Une plateforme multimodale d'aide à la décision

Application à la prise en charge du cancer du sein dans le cadre du projet DESIREE

Brigitte SEROUSSI^{1,2}, Jean-Baptiste LAMY¹, Coralie PREBET³, Charlotte NGO⁴, Luis TEIXEIRA⁵, Amaia UGARRIZA⁶, Booma Devi SEKAR⁷, Nekane LARBURU⁸, Naiara MURO^{1,8}, Gilles GUÉZENNEC¹, Jacques BOUAUD^{1,9}

1. *Laboratoire d'informatique médicale et d'ingénierie des connaissances pour la eSanté
Sorbonne Université, Université Paris 13, Sorbonne Paris Cité, Inserm
15 rue de l'école de médecine, F-75006, Paris, France
brigitte.seroussi@aphp.fr*
2. *Département de santé publique, Hôpital Tenon, AP-HP
4 rue de la Chine, F-75020, Paris, France*
3. *Service d'oncologie médicale, Hôpital Tenon, AP-HP
4 rue de la Chine, F-75020, Paris, France*
4. *Service de chirurgie cancérologique, gynécologique et du sein, Hôpital européen Georges Pompidou, AP-HP
20, rue Leblanc, F-75908 Paris cedex 15, France,
Faculté de médecine, Université Paris Descartes, Sorbonne Paris Cité
15, rue de l'école de médecine, F-75006 Paris, France*
5. *Service d'oncologie médicale, Hôpital Saint-Louis, AP-HP
1, avenue Claude Vellefaux, F-75010 Paris, France
Université Paris-Diderot, UFR de médecine, F-75890 Paris cedex 18, France*
6. *Bilbomatica
Santiago de Compostela 12, 4, Bilbao, Espagne*
7. *Computer Science Research Institute, Ulster University
Shore Road, Newtownabbey, Royaume Uni*
8. *eHealth and Biomedical Applications, Vicomtech-IK4
Paseo Mikeletegi 57 Parque Tecnologic, 20009 Donostia-San Sebastian, Espagne*

Biodonostia, Paseo Dr. Beguiristain s/n, 20014 Donostia-San Sebastian, Espagne

9. Délégation à la recherche clinique et à l'innovation, AP-HP
1, avenue Claude Vellefaux, F-75010 Paris, France

RESUME. Dans un objectif d'amélioration de la qualité des soins, nous avons développé une plateforme proposant trois systèmes d'aide à la décision clinique (SADC) basés sur des approches complémentaires, (i) un SADC fondé sur les guides de bonne pratique (GBP) dont l'objectif est de rappeler les recommandations issues de l'état de l'art dans une approche centrée patient, (ii) un SADC fondé sur l'expérience acquise lorsque les utilisateurs cliniciens informés décident de ne pas suivre les GBP et prennent des décisions non conformes pour lesquelles ils doivent fournir une justification, et (iii) un SADC fondé sur la mise en œuvre d'un raisonnement à partir de cas. Les interactions entre les trois SADC sont contrôlées par une ontologie du domaine. Utilisée comme une structure conceptuelle et terminologique, elle propose un modèle de données générique (entité-attribut-valeur) et un modèle de connaissances utilisés pour le raisonnement ontologique (subsumption) et l'aide à la décision déductive (inférence). La plateforme a été développée pour améliorer la prise en charge du cancer du sein dans le cadre du projet européen DESIREE.

ABSTRACT. In order to improve health care quality, we have developed a platform including three clinical decision support systems (CDSSs) based on complementary approaches, (i) a CDSS based on clinical practice guidelines (CPGs) that provides patient-centered state of the art recommendations, (ii) a CDSS based on the experience gained when informed clinician users decide not to follow CPGs and make non-compliant decisions that they are asked to justify, and (iii) a case-based reasoning CDSS. The three CDSSs are smoothly interacting under the control of a domain ontology. Used as a conceptual and terminological structure, the domain ontology proposes a generic data model (entity-attribute-value) and a knowledge model used for ontological reasoning (subsumption) and deductive decision support (inferences). The platform has been developed to improve the management of breast cancer patients within the European Commission funded project DESIREE.

MOTS-CLES : système d'aide à la décision, ontologie, guides de bonne pratique, raisonnement à partir de cas, cancer du sein.

KEYWORDS: clinical decision support system, ontology, clinical practice guidelines, case-based reasoning, breast cancer.

DOI:10.3166/tsi37-0001 © 2018 Lavoisier

1. Introduction

Le cancer du sein est le cancer féminin le plus fréquent en France avec environ 50 000 nouveaux cas par an¹. L'incidence du cancer du sein est actuellement globalement stable après avoir doublé entre 1985 et 2005. Toutefois, si elle diminue

¹ <https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Les-cancers-en-France-Edition-2017>

pour les femmes âgées de 50 à 79 ans qui bénéficient du dépistage organisé au niveau national, elle a augmenté de plus de 60 % chez les femmes âgées de 30 à 39 ans et de 40 à 49 ans. Avec 12 000 décès par an (chiffres consolidés en 2012), la mortalité par cancer du sein est en baisse. Néanmoins, si le cancer du sein est l'un des cancers de meilleur pronostic, avec des taux de survie à cinq et dix ans de 86 %, et 76 %, respectivement, la prise en charge des femmes atteintes d'un cancer du sein reste un challenge thérapeutique. Pour certains sous-types, en particulier les cancers du sein triple négatifs et les cancers du sein HER2+, des marges d'amélioration sont à la fois possibles et nécessaires (Elsamany et Abdullah, 2014).

Les guides de bonne pratique (GBP) ou *guidelines* sont élaborés afin de fournir pour un ensemble de profils patients sélectionnés, les meilleures recommandations de prise en charge. Ces recommandations sont établies à partir des résultats des essais cliniques publiés et représentent l'état de l'art scientifique en accord avec la médecine fondée sur les preuves ou *evidence-based medicine* (Sackett *et al.*, 1996). Des études ont montré que la mise en œuvre des GBP en oncologie avait un impact clinique avec une amélioration de la survie totale et de la survie sans récurrence (Hébert-Croteau *et al.*, 2004). Ceci est particulièrement vrai dans le cadre de la prise en charge du cancer du sein (Schwentner *et al.*, 2012). Néanmoins, malgré la publication de GBP et de la mise à disposition de recommandations de bonne pratique, on observe une variabilité des pratiques dans la prise en charge des patients atteints de cancer, et une conformité aux GBP des traitements reçus encore trop faible (Lebeau *et al.*, 2011 ; Wöckel *et al.*, 2010). Par exemple, Wöckel *et al.* (2010) ont rapporté un taux de conformité des décisions médicales portant sur le plan de soins complet de prise en charge de patients atteints de cancer du sein à 51,9 %.

Depuis quelques années, et depuis le Premier Plan Cancer en France 2003 – 2007², des mesures organisationnelles telles que les réunions de concertation pluridisciplinaire (RCP) ont été introduites afin de promouvoir la décision collective des différents professionnels de santé impliqués dans la prise en charge des patients atteints de cancer (voir Patkar *et al.*, 2011). L'objectif est de rassembler les spécialistes (chirurgiens, oncologues, radiothérapeutes, radiologues, anatomopathologistes, généticiens, etc.) afin qu'ils discutent de chaque patient et construisent ensemble le meilleur plan de soins. Les RCP sont supposées améliorer la mise en œuvre des GBP et le recrutement dans les essais cliniques. Si des études ont montré que les RCP permettaient d'améliorer la conformité des décisions aux GBP (Kesson *et al.*, 2012 ; van Hove *et al.*, 2014 ; Brar *et al.*, 2014), le bon fonctionnement des RCP est entravé par la complexité des cas discutés et la quantité des informations à gérer. En effet, il y a un grand nombre de dossiers à discuter et finalement peu de temps pour discuter de chacun d'eux. De fait, l'impact des RCP sur la qualité des soins a été remis en cause (Keating *et al.*, 2013 ; Johnson *et al.*, 2017), et le choix de l'utilisation de systèmes d'aide à la décision envisagé pour améliorer l'efficacité des RCP.

2 <https://www.e-cancer.fr/Plan-cancer/Les-Plans-cancer-de-2003-a-2013/Le-Plan-cancer-2003-2007>

Les GBP sont généralement élaborés par des sociétés savantes ou des agences nationales de la santé dans un format narratif. La seule diffusion des GBP sous la forme de documents textuels ne permet pas d'améliorer la mise en œuvre des recommandations et n'a pratiquement aucun impact sur les décisions des médecins (Giguère *et al.*, 2012). Au contraire, l'intégration des GBP dans des systèmes d'aide à la décision clinique (SADC) a mis en évidence une amélioration de la conformité des décisions des médecins avec les GBP (Jaspers *et al.*, 2011 ; Roshanov *et al.*, 2011). Différents prototypes de SADC ont été développés pour la prise en charge des patients souffrant d'un cancer du sein. Cependant, à la connaissance des auteurs, seuls quelques-uns ont été utilisés en routine au moment des RCP, et seulement un petit nombre d'entre eux ont démontré qu'ils amélioraient la conformité des décisions des RCP aux GBP. Parmi eux, citons MATE (Patkar *et al.*, 2012) qui utilise le langage PROforma, OncoCure (Eccher *et al.*, 2014) basé sur des protocoles codés en utilisant le formalisme Asbru (Shahar *et al.*, 1998), et OncoDoc (Séroussi *et al.*, 2001 ; Séroussi *et al.*, 2013) qui propose de naviguer à travers une base de connaissances structurée sous la forme d'un arbre de décision pour obtenir les meilleures recommandations centrées-patient. Différents formalismes dédiés ont été proposés pour traduire les recommandations textuelles en recommandations interprétables par un système informatisé (Peleg, 2013). D'autres travaux utilisant des approches issues du web sémantique ont également été proposés, par exemple, Abidi *et al.* (2007) ont proposé d'utiliser des ontologies pour guider un raisonnement basé sur des règles de production pour la prise en charge de patients atteints d'un cancer du sein. Plus récemment, des systèmes comme IBM Watson for Oncology (Lee *et al.*, 2018) ou the Oncology Expert Advisor (Simon *et al.*, 2018) cherchent à construire des outils d'aide à la décision en cancérologie s'appuyant sur des techniques d'intelligence artificielle apprenant sur les données rétrospectives de décisions de RCP.

DESIREE est un projet européen financé dans le cadre du programme H2020³. L'objectif est de développer un écosystème logiciel permettant la gestion collaborative et multidisciplinaire des RCP pour la prise en charge des patients atteints d'un cancer du sein non métastatique (Séroussi *et al.*, 2018). Le système met en œuvre un premier mode d'aide à la décision de prise en charge thérapeutique basé sur la proposition des recommandations issues des GBP dans une approche centrée patient. Néanmoins, les GBP présentent de nombreux défauts (Shiffman *et al.*, 2004) et sont souvent critiqués pour être incomplets. La liste des profils pour lesquels il existe des recommandations n'est pas exhaustive, et lorsqu'un patient n'est pas couvert par les GBP, il y a un silence du système (pas de proposition). Par ailleurs, les GBP sont ambigus, mis à jour à un rythme souvent incompatible avec la rapidité de l'évolution des connaissances, et ne tiennent pas compte des préférences patient. Aussi, certaines décisions prises en RCP peuvent ne pas être conformes aux recommandations et pourtant être légitimes et avoir une valeur clinique qu'il conviendrait de capitaliser comme source de connaissance à côté des GBP. Aussi,

³ Le projet DESIREE a reçu un financement de l'Union européenne dans le cadre de son programme recherche et innovation Horizon 2020 avec la convention de subvention n° 690238.

DESIREE propose un second mode d'aide à la décision basé sur l'expérience acquise à partir des décisions de RCP non conformes aux recommandations. Enfin, un troisième mode d'aide à la décision a été développé sur la base de la mise en œuvre d'un raisonnement à partir de cas. L'objectif est alors de déterminer la décision de prise en charge d'un nouveau patient en réutilisant les décisions prises pour des cas cliniques passés qui lui sont similaires.

Les trois modalités d'aide à la décision ont été mises en œuvre au sein de la plateforme DESIREE sous la forme de trois modules complémentaires et interopérables d'aide à la décision, notés SADC-G pour le module basé sur les GBP, SADC-E pour celui basé sur l'expérience capitalisée, et SADC-RC pour celui utilisant le raisonnement à partir de cas. L'ensemble a été développé en utilisant les outils du web sémantique. En particulier, nous avons construit une ontologie du domaine, le *Breast Cancer Knowledge Model* ou BCKM, utilisée comme une structure conceptuelle et terminologique fédérant les trois modules d'aide à la décision (Bouaud *et al.*, 2018). Nous avons également proposé une visualisation innovante des propositions du SADC-RC sous la forme de boîtes arc en ciel (Lamy *et al.*, 2017).

2. Description des modules d'aide à la décision

2.1. Architecture générale

L'originalité de la plateforme DESIREE est de proposer aux utilisateurs plusieurs modalités complémentaires d'aide à la décision (Séroussi *et al.*, 2018). Le SADC-G est un module s'appuyant sur une formalisation du contenu des GBP sous forme de règles de décision afin de produire des recommandations de prise en charge, essentiellement thérapeutique, conformes à l'état de l'art dans le domaine du cancer du sein et adaptées à la situation clinique du patient. Le second module, SADC-E, le complète en capitalisant sur l'expérience acquise lors des décisions prises par les membres de la RCP lorsque celles-ci n'étaient pas conformes aux recommandations du SADC-G. Le troisième module, le SADC-RC applique un raisonnement à partir de cas sur la base de décisions prises pour les cas les plus proches du cas considéré. Il capitalise également sur la base des cas pris en compte préalablement.

Les trois modules d'aide à la décision sont articulés autour d'un modèle de connaissances commun, le BCKM. Ce modèle de connaissances partagé est représenté sous la forme d'une ontologie qui inclut à la fois le modèle de structuration des données et les aspects termino-ontologiques pour décrire la sémantique des données, et permet de réaliser des raisonnements à différents niveaux d'abstraction.

Par ailleurs, un module externe, le DESIMS (*DESIREE Information Management System*) fait office de dossier patient électronique. Il assure le stockage des données patient et le pilotage des échanges avec les autres modules via son interface utilisateur.

Sur le plan de l'architecture logicielle, chacun des trois modules est mis en œuvre sous la forme d'un service web et fournit une API REST pour le piloter. Ils peuvent ainsi être distribués sur des machines différentes justifiant ainsi l'appellation *cloud-based* de la plateforme dans son intégralité. Chaque module dispose de son propre espace local de stockage utilisant un format interne. Par exemple, le SADC-G stocke les bases de règles correspondant aux différents GBP utilisés par le moteur de recommandations, mais ne mémorise pas de données patient. Le SADC-E maintient une copie des règles utilisées afin de calculer des coefficients tel que le poids de la preuve des règles apprises par expérience ainsi que les cas patients sur lesquels les règles sont apprises. Le SADC-RC stocke l'ensemble des cas dans un format optimisé lui permettant le calcul rapide des distances de similarité entre cas.

Les données patient – ses caractéristiques et celles de son cancer – sont saisies au niveau des formulaires de l'interface utilisateur et stockées dans l'entrepôt de données local du DESIMS. Celui-ci les fournit à la demande aux autres modules d'aide à la décision par le biais d'un serveur tampon s'appuyant sur le format d'échange FHIR (HL7.org, 2017) dédié à la normalisation des échanges d'information dans le domaine de la santé. La figure 1 décrit les interactions principales des trois modules d'aide à la décision, entre eux et entre leurs composants.

Figure 1. Architecture générale et flux d'information entre les différents modules d'aide à la décision et le DESIMS.

Le moteur de recommandation du SADC-G est appelé via l'interface utilisateur. Celui-ci récupère les données du serveur FHIR et applique les règles de décision formalisées qui correspondent aux GBP sélectionnés par les utilisateurs (plusieurs GBP ont été formalisées et sont exploitables par le SADC-G). Le résultat consiste en un ensemble de recommandations qui sont alors affichées dans l'interface utilisateur. Ces recommandations sont adaptées au cas du patient et correspondent à des plans de soins, comme par exemple : tumorectomie et ganglion sentinelle.

Les bases de règles sont gérées par un second composant, le gestionnaire de règles, dont le rôle est de traduire, stocker et sélectionner les règles de décision par GBP.

Le SADC-E récupère les données patient du serveur FHIR pour constituer et mettre à jour une base de données interne concernant les décisions prises en rapport avec les cas patient traités. Lorsqu'une décision non conforme est prise, son générateur de règles élabore une nouvelle règle intégrant des critères supplémentaires et utilise les services du gestionnaire de règles du SADC-G pour la mémoriser. De ce fait, les règles issues de l'expérience sont exécutées par le moteur de recommandations du SADC-G.

Le SADC-RC récupère les données des patients déjà traités pour constituer sa base de cas. Pour un nouveau patient, un algorithme de raisonnement à partir de cas (RAPC) retrouve les cas similaires et, en appliquant son analyseur de similarité, détecte les cas les plus proches et les propose aux participants de la RCP via l'interface utilisateur sous la forme d'une visualisation compacte utilisant la technique des boîtes arc-en-ciel (Lamy *et al.*, 2017).

2.2. Un modèle ontologique pour les connaissances et les données

Nous avons proposé de représenter dans un même cadre ontologique le modèle d'information clinique, ou modèle de données, et le modèle termino-ontologique. Ainsi, le BCKM est une spécification de l'ensemble des entités et concepts qui ont été identifiés comme nécessaires pour la prise en charge des patients atteints d'un cancer du sein. L'objectif du BCKM est d'être non seulement la référence commune unique, mais une ressource exploitable par tous les modules et composants où toutes les structures de données utilisées pour l'application et tous les concepts du domaine applicatif sont définis. Ainsi le BCKM est une ressource centrale statique leur permettant d'être interopérables sur le plan des structures de données et sur le plan sémantique. De plus, sa structuration en une ontologie formelle codée en OWL autorise la combinaison de raisonnements basés sur la relation de subsumption et sur les inférences issues de règles. L'organisation de l'ontologie s'appuie sur les concepts du modèle générique entité-attribut-valeur (EAV) pour la modélisation des données et sur l'identification des concepts liés au domaine, ici le cancer du sein (Bouaud *et al.*, 2018). Ainsi le BCKM n'a pas pour vocation de constituer une

ontologie de référence couvrant tous les aspects du cancer du sein mais plutôt de rassembler l'ensemble de connaissances utiles à la décision de prise en charge thérapeutique du cancer du sein, dans une modélisation explicite. Ceci suppose de permettre (i) la représentation de données patient issues d'un dossier patient électronique orienté « métier » – le métier étant celui des cliniciens – ainsi que (ii) la formalisation et l'opérationnalisation des connaissances contenues dans les GBP pour fournir l'aide à la décision.

2.2.1. Représentation ontologique du modèle de données EAV

Contrairement aux modèles d'information dédiés au domaine biomédical qui offrent des objets ou ressources prédéfinis à la base des systèmes d'information hospitaliers ou des dossiers médicaux électroniques (OpenEHR, FHIR), le modèle EAV est un modèle générique. D'un point de vue logique, les modèles de données, qu'ils soient relationnels ou objet, peuvent être traduits dans le modèle EAV. Il est considéré comme suffisamment flexible pour modéliser des données biomédicales (Nadkarni *et al.*, 1999 ; Löper *et al.*, 2013 ; Khan *et al.*, 2014). Nous avons donc structuré l'ontologie au travers du prisme du modèle EAV et nous avons choisi de représenter explicitement sous forme de classes/concepts les trois éléments du modèle EAV, les entités, les attributs et les valeurs.

- Les sous-classes de *ModelEntity* permettent de définir les composants du modèle de données. Les entités du modèle sont celles qui sont pertinentes pour la prise de décision. Elles correspondent dans notre application aux entités anatomiques dont les descriptions caractérisent la maladie, ici le patient (*PatientEntity*), le côté (*SideEntity*) et la lésion (*LesionEntity*) (figure 2), mais également à d'autres entités associées au patient comme ses parents ou les traitements reçus et examens réalisés.
- Les sous-classes de *ModelAttribute* répertorient les attributs des différentes entités, comme l'âge pour le patient (*Age*), la présence d'adénopathies pour le côté (*ClinicalLymphNodes*) ou le type histologique d'une lésion (*HistologicType*). Chaque classe d'un attribut est déclarée appartenir à une entité au moyen de l'*ObjectProperty isAttributeOf*.
- Les sous-classes de *ModelValue* représentent les différents types de valeur possibles pour les attributs. Des sous-classes correspondent aux types primitifs comme les entiers, les flottants, les booléens, les dates, les chaînes de caractères. Une sous-classe distincte, *HierarchicalValue*, possède des sous-classes correspondant à des hiérarchies de valeurs possibles organisées par la relation de subsumption. Ainsi, les types de cancers du sein sont décrits hiérarchiquement incluant le cancer invasif, le cancer in situ, la maladie de Paget, etc. L'intérêt de ces ensembles de valeurs structurés par subsumption est d'autoriser le recueil d'information à différents niveaux d'abstraction et le raisonnement sur ces différents niveaux.

Dans l'ontologie, la spécification du type d'un attribut se fait au moyen de l'*ObjectProperty* unique, 'hasRange', entre la classe de l'attribut et la classe du type de valeurs. Par exemple, l'attribut *Age* d'un patient est lié au type

IntegerValue, l'attribut ClinicalLymphNodes d'un côté est lié à BooleanValue et l'attribut HistologicType d'une lésion à pour type BreastCarcinoma.

Les relations entre entités sont représentées dans l'ontologie par des sous-ObjectProperties de l'ObjectProperty isRelatedTo. Par exemple, une entité LesionEntity est reliée à l'entité SideEntity par l'ObjectProperty hasSide. Ce niveau de hiérarchie permet de distinguer les relations propres au modèle de données et d'en obtenir la liste en une seule requête.

Figure 2. Extrait du diagramme de classes représentant les trois principales entités cliniques utilisées pour décrire un cas de cancer du sein.

2.2.2. Intégration des connaissances du domaine à l'ontologie

Les éléments présentés précédemment fournissent le cadre pour la représentation d'un modèle EAV et forment le noyau du BCKM. Les connaissances, c'est-à-dire, l'ensemble des concepts liés à l'application à développer et au domaine à proprement parler, ici le domaine de la prise en charge du cancer du sein, sont représentées par des parties d'ontologies venant se greffer de manière modulaire à l'ontologie noyau. Parmi les concepts modélisés, on identifie ceux correspondant aux entités du modèle de données, comme le patient, les côtés, les lésions, mais aussi les procédures, les recommandations. On identifie également les différents attributs caractérisant chaque entité, comme l'âge, le type histologique, la taille de la tumeur, etc. Enfin, les ensembles de valeurs hiérarchiques propres au domaine d'application doivent être intégrés à l'ontologie. Ces ensembles caractérisent par exemple des ensembles non structurés de valeurs simples, comme le sexe avec les valeurs masculin et féminin, ou encore le grade avec les valeurs grade faible, grade intermédiaire et haut grade. Ils permettent de caractériser également des ensembles structurés par subsumption, plus riches, comme les différents types histologiques de cancer du sein mentionnés précédemment, la classification de gènes, les différentes procédures thérapeutiques. Pour construire ces ensembles de valeurs, et lorsque cela était possible, nous avons extrait des sous parties d'ontologies existantes, en particulier de l'ontologie sur le cancer du thésaurus du National Cancer Institute

(NCI thesaurus)⁴. Nous avons utilisé l'outil Flat OWL Editor (Desprès et Guézennec, 2017) pour extraire des hiérarchies de concepts du NCI thesaurus (NCIt) puis pour les injecter dans le noyau du BCKM. La figure 3 schématise l'organisation de l'ontologie résultante en distinguant l'ontologie noyau, dédiée au modèle de données, et la sous-ontologie des concepts du domaine stricto sensu.

Figure 3. Schéma de la structuration du haut de l'ontologie BCKM.

Dans l'état actuel, le modèle construit pour DESIREE s'appuie sur 22 entités et 394 attributs. Les attributs se répartissent en fonction de leur type de la manière suivante : 49 % de booléens, 9 % d'entiers, 4 % de flottants, 5 % de chaînes de caractère, 4 % de dates et 33 % font référence à des valeurs hiérarchiques. L'ontologie résultante, le BCKM, contient 1445 classes, 2305 axiomes, 25 ObjectProperties, 15 DataProperties. Au total, 658 classes du BCKM sont issues du NCI thesaurus.

La figure 4 montre un extrait du BCKM dans l'éditeur Protégé. On y voit que le concept HistologicType est un attribut (isAttributeOf) de l'entité SideEntity et a pour type de valeur (hasRange) BreastCarcinoma. On voit également que BreastCarcinoma est sous-classe de HierarchicalValue et possède une arborescence de sous-classes.

⁴ <https://ncit.nci.nih.gov/ncitbrowser/>

2.3. Le SADC-G

2.3.1. Guides de bonne pratique et structuration des bases de règles

La première étape a consisté à sélectionner les GBP dédiés à la prise en charge du cancer du sein afin de couvrir l'état de l'art international et de tenir compte des pratiques des partenaires cliniques du projet. Aussi, nous avons sélectionné les recommandations NCCN (*Version 3.2017 – November 10, 2017*) publiées par le réseau des 27 centres de cancérologie aux Etats-Unis (NCCN.org, 2017), les recommandations publiées par l'Assistance Publique – Hôpitaux de Paris en Mars 2016 (AP-HP, 2016) et celles publiées par Onkologikoa, centre de prise en charge du cancer, situé à San Sebastian (Espagne), et partenaire du projet DESIREE. Les travaux de modélisation des GBP se sont déroulés en plusieurs étapes à partir de la version textuelle des GBP (cf. Figure 5). Les travaux portant sur les règles espagnoles ont été conduits par les partenaires espagnols (NL, NM) et ne seront pas décrits ici.

Figure 4. Extrait de l'ontologie du BCKM dans Protégé.

C - Indications de reprise chirurgicale

1 • Berges

- Cancer in situ, berges de marges $\leq 2\text{mm}$
 - Cancer infiltrant : berges atteintes au niveau de la surface encre
- Les indications de reprise chirurgicale doivent toutes tre discutes en RCP.

Figure 5. Extrait des GBP textuels de l'AP-HP (Mars 2016).

Pour construire les bases de connaissances, la premire tape a consist construire une version semi-formelle du texte des GBP sous la forme de rgles SI-ALORS avec le lien vers la page du texte o la proposition de prise en charge tait dcrite et le niveau de preuve de la proposition (proposition imprative, recommande, possible, contre-indique, interdite), par exemple sur la base du texte de la figure 5, « SI Cancer du sein in situ ET Tumorectomie ralise ET il existe une lsion avec marges $\leq 2\text{mm}$ ALORS Tumorectomie de reprise recommande (p13) OU Mastectomie possible (p13). Cette tape a t ralise par un oncologue mdical spcialis dans la prise en charge du cancer du sein (CP). Les rgles ont t valides par les cliniciens du projet (CN, LT). La deuxime tape a repos sur la compltion et la normalisation des notions utilises dans les rgles semi-formelles en conformit avec les concepts dfinis dans l'ontologie et la rcriture des rgles dans un pseudo langage logique restant toutefois comprhensible par les cliniciens, ce qui donne sur la rgle prcdente : IF InSituBreastCarcinoma AND PriorBreastConservationSurgery AND InSituCarcinomaMarginWidth ≤ 2 THEN BreastReExcisionForPositiveMargins using 'SHOULD' OR Mastectomy using 'MAY'. Cette tape a t ralise par un informaticien mdical ayant une bonne connaissance de la prise en charge du cancer du sein (BS).

2.3.2. Reprsentation des donnes et langage(s) de reprsentation des rgles

L'tape suivante a port sur le codage des rgles dans un langage driv de NRL (*Natural Rule Language*) (Nentwich et James, 2010), originellement un langage destin au contrle de validit de modles de donnes (Bosca *et al.*, 2014) inspir d'OCL (Object Constraint Language). NRL permet ainsi l'criture d'expressions logiques en conformit avec les composants du modle EAV et indpendamment de la spcification d'un moteur d'infrences. La figure 6 donne en exemple la traduction en NRL de la rgle prcdemment utilise.

```

Context: PatientEntity
-- =====
Action Rule : "SI Cancer du sein in situ ET Tumorectomie réalisée ET il existe une lésion
avec marges ≤ 2mm ALORS Tumorectomie de reprise recommandée (p13) ou
Mastectomie possible (p13)"

"theSide" is a SideEntity,
"thePatient" represents theSide.hasPatient,

IF
    thePatient.BUScenario = ScenarioD
AND   theSide.BreastHistologicType is a kind of InSituBreastCarcinoma
AND   theSide.PriorBreastConservationSurgery = true
AND   theSide.InSituCarcinomaMarginWidth ≤ 2

THEN
[buildcontainer ] RecoEntity from
    [build] OrderEntity with BreastReExcisionForPositiveMargins to theSide
    with '' using 'SHOULD',
[buildcontainer ] RecoEntity from
    [build] OrderEntity with Mastectomy to theSide with '' using 'MAY';

```

Figure 6. Extrait d'une règle de décision codée en NRL respectant le modèle du BCKM.

Le BCKM est le modèle de référence à partir duquel un modèle XSD est automatiquement dérivé qui sert de référence pour le parseur NRL qui valide les règles de décision. Les règles NRL sont ensuite transformées en une représentation intermédiaire en XML, indépendante du langage d'expression des règles et indépendante du langage d'exécution des règles, analogue aux *Knowledge artifacts* proposés par l'organisation HL7 (HL7.org, 2015). Cette représentation XML est utilisée pour les échanges avec les autres modules, notamment le SADC-E.

Enfin, la dernière étape a consisté à traduire automatiquement cette représentation en une représentation interne des règles en notation N3 (Berners-Lee *et al.*, 2008), langage du web sémantique, destinée à alimenter le moteur d'inférences.

2.3.3 Architecture pour l'exécution et la production des recommandations

Le moteur de recommandations est le composant central du SADC-G (Séroussi *et al.*, 2017). Afin de traiter à la fois le raisonnement à partir de règles et le raisonnement ontologique (mise en œuvre de la subsumption), nous avons utilisé le moteur d'inférence sémantique EULER/EYE (Verborgh et De Roo, 2015). Euler utilise N3, il est donc nativement interopérable avec l'ontologie du domaine. Les deux outils appartiennent à la même famille d'outils du web sémantique capables de manipuler des graphes de triplets. Le moteur prend en entrée les règles N3, un

2.3.4. Production des recommandations

The diagram illustrates the DESIMS architecture, which is a decision support system for patient care. It is organized into several interconnected components:

- Entrepôt de données patient (DESIMS)**: The central data warehouse, represented by a cylinder, which serves as the primary data source for the system.
- Serveur HAPI FHIR**: A server component that interacts with the data warehouse and the Patient Data Bundle.
- Patient Data Bundle**: A component that receives data from the HAPI FHIR server and feeds into the FHIR Parser.
- Moteur de recommandations**: The core recommendation engine, enclosed in a dashed box. It contains:
 - FHIR Parser**: Processes the Patient Data Bundle into N3 facts.
 - Fact Builder**: Takes input from the FHIR Parser and outputs N3 facts.
 - Rule Extractor**: Extracts rules from the Rules.db and outputs N3 rules.
 - N3 facts** and **N3 rules**: Intermediate data products that feed into the Inference Engine.
 - Inference Engine**: The central processing unit that generates N3 new facts based on the input facts and rules.
 - N3 model**: A component that receives input from the Inference Engine and feeds into the Recommendation Builder.
 - Recommendation Builder**: Generates recommendations based on the N3 model and outputs them to the Interface graphique.
- Interface graphique**: The user interface component that displays recommendations to the user.
- Rules.db N3 + XML (SQLite)**: A database component that stores rules and provides input to the Rule Extractor.
- Traducteur de règles**: A component that translates rules from the Rules.db into a format usable by the Rule Extractor.
- CPG 1, CPG 2, CPG 3**: Clinical Practice Guidelines represented as stacked boxes, which feed into the Traducteur de règles.
- Knowledge Model OWL**: A component represented by a blue molecular-like structure that feeds into the Traducteur de règles.

The flow of data is as follows: The Entrepôt de données patient (DESIMS) feeds into the Serveur HAPI FHIR and the Interface graphique. The Serveur HAPI FHIR feeds into the Patient Data Bundle, which then feeds into the FHIR Parser. The FHIR Parser feeds into the Fact Builder, which outputs N3 facts. The Rule Extractor feeds into N3 rules. Both N3 facts and N3 rules feed into the Inference Engine. The Inference Engine outputs N3 new facts, which feed into the N3 model. The N3 model feeds into the Recommendation Builder, which outputs Recommendations to the Interface graphique. The Rules.db N3 + XML (SQLite) feeds into the Rule Extractor. The Traducteur de règles feeds into the Rule Extractor and receives input from CPG 1, CPG 2, CPG 3, and the Knowledge Model OWL.

Figure 7. Schéma général et visualisation des flux de données au niveau du moteur de recommandations du SADC-G.

2.4. Le SADC-E

Le SADC-E vise à capitaliser l'expérience clinique contenue dans les décisions de RCP non conformes aux propositions du SADC-G. Ces décisions non conformes doivent être justifiées par la donnée des critères patient spécifiques qui ont conduit la RCP à ne pas suivre les recommandations des GBP. Ces critères caractérisent des exceptions aux recommandations, ils peuvent correspondre à des préférences patient (p. ex. refus de mastectomie, de chimiothérapie), à des cas particuliers (p. ex. antécédents de radiothérapie, membre supérieur amputé), à des pratiques spécifiques locales (p. ex., différence de seuil utilisé pour classer, non disponibilité d'un médicament recommandé).

2.4.1. Apprentissage de nouvelles règles

Les décisions non conformes sont représentées par un événement décisionnel noté D et sont à l'origine de la création d'une nouvelle règle de décision R apprise par expérience. Pour un patient P , on note $D = \langle \{P_i\}, \{R_j\}, FD, \{C_k\}, E, O(t) \rangle$ où $\{P_i\}$ est l'ensemble des données du patient P , $\{R_j\}$ est l'ensemble des règles de décision déclenchées par le SADC-G pour proposer les recommandations adaptées au patient P , FD est la décision (non conforme) prise par la RCP pour le patient P , $\{C_k\}$ est l'ensemble des critères intervenant dans la justification de la non conformité, E est le traitement effectivement reçu par le patient P (la plupart du temps, $FD = E$).

La nouvelle règle R apprise par expérience a une partie SI constituée des données du patient P enrichies des critères de justification de la non-conformité $\{C_k\}$, et une partie ALORS qui correspond à la décision finale FD . On a alors : $R = SI \{P_i\} \cup \{C_k\} \text{ ALORS } FD$ avec $O(t)$ où $O(t)$ représente le poids de la preuve de cette nouvelle règle, calculé par la fréquence de l'impact positif de son application (Muro *et al.*, 2017). C'est un poids qui varie en fonction du temps : il est dynamiquement actualisé en fonction de l'application de cette règle et de l'évaluation clinique de son utilisation pour décider. Des questionnaires permettant le recueil de l'évaluation de la qualité de vie, de la survie, etc. par les patients ayant reçu le traitement préconisé par cette règle (*patient-reported outcomes*) sont utilisés pour le calcul (cf. ICHOM⁵).

2.4.2. Interactions du SADC-E et exécution des règles

Le SADC-E reçoit du DESIMS les décisions prises par la RCP et les recommandations issues du SADC-G. La production des nouvelles règles dépend du site clinique où sont prises les décisions afin de fonder la pertinence de ces règles sur les pratiques de la RCP du site. Cela dépend également des GBP utilisés que les

⁵ <https://www.ichom.org>

règles construites par expérience vont compléter. Par ailleurs, les règles apprises par expérience sont générées conformément aux concepts du BCKM et sont ainsi en parfaite cohérence formelle avec les règles issues de la modélisation des GBP qu'elles viennent spécifier.

Pour stocker les règles apprises, le SADC-E utilise les services du gestionnaire de règles développé pour le SADC-G. Elles y sont indexées différemment des règles issues des GBP. L'exécution des règles issues de l'expérience se fait au moyen du moteur de production des recommandations du SADC-G. Dans ce cas, le SADC-G est appelé depuis le DESIMS, pour un cas patient donné, en précisant en arguments les GBP à utiliser et le site dont on souhaite les propositions issues de l'expérience. Ces dernières viennent alors compléter au niveau de l'interface les recommandations issues des GBP.

2.5. Le SADC-RC

Le raisonnement à partir de cas est un paradigme de l'intelligence artificielle issu des travaux de Minsky (1975), et de Schank et Abelson (1977). Ce type de raisonnement consiste à réutiliser les solutions de problèmes passés, déjà résolus, dans le but de résoudre de nouveaux problèmes. La réutilisation va de la simple ré-application de la solution précédente à la révision d'une solution précédente afin de l'adapter aux spécificités du nouveau problème. Un système d'aide à la décision basé sur le raisonnement à partir de cas dispose d'une base de cas, chaque cas étant décrit par un ensemble de critères et une solution. Le cycle du raisonnement à partir de cas est constitué de quatre processus consécutifs : la recherche de cas similaires, la réutilisation des solutions passées, la révision des solutions passées et l'apprentissage de nouvelles solutions. Dans le cadre de DESIREE, les cas sont des cas cliniques décrits sous la forme d'un vecteur de variables représentées par un couple attributs-valeurs et de la solution du cas qui représente la décision de prise en charge thérapeutique prise par la RCP (Sekar *et al.*, 2018). Les couples attributs-valeurs sont pris dans le modèle de données du patient tel que décrit par le BCKM. On utilise uniquement la première étape du cycle, c'est-à-dire, la recherche des cas similaires, le choix de la décision pour le nouveau cas clinique étant laissé à l'utilisateur sur la base d'une visualisation de la décision prise pour les cas les plus similaires.

2.5.1. Recherche de cas similaires

La recherche des cas similaires repose sur l'utilisation d'une distance de similarité notée *Sim*. Soit $X = \{x_j\}$ le cas clinique à résoudre et $Y_i = \{\{y_{ij}\}, d_i\}$ les cas de la base de cas, où les d_i représentent les décisions prises pour les cas Y_i . La distance de similarité est définie comme une distance Euclidienne qui considère chaque attribut comme une dimension pondérée par le poids w associé à chaque attribut :

$$Sim(X, Y) = \sqrt{\sum_{j=1}^{j=n} w_j \times dist(x_j, y_j)^2}$$

Au niveau de la base de cas initiale, seuls 51 attributs avaient été utilisés pour caractériser les cas cliniques (22 booléens, 16 numériques et 13 catégoriels). Les 51 attributs ont été sélectionnés et pondérés (calcul des w_j) en utilisant les algorithmes *BestFirst* et *Ranker* de Weka (Witten *et al.*, 1999) conduisant à ne retenir que 12 attributs comme « pertinents » pour le calcul de la similarité. Ces résultats ont ensuite été soumis aux avis de trois experts dans la prise en charge du cancer du sein ce qui a conduit à augmenter la valeur de trois poids et à rajouter un attribut qui n'avait pas été initialement sélectionné (HistologicType).

La fonction *dist* calcule la distance entre deux valeurs d'un même attribut. Trois fonctions *dist* ont été implémentées, selon le type de l'attribut. Pour les booléens, il s'agit simplement de comparer les deux valeurs. Pour les nombres, il s'agit d'une distance Euclidienne classique. Pour les valeurs catégorielles, nous avons utilisé la distance sémantique cosinus qui s'appuie sur la position hiérarchique des attributs dans le BCKM (Sekar *et al.*, 2018), et calcule la distance comme un degré de proximité taxonomique dans l'ontologie.

$$dist_{booléen}(x_1, x_2) = \begin{cases} 1 & \text{si } x_1 \neq x_2 \\ 0 & \text{si } x_1 = x_2 \end{cases}$$

$$dist_{numérique}(x_1, x_2) = \frac{|x_1 - x_2|}{\max(x) - \min(x)}$$

$$dist_{sémantique}(x_1, x_2) = 1 - \frac{|super(x_1) \cap super(x_2)|}{\sqrt{|super(x_1)|} \times \sqrt{|super(x_2)|}}$$

où *super(x)* est une fonction qui retourne, pour un concept *x*, l'ensemble des classes auxquelles *x* appartient dans l'ontologie, y compris les superclasses.

À partir de cette distance de similarité, on utilise ensuite la méthode des *k* plus proches voisins (Rani et Vashishtha, 2017) pour sélectionner les *k* cas les plus similaires au cas à résoudre qui sont alors visualisés à l'intention des utilisateurs.

2.5.2. Visualisation des cas similaires

La visualisation des *k* cas les plus similaires utilise la méthode des « boîtes arc en ciel » (Lamy *et al.*, 2017 ; Lamy *et al.*, 2019) et laisse à l'utilisateur du système le choix de la décision thérapeutique qui lui paraît appropriée. Avec les boîtes arc en ciel, les cas similaires sont représentés graphiquement par des colonnes. La largeur des colonnes est proportionnelle à la similarité des cas qu'elles représentent au cas à résoudre. Les colonnes sont regroupées par classes de décisions thérapeutiques et la tête de la colonne est colorée selon la décision thérapeutique adoptée pour le cas similaire représenté. Le cas à résoudre est au milieu. Les attributs significatifs pour le calcul de la distance de similarité sont représentés par des « boîtes ». Chaque boîte recouvre les colonnes correspondant aux cas qui partagent l'attribut. La couleur de la boîte est la moyenne des couleurs associées aux colonnes recouvertes. On décide sur la base de la couleur dominante au niveau de la colonne qui représente le nouveau cas à résoudre.

3. Utilisation de DESIREE sur un cas clinique

3.1. Description du cas et saisie des données

On considère le cas d'une patiente de 58 ans présentant à la palpation clinique une lésion mammaire droite de 20 mm, et une adénopathie axillaire à droite. La mammographie a mis en évidence une lésion du quadrant supéro-externe du sein droit de 20 mm. L'IRM confirme une lésion de 21 mm, dont la biopsie fait le diagnostic d'un carcinome canalaire infiltrant de grade 2, triple négatif (récepteurs aux œstrogènes à 0 %, récepteurs à la progestérone à 0 %, et HER2 négatif), avec un KI67 à 62 %. La cytoponction de l'adénopathie confirme la présence de cellules tumorales. Le cancer est classé cT2N1M0, soit un stade IIB. Les données sont saisies au niveau de l'interface DESIMS (cf. Figure 8) lors de la prise en charge initiale de la patiente.

QUICK VIEW - STATUS - EVOLUTION - DECISION

Lesions

Hide

Lesions

Lesion1: RB, Upper outer quadrant

Patient file

Show

First visit

Hide

Visit date: 2016-06-06

Clinical diagnosis *** Radiological TNM prevailed ***

cTNM: T2 N1 M0

Stage: IIB

T2: Tumor size more than 20 mm but no more (or equal) than 50 mm in greatest dimension
 N1: Micrometastasis / Metastasis 1-3 axillar lymph nodes/ Internal mammary nodes metastasis with detection by sentinel node procedure, no clinical detection
 M0: No metastasis can be assessed (clinical or by diagnostic imaging methods)

Breast info

Bra size: Bra cup:

Lesion	Location	Palpation	Tumor side	Tumor number	Diameter	Ulceration	Skin metastasis
Lesion1	RB: Upper outer	Positive	Right	1	21.0 mm	No	No

Lymph nodes info

	Enlarged	Side
Axillary	Yes	Right

Figure 8. Visualisation des données de la patiente saisies dans le DESIMS.

3.2. Production de recommandations

L'activation du SADC-G produit les différentes options thérapeutiques à partir du GBP NCCN (cf. figure 9). Deux protocoles de chirurgie sont proposés, une mastectomie associée à une exploration axillaire, spécialisée en curage axillaire (de niveau *should* donc recommandé), et à une reconstruction mammaire immédiate (de niveau *may* donc possible) ou bien une tumorectomie, complétée par la pose de clips sur le lit tumoral, associée à une exploration axillaire spécialisée en curage axillaire (de niveau *should* donc recommandé). Un protocole de traitement systémique néo-adjuvant est également possible et dans le cas présent, il est spécialisé par différents

protocoles de chimiothérapie, comportant plusieurs agents cytotoxiques, qui sont recommandés (*should*) ou possibles (*may*).

Dans le cas de l'exécution de règles issues du SADC-E, les résultats peuvent être similaires, seul le statut des propositions de prise en charge change puisqu'elles ne sont pas des recommandations issues de GBP mais l'expression d'une expertise locale comblant les « trous » des GBP.

<p>Recommendation from: R-NCCN-A-Inv-I-IIaT3N1M0-M+LN-Surgery+/-Reconstruction: IF Stage I-II-IIaT3N1M0 THEN Mastectomy + surgical axillary is recommended (Cat 1) + Breast reconstruction is possible (cat 2A) BINV-2/3 [NCCN-GR10-RL02].</p> <p>1) Mastectomy [on 'bs2', SHOULD (1)]</p> <p>1) AxillaryLymphNodeSurgicalProcedure</p> <p>Refinement recommendation from: R-NCCN-A-Inv-Operable-LN-Confirmed-ALND: IF Invasive breast cancer AND Operable breast cancer AND Confirmed Positive nodes THEN AxillaryLymphNodeDissection is recommended (Cat 2A) BINV-D [NCCN-GR09-RL02].</p> <p>1) AxillaryLymphNodeDissection [on 'bs2', SHOULD (2A)]</p> <p>[on 'bs2', SHOULD (1)]</p> <p>1) BreastReconstruction [on 'bs2', MAY (2A)]</p>
<p>Recommendation from: R-NCCN-A-Inv-I-II-L+LN-Surgery: IF Stage I-II THEN Lumpectomy AND surgical axillary staging is recommended (Cat 1) BINV-2 [NCCN-GR10-RL01].</p> <p>1) Lumpectomy</p> <p>Complement recommendation from: R-NCCN-A-G-Clips-after-Lumpectomy: IF breast cancer AND Lumpectomy (recommended OR possible) THEN Clips on the tumor bed during surgery is recommended (Cat 2A) [NCCN-GR06-RL04].</p> <p>1) ClipsOnTumorBed [on 'Lesion1', SHOULD (2A)]</p> <p>[on 'Lesion1', SHOULD (1)]</p> <p>1) AxillaryLymphNodeSurgicalProcedure</p> <p>Refinement recommendation from: R-NCCN-A-Inv-Operable-LN-Confirmed-ALND: IF Invasive breast cancer AND Operable breast cancer AND Confirmed Positive nodes THEN AxillaryLymphNodeDissection is recommended (Cat 2A) BINV-D [NCCN-GR09-RL02].</p> <p>1) AxillaryLymphNodeDissection [on 'bs2', SHOULD (2A)]</p> <p>[on 'bs2', SHOULD (1)]</p>
<p>Recommendation from: R-NCCN-A-Inv-II-IIaT3N1M0-L-Not-CI-except-Tumor-size-NEO-Adj: IF Stage II or IIaT3N1M0 AND Lumpectomy not contraindicated except for tumor size THEN Preoperative systemic therapy is possible (Cat 2A) BINV-1, BINV-10 [NCCN-GR11-RL02].</p> <p>1) BreastCancerSystemicTherapy</p> <p>Refinement recommendation from: R-NCCN-A-Inv-NEO-Adj-HER2-NotT4d-regimens: IF Preoperative systemic therapy is (Possible OR Recommended) AND HER2-negative AND not cT4d THEN ... [NCCN-GR11-RL14].</p> <p>1) Cyclophosphamide x 4 (21 days) [on '060004', SHOULD (1)]</p> <p>1) Docetaxel x 4 (21 days) [on '060004', SHOULD (1)]</p> <p>Refinement recommendation from: R-NCCN-A-Inv-NEO-Adj-HER2-NotT4d-regimens: IF Preoperative systemic therapy is (Possible OR Recommended) AND HER2-negative AND not cT4d THEN ... [NCCN-GR11-RL14].</p> <p>1) Cyclophosphamide x 4 (14 days) [on '060004', SHOULD (1)]</p> <p>1) Doxorubicin x 4 (14 days) [on '060004', SHOULD (1)]</p> <p>2) Paclitaxel x 12 (7 days) [on '060004', SHOULD (1)]</p> <p>Refinement recommendation from: R-NCCN-A-Inv-NEO-Adj-HER2-NotT4d-regimens: IF Preoperative systemic therapy is (Possible OR Recommended) AND HER2-negative AND not cT4d THEN ... [NCCN-GR11-RL14].</p> <p>1) Doxorubicin x 4 (14 days) [on '060004', SHOULD (1)]</p> <p>1) Cyclophosphamide x 4 (14 days) [on '060004', SHOULD (1)]</p> <p>2) Paclitaxel x 4 (14 days) [on '060004', SHOULD (1)]</p> <p>Refinement recommendation from: R-NCCN-A-Inv-NEO-Adj-HER2-NotT4d-regimens: IF Preoperative systemic therapy is (Possible OR Recommended) AND HER2-negative AND not cT4d THEN ... [NCCN-GR11-RL14].</p> <p>1) Cyclophosphamide x 4 (21 days) [on '060004', MAY (1)]</p> <p>1) Doxorubicin x 4 (21 days) [on '060004', MAY (1)]</p> <p>2) Paclitaxel x 12 (7 days) [on '060004', MAY (1)]</p>

Figure 9. Visualisation des recommandations produites par le SADC-G pour le cas clinique décrit.

3.3. Visualisation des cas similaires

On visualise les résultats du raisonnement à partir de cas par la méthode des boîtes arc en ciel. Un exemple (différent de l'exemple précédent) est proposé en figure 10. On voit au niveau du nuage de points centrés sur le cas à résoudre (en blanc) que les cas les plus similaires sont rouge ou jaune, ce qui correspond à une décision de chirurgie (rouge) ou d'hormonothérapie néo-adjuvante (jaune). La lecture des couleurs sur la colonne du cas à résoudre donne une prédominance du rouge ce qui devrait plutôt orienter vers le choix d'une chirurgie.

Figure 10. Visualisation des cas similaires avec les boîtes arc-en-ciel.

4. Discussion et conclusion

Dans le cadre du projet DESIREE, nous avons développé trois modules d'aide à la décision clinique afin de proposer une aide à la décision fondée sur les GBP, une aide à la décision fondée sur l'expérience et une aide à la décision fondée sur la mise en œuvre d'un raisonnement à partir de cas. Les GBP sélectionnées sont les GBP américains (NCCN), ceux de l'Assistance Publique – Hôpitaux de Paris, et des GBP espagnols élaborés par le centre de soins Onkologikoa, partenaire du projet.

Les développements ont été réalisés en utilisant les outils du web sémantique. Nous avons construit une ontologie du domaine utilisée comme structure conceptuelle et terminologique afin de proposer un modèle de données (EAV) et un modèle de connaissances, utilisée pour le raisonnement et l'aide à la décision. Le BCKM joue ainsi un rôle pivot dans la gestion des données et des connaissances. D'autres travaux visent également à combiner modèles d'information et ontologies (Ethier *et al.*, 2015 ; Martínez-Costa et Schulz, 2017). Le BCKM permet d'assurer la conformité au modèle de la représentation des données patient et de l'écriture des règles de décision des bases de connaissances, que celles-ci viennent de la formalisation des GBP pour le SADC-G, ou aient été apprises par l'expérience avec le SADC-E. Les transformations des données et des règles de décision en notation N3 permettent l'utilisation d'un moteur d'inférence sémantique, Euler/EYE, combinant raisonnement ontologique et par règles. Cette approche permet notamment des raisonnements autorisant des recommandations de prise en charge thérapeutique des patients à niveaux d'abstraction variables, patients dont les profils cliniques peuvent également être décrits à des niveaux d'abstraction variables. Par ailleurs, le BCKM est aussi utilisé pour la recherche de patients similaires par le SADC-RC, la distance entre attributs catégoriels étant calculée comme un degré de proximité taxonomique.

Le DESIMS dans lequel sont entrées et stockées les données a été conçu préalablement au BCKM avec un modèle d'information différent de celui adopté dans le BCKM et un référentiel terminologique également différent, illustrant le problème, récurrent dans le domaine biomédical, de l'interopérabilité syntaxique et

sémantique. Si l'utilisation du standard FHIR a permis de simplifier la résolution des aspects syntaxiques de l'interopérabilité, les aspects sémantiques ont nécessité de procéder manuellement à des alignements entre les termes du référentiel du DESIMS et les concepts du BCKM. Ce constat plaide une fois encore en faveur du partage de référentiels termino-ontologiques, mais également de modèles d'informations cliniques (Becnel *et al.*, 2017), éventuellement dédiés à un domaine particulier comme le cancer (Hochheiser *et al.*, 2016).

La conception du BCKM s'est faite dans le cadre du développement d'un prototype de recherche avec la contrainte forte de représenter des données réelles issues d'un dossier s'appuyant sur un modèle relationnel (le DESIMS). Nous avons choisi d'utiliser le modèle de données EAV, qui est simple mais générique. Par ailleurs, nous avons choisi d'explicitier chaque élément du modèle EAV sous forme de classes et de pas exploiter les possibilités d'OWL où typiquement les attributs auraient pu être représentés par des DataProperties pour les types primitifs et des ObjectProperties pour les types hiérarchiques. Ceci permet de rendre générique la relation `hasRange` entre une classe `ModelAttribute` et une classe `ModelValue` et de la traiter uniformément quel que soit le type de valeur. Cette particularité est exploitée par exemple lors de la construction automatique de formulaires de saisie basés sur le modèle ontologique (Sadki *et al.*, 2018). Par ailleurs, cette représentation par classes permet d'importer des concepts/classes d'autres ontologies et de leur donner dans le BCKM le rôle que l'on souhaite : entité, attribut ou valeur. Ainsi, nous avons importé dans le BCKM de nombreux concepts et hiérarchies du NCI thesaurus avec l'outil Flat OWL Editor. On notera que dans ce cas, le lien entre le NCI thesaurus et ses classes « clonées » n'est pas conservé dynamiquement comme cela peut être le cas avec un outil comme OntoFox (Xiang *et al.*, 2010) ; il l'est cependant de manière statique puisque le `NCI_CODE` est conservé en annotation. Toutefois, la conservation dynamique du lien d'importation n'a pas été recherchée d'une part parce que la structuration du NCI thesaurus n'était pas en l'état satisfaisante du point de vue de notre application (l'organisation de certaines hiérarchies a dû être revue manuellement), et d'autre part, le codage des données existantes pour une certaine version du BCKM et de ses concepts pourrait ne pas résister aux évolutions du BCKM dont les évolutions doivent être contrôlées. Dans l'état actuel, la qualité de l'ontologie est satisfaisante dans le sens où elle permet la représentation des informations utiles pour les prises de décision, qu'elles proviennent des données ou des GBP, et la production automatique de recommandations. Cette satisfaction est obtenue au travers des étapes de validation technique et clinique du système.

Le SADC-G produit les recommandations telles qu'elles sont exprimées dans les GBP. Des études ont décrit des incohérences intra-GBP qu'il s'agit de gérer. Par ailleurs, en proposant les recommandations issues de trois GBP différents, on s'expose à une incohérence inter-GBP. Des solutions de réconciliations ont été développées au sein des règles de décision (*complement* et *refine*). Il faudra améliorer la proposition des recommandations par un post traitement permettant la résolution des conflits tout en bénéficiant de la complémentarité des GBP. Le SADC-E est actuellement développé pour servir la construction des règles apprises par l'expérience. La mise en œuvre de questionnaires patients permettant d'évaluer

la qualité clinique des règles apprises (calcul du poids de la preuve) doit être organisée. Enfin, le SDC-RC utilise une mesure globale de la similarité entre les cas basée sur une moyenne arithmétique des similarités locales. Les critères n'ayant pas le même poids pour établir la similarité sur le plan médical, il est prévu d'apprendre les poids des critères lorsque le peuplement de la base de cas sera suffisant.

La plateforme DESIREE a fait l'objet d'une évaluation sur cas simulés par un échantillon de huit cliniciens. Les services offerts ont reçu un bon accueil mais des améliorations ont été demandées. Elles sont actuellement en cours de développement. L'évaluation sur vrais patients mais sur RCP simulées est prévue à partir de février 2019, et l'évaluation en vraie grandeur sera organisée sur quatre sites (Onkologikoa à San Sébastien en Espagne, ERESA à Valence en Espagne, l'hôpital Saint-Louis et l'hôpital européen Georges Pompidou de l'Assistance Publique – Hôpitaux de Paris) à compter du mois d'avril 2019.

Remerciements

Les auteurs remercient l'ensemble des partenaires du projet DESIREE.

Bibliographie

- Abidi S.R., Abidi S.S., Hussain S., Shepherd M. (2007). Ontology-based modeling of clinical practice guidelines: a clinical decision support system for breast cancer follow-up interventions at primary care settings. *Stud Health Technol Inform*, vol 129(Pt 2), p. 845–849.
- AP-HP. (2016). Référentiel cancer du sein, mars 2016. Référentiels de l'AP-HP. https://www.aphp.fr/sites/default/files/referentiel_cancers_du_sein_-_juin_2016_1.pdf
- Becnel L.B., Hastak S., Ver Hoef W., Milius R.P., Slack M., Wold D., et al. (2017). BRIDG: a domain information model for translational and clinical protocol-driven research. *J Am Med Inform Assoc*, vol 24, n° 5, p. 882–890.
- Berners-Lee T., Connolly D., Kagal L., Scharf Y., Hendler J. (2008). N3Logic: A logical framework for the World Wide Web. *Theory and Practice of Logic Programming*, vol 8, n° 3, p. 249–269.
- Boscá D., Maldonado J.A., Moner D., Robles M. (2015). Automatic generation of computable implementation guides from clinical information models. *J Biomed Inform*, vol 55, p. 143–152.
- Bouaud J., Guézennec G., Séroussi B. (2018). Combining the Generic Entity-Attribute-Value Model and Terminological Models into a Common Ontology to Enable Data Integration and Decision Support. *Stud Health Technol Inform*, vol 247, p. 541–545.
- Brar S.S., Hong N.L., Wright F.C. (2014). Multidisciplinary cancer care: Does it improve outcomes? *J Surg Oncol*, vol 110, n° 5, p. 494–499.
- Despres S., Guézennec G. (2017). Flat OWL Editor : un outil utilisant des feuilles de calcul pour découpler les tâches des acteurs impliqués dans la gestion d'une ontologie. Terminologie & Ontologie: Théories et Applications, Actes de la conférence TOTH 2017, Jun 2017, Chambéry, France. Université Savoie Mont Blanc, p. 181-194.

- Ethier J.F., Curcin V., Barton A., McGilchrist M.M., Bastiaens H., Andreasson A., et al. (2015). Clinical data integration model. Core interoperability ontology for research using primary care data. *Methods Inf Med*, vol 54, n° 1, p. 16–23.
- Elsamany S., Abdullah S. (2014). Triple-negative breast cancer: future prospects in diagnosis and management. *Med Oncol*, vol 31, n° 2, 834.
- Eccher C., Seyfang A., Ferro A. (2014). Implementation and evaluation of an Asbru-based decision support system for adjuvant treatment in breast cancer. *Comput Methods Programs Biomed*, vol 117, n° 2, p. 308–21.
- Giguère A., Légaré F., Grimshaw J., Turcotte S., Fiander M., Grudniewicz A., et al. (2012). Printed educational materials: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;10:CD004398.
- Hébert-Croteau N., Brisson J., Latreille J., Rivard M., Abdelaziz N., Martin G. (2004). Compliance with consensus recommendations for systemic therapy is associated with improved survival of women with node-negative breast cancer. *J Clin Oncol*, vol 22, n° 18, p. 3685–3693.
- HL7.org. (2015). HL7 standard: Clinical decision support knowledge artifact specification, release 1.3. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=337
- HL7.org. (2017). FHIR Release 3 (STU; v3.0.1-11917), <http://hl7.org/fhir/STU3/>.
- Hochheiser H., Castine M., Harris D., Savova G., Jacobson R.S. (2016). An information model for computable cancer phenotypes. *BMC Med Inform Decis Mak*, vol 16 n° 1, p. 121.
- Jaspers M.W., Smeulders M., Vermeulen H., Peute L.W. (2011). 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*, vol 18, n° 3, p. 327–334.
- Johnson C.E., Slavova-Azmanova N., Saunders C. (2017). Development of a peer review framework for cancer multidisciplinary meetings. *Intern Med J*, vol 47, n° 5, p. 529–535.
- Keating N.L., Landrum M.B., Lamont E.B., Bozeman S.R., Shulman L.N., McNeil B.J. (2013). Tumor boards and the quality of cancer care. *J Natl Cancer Inst*, vol 105, n° 2, p. 113–21.
- Kesson E.M., Allardice G.M., George W.D., Burns H.J., Morrison D.S. (2012). Effects of multidisciplinary team working on breast cancer survival: retrospective, comparative, interventional cohort study of 13 722 women. *BMJ*, vol 344, e2718.
- Khan O., Lim Choi Keung S.N., Zhao L., Arvanitis T.N. (2014). A Hybrid EAV-Relational Model for Consistent and Scalable Capture of Clinical Research Data. *Stud Health Technol Inform*, vol 202, p. 32–35.
- Lamy J.B., Berthelot H., Capron C., Favre M. (2017). Rainbow boxes: a new technique for overlapping set visualization and two applications in the biomedical domain, *Journal of Visual Language and Computing*, vol. 43, p. 71–82.
- Lamy J.B., Sekar B., Guézennec G., Bouaud J, Séroussi B. (2019). Explainable artificial intelligence for breast cancer: a visual case-based reasoning approach. *Artif Intell Med*, vol 94, March 2019, p. 42-53.

- Lebeau M., Mathoulin-Pélissier S., Bellera C., Tunon-de Lara C., Daban A., Lipinski F., et al. (2011). Breast cancer care compared with clinical guidelines: an observational study in France. *BMC Public Health*, vol 11, 45.
- Lee W.S., Ahn S.M., Chung J.W., Kim K.O., Kwon K.A., Kim Y., et al. (2018). Assessing Concordance With Watson for Oncology, a Cognitive Computing Decision Support System for Colon Cancer Treatment in Korea. *JCO Clin Cancer Inform*, vol 2, p. 1-8.
- Löper D., Klettke M., Bruder I., Heuer A. (2013). Enabling flexible integration of healthcare information using the entity-attribute-value storage model. *Health Inf Sci Syst*, vol 1, 9.
- Martínez-Costa C., Schulz S. (2017). Validating EHR clinical models using ontology patterns. *J Biomed Inform*, vol 76, p. 124-137.
- Minsky M. (1975). A framework for representing knowledge. *The Psychology of Computer Vision*. Ed. Patrick Winston, Mc Graw Hill.
- Muro N., Larburu N., Bouaud J., Seroussi B. (2017). Weighting Experience-Based Decision Support on the Basis of Clinical Outcomes' Assessment. *Stud Health Technol Inform*, vol 244, p. 33-37.
- Nadkarni P.M., Marengo L., Chen R., Skoufos E., Shepherd G., Miller P. (1999). Organization of Heterogeneous Scientific Data Using the EAV/CR Representation. *J Am Med Inform Assoc*, vol 6, n° 6, p. 478-493.
- Nentwich C., James R. (2010). Natural rule language (NRL) specification 1.4.0. <http://nrl.sourceforge.net/spec/>
- NCCN.org. (2017). *Breast Cancer. Version 3.2017*. NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®).
- Patkar V., Acosta D., Davidson T., Jones A., Fox J., Keshtgar M. (2011). Cancer multidisciplinary team meetings: evidence, challenges, and the role of clinical decision support technology. *Int J Breast Cancer*, vol 2011, 831605.
- Patkar V., Acosta D., Davidson T., Jones A., Fox J., Keshtgar M. (2012). Using computerised decision support to improve compliance of cancer multidisciplinary meetings with evidence-based guidance. *BMJ Open*, vol 2, n° 3, e000439.
- Peleg M. (2013). Computer-interpretable clinical guidelines: a methodological review. *J Biomed Inform*, vol 46, n° 4, p. 744-763.
- Rani P., Vashishtha J. (2017). An appraisal of KNN to the perfection. *International Journal of Computer Applications*, vol 170, n° 2, p. 13-17.
- Roshanov P.S., Misra S., Gerstein H.C., Garg A.X., Sebaltd RJ, Mackay JA, et al. (2011). Computerized clinical decision support systems for chronic disease management: a decision-maker-researcher partnership systematic review. *Implement Sci*, vol 6, 92.
- Sadki F., Bouaud J., Guézennec G., Séroussi B. (2018). Semantically structured web form and data storage: a generic ontology-driven approach applied to breast cancer. *Stud Health Technol Inform*, vol 255, p. 205-209.
- Sackett D.L., Rosenberg W.M., Gray J.A., Haynes R.B., Richardson W.S. (1996). Evidence based medicine: what it is and what it isn't. *BMJ*, vol 312, n° 7023, p. 71-72.
- Schank R.C., Abelson R.P. (1977). *Scripts, plans, goals and understanding: an inquiry into human knowledge structures*. L. Erlbaum, Hillsdale, NJ.

- Schwentner L., Wolters R., Koretz K., Wischnewsky M.B., Kreienberg R., Rottscholl R., et al. (2012). Triple-negative breast cancer: the impact of guideline-adherent adjuvant treatment on survival—a retrospective multi-centre cohort study. *Breast Cancer Res Treat*, vol 132, n° 3, p. 1073–1080.
- Sekar B.D., Lamy J.B., Larburu N., Séroussi B., Guézennec G., Bouaud J., Muro N., Wang H., Liu J. (2018). Case-Based Decision Support System for Breast Cancer Management. *International Journal of Computational Intelligence Systems*, Vol. 12 (2018) 28-38.
- Séroussi B., Bouaud J., Antoine EC. (2001). OncoDoc: a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. *Artif Intell Med*, vol 22, n° 1, p. 43–64.
- Séroussi B., Laouénan C., Gligorov J., Uzan S., Mentré F., Bouaud J. (2013). Which breast cancer decisions remain non-compliant with guidelines despite the use of computerized decision support? *Br J Cancer*, vol 109, n° 5, p. 1147–1156.
- Séroussi B., Guézennec G., Lamy J.B., Muro N., Larburu N., Sekar B., et al. (2017). Reconciliation of multiple guidelines for decision support: a case study on the multidisciplinary management of breast cancer within the DESIREE project. *AMIA Annu Symp Proc* 2017, p. 1527–1536.
- Séroussi B., Lamy J.B., Muro N., Larburu N., Sekar B.D., Guézennec G., Bouaud J. (2018). Implementing Guideline-Based, Experience-Based, and Case-Based Approaches to Enrich Decision Support for the Management of Breast Cancer Patients in the DESIREE Project. *Stud Health Technol Inform*, vol 255, p. 190–194.
- Shahar Y., Miksch S., Johnson P. (1998). The Asgaard project: a task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artif Intell Med*, vol 14 n° 1-2, p. 29–51.
- Shiffman R.N., Michel G., Essaihi A., Thornquist E. (2004). Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. *J Am Med Inform Assoc*, vol 11, n° 5, p. 418–426.
- Simon G., DiNardo C.D., Takahashi K., Cascone T., Powers C., Stevens R., et al. (2018). Applying Artificial Intelligence to Address the Knowledge Gaps in Cancer Care. *Oncologist*. pii: theoncologist.2018-0257.
- van Hove J., de Munck L., Otter R., de Vries J., Siesling S. (2014). Quality improvement by implementing an integrated oncological care pathway for breast cancer patients. *Breast*, vol 23, n° 4, p. 364–370.
- Verborgh R., De Roo J. (2015). Drawing conclusions from linked data on the web: The EYE reasoner. *IEEE Software*, vol 32, n° 3, p. 23–27.
- Witten I. H., Frank E., Trigg L., Hall M., Holmes G., Cunningham S. J. (1999). Weka: Practical Machine Learning Tools and Techniques with Java Implementations, Proceedings of the ICONIP/ANZIIS/ANNES'99 Workshop on Emerging Knowledge Engineering and Connectionist-Based Information Systems, p. 192–196.
- Wöckel A., Kurzeder C., Geyer V., Novasphenny I., Wolters R., Wischnewsky M., et al. (2010). Effects of guideline adherence in primary breast cancer—a 5-year multi-center cohort study of 3976 patients. *Breast*, vol 19, n° 2, p. 120–127.

Xiang Z., Courtot M., Brinkman, R.R., Ruttenberg A., He Y. (2010). OntoFox: web-based support for ontology reuse. *BMC research notes*, vol 3, n°175.

Brigitte Séroussi est titulaire d'un diplôme d'ingénieur de l'école Centrale Paris, d'un doctorat en médecine de l'Université Paris 5 (France) et d'un doctorat de Biomathématiques de l'Université Paris 7 (France). Elle est actuellement Professeur des Universités à Sorbonne Université (Paris, France) et Praticien Hospitalier à l'Assistance Publique - Hôpitaux de Paris (Hôpital Tenon) en charge de l'évaluation des pratiques et de la qualité des soins. Elle est membre du LIMICS (Sorbonne Université, Paris 13, INSERM). Ses travaux de recherche sont centrés sur la modélisation, le développement, la mise en œuvre et l'évaluation des systèmes d'aide à la décision clinique basés sur les guides de pratique clinique. Elle a coordonné et participé à de nombreux projets nationaux et internationaux (dont européens).

Jean-Baptiste Lamy est docteur en pharmacie et en santé publique, et enseignant-chercheur en informatique à l'Université Paris 13 (Bobigny, France) et au laboratoire LIMICS (Paris, France). Ses recherches portent sur l'aide à la décision, la représentation des connaissances, les ontologies et la visualisation des connaissances, avec un domaine d'application médical. Il conçoit notamment des systèmes visuels pour aider à la décision thérapeutique, ainsi que des langages iconiques pour la médecine.

Coralie Prebet est médecin, assistante en Oncologie Médicale à l'Hôpital Tenon, Assistance Publique - Hôpitaux de Paris (Paris, France) et Sorbonne Université (Paris, France), spécialisée dans la prise en charge du cancer du sein et des cancers urologiques. Elle a participé au projet DESIREE H2020 en contribuant à l'implémentation des recommandations de bonne pratique internationales sur la prise en charge du cancer du sein. Ses domaines d'intérêts portent notamment sur l'amélioration de la prise en charge des patients atteints de cancer en facilitant l'accès et l'utilisation des données actualisées de la science par les praticiens.

Charlotte Ngo est Maître de Conférences des Universités et Praticien Hospitalier en chirurgie cancérologique gynécologique et mammaire à l'Hôpital Européen Georges Pompidou au sein de l'Assistance Publique – Hôpitaux de Paris (Paris, France). Depuis plus de 10 ans, son activité clinique est centrée sur la prise en charge des cancers féminins et plus particulièrement du cancer du sein. Ses projets de recherche explorent les champs de l'innovation chirurgicale et de la réduction des inégalités sociales de santé. Elle enseigne également la cancérologie à l'Université Paris-Descartes, Université de Paris (France).

Luis Teixeira est Maître de Conférences des Universités et Praticien Hospitalier en Oncologie Médicale à l'Université de Paris (France). Il exerce au Sénopôle Saint Louis Territoire Cancer Nord de l'Hôpital Saint Louis à Paris (France), centre expert et premier centre en sénologie pour l'Assistance Publique – Hôpitaux de Paris. Il est affilié à l'Equipe 6 "Physiopathologie des cancers du sein", INSERM U976. Il a effectué ses études médicales à l'Université Pierre et Marie Curie à Paris. Il a une formation initiale en Médecine Interne, puis en Cancérologie. Ses principaux centres d'intérêts portent sur la recherche translationnelle et la recherche clinique dans le cancer du sein.

Amaia Ugarriza est diplômé en informatique. Elle a été responsable de nombreux projets au cours des douze dernières années dans l'unité de Recherche et Développement de l'entreprise espagnole Bilbomática. Ses travaux de recherche portent sur les ontologies, les standards FHIR-HL7 et l'expérience utilisateur UX, en particulier dans le domaine des traitements multidisciplinaires et de la gestion de l'information clinique hétérogène dans le milieu

hospitalier. Ses recherches ont été appliquées à la prise en charge de la maladie d'Alzheimer, du cancer du sein, de la cardiologie et du déclin cognitif.

Booma Devi Sekar est titulaire d'un doctorat de l'Université de Macao (Macao) et travaille actuellement comme chercheur associé à la School of Computing de l'Université d'Ulster (Royaume-Uni). Ses sujets de recherche portent sur l'analyse des données, l'apprentissage automatique, les méthodes d'intelligence artificielle, le traitement du signal et la théorie de l'information, ainsi que sur la recherche appliquée dans les domaines des machines électriques, des télécommunications, et des soins de santé.

Nekane Larburu est titulaire d'un Master en Ingénierie des Télécommunications de l'Université de Navarre (Espagne). Elle s'est fortement impliquée dans le projet européen Mobiguide qui visait à élaborer un système d'aide à la décision personnalisée "mobile". Actuellement, ses activités de recherche se concentrent sur l'application des techniques d'ingénierie des connaissances pour l'élaboration de systèmes personnalisés dans le domaine de la santé au sein du Département des Applications Biomédicales pour la eSanté du Centre de Recherche Vicomtech (Espagne).

Naiara Muro Amuchastegui est chercheur au sein du Département des Applications Biomédicales pour la eSanté du Centre de Recherche Vicomtech (Espagne) depuis 2014. Elle est titulaire d'un diplôme d'ingénieur biomédical de l'Université de Navarre (Espagne) et d'une maîtrise en génie informatique et systèmes intelligents de l'Université du Pays Basque (Espagne). Elle est affiliée au groupe eHealth du Département de bio-ingénierie de l'Institut de recherche biosanitaire Biodonosti. Ses travaux de recherche portent sur l'extraction et l'application de nouvelles connaissances dans les systèmes d'aide à la décision médicale.

Gilles Guezennec a longtemps exercé en tant qu'ingénieur logiciel dans les domaines de l'informatique éditoriale et du traitement de données. Il a rejoint le LIMICS en 2013 comme Ingénieur de Recherche où ses travaux se sont spécialisés dans la mise en œuvre des techniques du Web Sémantique et des Ontologies dans différents projets de recherche français et européens. La construction d'ontologies, l'annotation sémantique et le développement d'outils d'aide à la décision constituent ses champs d'activité actuels.

Jacques Bouaud est titulaire d'un doctorat en biomathématiques de l'Université Paris-Diderot. Il est chargé de mission à la Délégation de la Recherche Clinique et de l'Innovation de l'Assistance Publique-Hôpitaux de Paris et membre du LIMICS. Son intérêt porte sur la modélisation de connaissances, l'implémentation d'outils d'aide à la décision dans le domaine médical s'appuyant sur les recommandations de bonne pratique, ainsi que sur l'évaluation de ces outils. Les domaines d'application ciblés sont principalement les prises en charge thérapeutiques pour le cancer du sein et les maladies chroniques.