par des non experts en Facteurs Humains, en proposant des outils méthodologiques faciles d’utilisation, non chronophages en temps et en ressources humaines et fournissant les éléments nécessaires au recueil des informations pour éviter les biais d’interprétation.

Remerciements

Ce projet UPRES est financé par la Direction Générale de l’Offre des Soins (DGOS) et l’Agence Nationale d’Appui à la Performance des établissements de santé et médico-sociaux (ANAP) dans le cadre du PREPS 2012.

Références

- [1] Maslove DM, Leiter RE, Griesman J, Arnott C, Mourad O, Chow CM, Bell CM. Electronic versus dictated hospital discharge summaries: a randomized controlled trial. *J Gen Intern Med.* 2009 Sep;24(9):995-1001
- [2] Alderton M, Callen J. Are general practitioners satisfied with electronic discharge summaries? *HIM J.* 2007;36(1):7-12.
- [3] Bürkle T, Engel PA. When usage and user satisfaction differ: the case of an electronic discharge summary. *Stud Health Technol Inform.* 2007;129(Pt 2):1093-7.
- [4] Van Walraven C, Laupacis A, Seth R, Wells G. Dictated versus database-generated discharge summaries: a randomized clinical trial. *CMAJ.* 1999 Feb 9;160(3):319-26.

- [5] Archbold RA, Laji K, Suliman A, Ranjadayan K, Hemingway H, Timmis AD. Evaluation of a computer-generated discharge summary for patients with acute coronary syndromes. *Br J Gen Pract.* 1998 Apr;48(429):1163-4.
- [6] Stainkey L, Pain T, McNichol M, Hack J, Roberts L. Matched comparison of GP and consultant rating of electronic discharge summaries. *HIM J.* 2010;39(3):7-15.
- [7] Callen JL, Alderton M, McIntosh J. Evaluation of electronic discharge summaries: a comparison of documentation in electronic and handwritten discharge summaries. *Int J Med Inform.* 2008 Sep;77(9):613-20.
- [8] O'Leary KJ, Liebovitz DM, Feinglass J, Liss DT, Evans DB, Kulkarni N, Landler MP, Baker DW. Creating a better discharge summary: improvement in quality and timeliness using an electronic discharge summary. *J Hosp Med.* 2009 Apr;4(4):219-25.
- [9] Chaudet H, Anceaux F, Beuscart MC., Pelayo S., Pellegrin L. Facteurs humains et ergonomie en informatique médicale. In *Informatique médicale, e-santé, fondements et applications 2013* : 495-520.
- [10] Leplat J, 1993, l'analyse du travail en psychologie ergonomique. Tome 1. Toulouse, Octarès.
- [11] Garcia Rojo M., Rolon E., Calahorra L., Oscar Garcia F., Paloma Sanchez R., Ruiz F., Ballester N., Armenteros M., Rodriguez T., Martin Espartero R. (2008). Implementation of the Business Process Modeling Notation, (BPMN) in the modeling of anatomic pathology processes, *Diagnostic Pathology*, vol. 3(Suppl I).
- [12] Ruiz F., Garcia F., Calahorra L., Llorente C., Goncalves L., Daniel C., Blobel B. (2012). Business Process Modeling in Healthcare, *Studies in Health Technology and Informatics*, vol. 179, p. 75-87.
- [13] Jun G. T., Ward J., Morris Z., Clarkson J. (2009). Healthcare process modeling : Which method when? , *International Journal for Quality in Healthcare*, vol. 21, n°3, p. 214-224.
- [14] Frimpong JA, Jackson BE, Stewart LM, Singh KP, Rivers PA, Bae S. Health information technology capacity at federally qualified health centers: a mechanism for improving quality of care. *BMC Health Serv Res.* 2013 Jan 31; 13:35.
- [15] Liaw ST, Chen HY, Maneze D, Taggart J, Dennis S, Vagholkar S, Bunker J. Health reform: is routinely collected electronic information fit for purpose? *Emerg Med Australas.* 2012 Feb;24(1):57-63.
- [16] Callen J, McIntosh J, Li J. Accuracy of medication documentation in hospital discharge summaries: A retrospective analysis of medication transcription errors in manual and electronic discharge summaries. *Int J Med Inform.* 2010 Jan;79(1):58-64.
- [17] Brankline AL, Coyle CM, Jencks KA, Mullegama A, O'Brien MW. Practical innovations: technology-assisted referrals. *Soc Work Health Care.* 2009;48(8):768-76
- [18] Stetson PD, Morrison FP, Bakken S, Johnson SB; eNote Research Team. Preliminary development of the physician documentation quality instrument. *J Am Med Inform Assoc.* 2008 Jul-Aug;15(4):534-41.
- [19] Navas H, Osornio AL, Baum A, Gomez A, Luna D, de Quiros FG. Creation and evaluation of a terminology server for the interactive coding of discharge summaries. *Stud Health Technol Inform.* 2007;129(Pt 1):650-4.
- [20] Kirby J, Barker B, Fernando DJ, Jose M, Curtis C, Goodchild A, Dickens C, Olla E, Cooke R, Idris I, Thomson GA. A prospective case control study of the benefits of electronic discharge summaries. *J Telemed Telecare.* 2006;12 Suppl 1:20-1.
- [21] Devine EG, Gaehde SA, Curtis AC. Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. *J Am Med Inform Assoc.* 2000 Sep-Oct;7(5):462-8.
- [22] Carayon P, In *Human Factors and Ergonomics in Health Care and Patient Safety 2012*: 3 - 16

Adresse de correspondance

Ludivine Watbled

CIC-IT Biocapteurs et e-santé, Innovations et usages / EVALAB, CHRU Lille

Maison Régionale de la Recherche Clinique

6 rue du Professeur Laguesse

59045 LILLE Cedex

Courriel : ludivine.watbled@chru-lille.fr

Index

- Alami, Hassane, 9
- Ben Loubir, Dalal, 188
- Bennani Othmani, Mohammed, 188
- Bernonville, Stéphanie, 219
- Beuscart, Marie-Catherine, 219
- Beuscart, Régis, 219
- Bittar, André, 152
- Bouaud, Jacques, 21, 71
- Brechat, Bérénice, 34
- Burgun, Anita, 83
- Charlet, Jean, 44
- Chniti, Amina, 44
- Choquet, Rémy, 198
- Cuggia, Marc, 59, 139, 162
- Dahamna, Badisse, 139
- Dalleau, Kevin, 59
- Daniel, Christel, 44
- Darmoni, Stéfan J., 44, 127, 139, 152
- de Carrara, Albane, 198
- Degoulet, Patrice, 115
- Delamarre, Denis, 59
- Dhombres, Ferdinand, 198
- Dini, Luca, 152
- Djossa Adoun, Serge, 204
- Dupuis, Isabelle, 115
- Ehrler, Frédéric, 211
- Ertel-Pau, Valérie, 21
- Falcoff, Hector, 21
- Fonjallaz, Yannick, 198
- Fortin, Jean-Paul, 9, 94
- Gagnon, Marie-Pierre, 9, 94, 204
- Galopin, Alexandre, 71
- Garcelon, Nicolas, 83
- Ghandour, El Kebir, 94
- Gicquel, Quentin, 152
- Grabar, Natalia, 139
- Griffon, Nicolas, 44, 127, 139
- Grosjean, Julien, 139
- Guerlinger, Sandra, 219
- Hadji, Brahim, 115
- Heudes, Didier, 115
- Housbane, Samy, 188
- Hussain, Sajjad, 44
- Jouhet, Vianney, 34
- Joulakian, Mher B., 127, 139
- Landais, Paul, 198
- Ledieu, Thibault, 59
- Lelong, Romain, 139
- Leneveut, Laurence, 115
- Lepage, Éric, 44, 127, 219
- Lovis, Christian, 211
- Maaroufi, Meriem, 198
- Massari, Philippe, 127, 139, 219
- Merabti, Tayeb, 139
- Metzger, Marie-Hélène, 152
- Mougin, Fleur, 34
- Ouagne, David, 44
- Pelayo, Sylvia, 219
- Pereira, Suzanne, 71, 139
- Rabarijaona, Domoïna, 152
- Raimond, Véronique, 21
- Riou, Christine, 162
- Rodrigues, Jean-Marie, 44
- Roger France, Francis, 1
- Sadou, Éric, 44

Salomon, Rémi, 83
Sarrey, Évelyne, 211
Savoy-Collet, Céline, 127
Schuers, Matthieu, 127
Segond, Frédérique, 152
Serhier, Zeineb, 188
Séroussi, Brigitte, 21, 71
Skalli, Soumaya, 127
Soualmia, Lina F., 152

Thiessard, Frantz, 34, 139
Traore, Lamine, 44

Vandenbussche, Pierre-Yves, 198

Wagner, Jean-François, 115
Walesa, Magali, 211
Watbled, Ludivine, 219
Wipfli, Rolf, 211