



Disponible en ligne sur

ScienceDirect
www.sciencedirect.com

Elsevier Masson France

EM|consulte
www.em-consulte.com



ATELIERS DE GIENS 2016/PHARMACOLOGIE SOCIALE

Intérêts et limites des communautés virtuelles de patients pour la recherche sur les produits de santé[☆]

Sophie Ravoire^{a,*}, Marie Lang^b, Elena Perrin^c, les participants à la table ronde n° 6 de Giens XXXII, Antoine Audry^d, Pascal Bilbault^e, Michael Chekroun^f, Lauren Demerville^g, Thierry Escudier^h, Laure Guérout-Accolasⁱ, Caroline Guillot^j, Muriel Malbezin^k, Philippe Maugendre^l, Joëlle Micallef^m, Mathieu Molimardⁿ, François Montastruc^o, Evelyne Pierron^p, Lionel Reichardt^q, Frantz Thiessard^r

^a SR Consulting, 54, rue Sébastien-Mercier, 75015 Paris, France

^b GCS CNCR, 75014 Paris, France

^c Lilly France, 92521 Neuilly-sur-Seine, France

^d Medtronic France, 92513 Boulogne Billancourt, France

^e Boehringer-Ingelheim France, 51721 Reims, France

^f Carenity, 75008 Paris, France

^g DRCD, AP–HP, 75010 Paris, France

^h Institut de recherche Pierre-Fabre, 31035 Toulouse, France

ⁱ Association patients en réseau, 78100 Saint-Germain-en-Laye, France

^j Diabète LAB, 75012 Paris, France

^k Hospices civils de Lyon, DRCL, 69229 Lyon, France

^l Sanofi, 94255 Gentilly, France

^m Hôpital de la Timone, AP–HM, 13005 Marseille, France

ⁿ Université de Bordeaux, CHU de Bordeaux, 33076 Bordeaux, France

^o Université de Toulouse, CHU de Toulouse, 31000 Toulouse, France

^p EP consultants, 75009 Paris, France

^q 7C'S health, 92190 Meudon, France

^r ISPED, université de Bordeaux, CHU de Bordeaux, 33000 Bordeaux, France

Reçu le 22 novembre 2016 ; accepté le 28 novembre 2016

Disponible sur Internet le 6 décembre 2016

[☆] Les articles, analyses et propositions issus des Ateliers de Giens n'engagent que leurs auteurs et ne préjugent pas de la position de leur organisme de tutelle.

* Auteur correspondant.

Adresse e-mail : s.ravoire@sr-consulting.fr (S. Ravoire).

<http://dx.doi.org/10.1016/j.therap.2016.11.055>

0040-5957/© 2016 Société française de pharmacologie et de thérapeutique. Publié par Elsevier Masson SAS. Tous droits réservés.

MOTS CLÉS

France ;
 Recherche sur les produits de santé ;
 Réseaux sociaux ;
 Communautés virtuelles ;
 Communautés Web ;
 Patients ;
 Pharmacovigilance ;
 Addictovigilance ;
 Données sur Internet ;
 Post-inscription ;
 Pharmaco-épidémiologie

Résumé Le partage et l'échange d'informations sur des sujets de santé, les paroles des patients et de leurs aidants sur différentes plateformes Internet, ainsi que leurs savoirs expérientiels représentent potentiellement de nouveaux environnements de recherche, en particulier sur les produits de santé. Les informations fournies dans le cadre d'une démarche individuelle et volontaire par les patients inscrits sur ces communautés « virtuelles » ou réseaux en ligne, sont autant de sources pour appréhender et comprendre finement l'usage des produits de santé, évaluer leur efficacité et quantifier les possibles effets indésirables en conditions réelles, détecter des signaux faibles de pharmacovigilance ou « d'addictovigilance », mais aussi développer de nouveaux outils d'évaluation afin de générer de nouvelles hypothèses de travail. L'expression libre du vécu, du ressenti des patients et de la réalité des expériences partagées ouvrent également le champ aux recherches sociétales autour des produits de santé, champ encore peu exploré à ce jour. La recherche sur les produits de santé est soumise à une réglementation établie et fait appel à des moyens et des méthodologies qui jusqu'ici n'ont que peu évolué. Dans ce domaine, le développement des communautés de patients en ligne offre de nouvelles perspectives, l'enjeu aujourd'hui étant de définir leur place au sein des moyens de recherche conventionnels. Cette place ne pourra être acceptée par toutes les parties prenantes qu'à condition d'une bonne compréhension de leurs atouts, limites et contraintes. La table ronde réunie sur ce sujet s'est attachée à explorer et mieux comprendre l'ampleur du phénomène, les différentes typologies existantes des communautés virtuelles et des réseaux en ligne de patients, d'identifier les avantages possibles, les spécificités et les limites méthodologiques, réglementaires et déontologiques auxquelles les chercheurs sont confrontés actuellement et enfin, d'émettre les premières recommandations dans ce domaine de recherche en croissance.

© 2016 Société française de pharmacologie et de thérapeutique. Publié par Elsevier Masson SAS. Tous droits réservés.

Abréviations

ANSM	Agence nationale de sécurité des médicaments et des produits de santé
CPP	Comité consultatif de protection des personnes
DRCI	délégation à la recherche clinique et à l'innovation
FDA	Food and Drug Administration
FMTI	<i>French multi-terminology indexer</i>
MedDRA	<i>medical dictionary for regulatory activities</i>
PRO/PRI	<i>patient reported outcomes/patient reported information</i>
SMTS	serveur multi-terminologies en santé
SNOMED	<i>systematic nomenclature medical clinical terms</i>
TAL	traitement automatique du langage
UMLF	<i>unified medical lexicon for French</i>
UMLS	<i>unified medical language system</i>

Introduction

L'appropriation par les usagers des nouvelles fonctionnalités du web, connues sous le nom de web 2.0, en offrant plus de possibilités d'interactions, d'échanges d'informations a permis ces dernières années l'émergence de différents systèmes sociaux, tels le réseautage social, les blogs, les wikis et les plateformes interactives communautaires. La santé au cœur des préoccupations quotidiennes est rapidement devenue un sujet de prédilection des internautes surfant sur ces sites, traduisant la forte volonté des individus à devenir acteurs de leur prise en charge médicale. La simplicité de l'outil, la multiplicité des possibilités, de

la simple recherche d'informations jusqu'à la participation active de la construction des savoirs, en passant par le partage d'expériences – engagent les patients et leurs aidants à sortir de leur sphère intime et à devenir eux-mêmes source d'informations. La table ronde réunie sur ce sujet d'actualité en forte évolution s'est fixée pour objectifs de clarifier les différentes typologies de systèmes ou communautés virtuelles de patients identifiables sur la toile, d'en comprendre pour chaque type les spécificités, d'en dégager les potentialités, les intérêts et les limites dans une perspective d'utilisation et de valorisation pour la recherche de ces supports et des données qu'ils produisent et de création de partenariats de recherche. Ces réflexions ont été conduites en ayant conscience que le sujet recouvrait certains aspects de propriété et d'accessibilité des données et en général de la problématique des entrepôts de données ou base de données patients et des *big data* ou *open data* traitée en 2015 par une autre table ronde [1]. Il est aussi certain que la perpétuelle créativité des usagers, leur place dans l'évaluation des solutions de santé, les évolutions technologiques informatiques mais également les innovations issues de la télémédecine, l'e-santé, l'intelligence artificielle, les objets connectés, continueront d'alimenter et amener ces réflexions.

Les membres de la table ronde ont également tenté, en croisant leurs expériences issues de divers milieux professionnels (associations de patients, plateformes commerciales, consultants, cliniciens, directions de recherche clinique, pharmaco-épidémiologistes, industriels de la santé, etc.) et après analyse de la littérature disponible, de faire émerger certains des facteurs clés de l'apport des communautés virtuelles de patients à la recherche, ceci

dans le but de proposer des recommandations pouvant servir de référence pour l'ensemble des opérateurs du domaine et de la communauté des chercheurs.

Le paysage actuel des communautés virtuelles de patients

Ce paysage semble aujourd'hui très vaste et sans limite en raison de la croissance exponentielle des outils ainsi que des données, de l'appétence décomplexée des usagers pour le domaine de la santé et de l'explosion du nombre de supports pour faciliter l'expression des patients et leurs aidants, l'ensemble contribuant à un « data déluge », largement inexploité, ainsi :

- 90 % des données de l'humanité ont été générées au cours des deux dernières années (source IBM) ;
- en 2015, 86 % des français souhaitaient, en cas de maladie, échanger avec des patients ayant la même pathologie [2] ;
- en 2016, 68 % des français ont consulté Internet afin de trouver des informations médicales. Les principaux sites consultés sont des sites informels et collaboratifs (87 %), les sites généralistes dédiés à la santé tels que Doctissimo (69 %), les sites uniquement collaboratifs (blogs, forums et réseaux sociaux) dans 26 % des cas [3].

Sous des dénominations diverses : communautés virtuelles, communautés en ligne, communautés de pratiques, sites collaboratifs, blogs, forums, réseaux sociaux, ces initiatives collectives de patients et/ou d'aidants, sont plus ou moins structurées pour permettre des « livrables » à des fins de recherche, complémentaires des données issues des méthodes plus classiques de la science et de la pratique médicale.

La grande majorité de ces communautés n'ont pas pour finalité des activités de recherche. Il est cependant possible d'identifier un gradient selon leur degré de maturation et de structuration des données, une discrimination entre une composante commerciale (ex. : PatientsLikeMe, Carenity), associative (ex. : « Mon réseau cancer du sein », « Diabète LAB », « Renaloo »), ou académique (ex. : « ComPaRe », « Seintinelles ») n'étant pas retenue comme pertinente dans l'approche méthodologique choisie par la table ronde.

Ce qui caractérise l'ensemble de ces communautés « virtuelles » de patients, de toute taille, est la démarche active individuelle des patients et/ou de leurs aidants (pour s'y inscrire, s'y connecter et participer), d'où la terminologie *patient reported outcome* (PRO) ou *patient reported information* (PRI) fréquemment rencontrée et dont la traduction francophone pourrait être : « résultat ou information rapporté(e) par le patient lui-même » [4].

C'est, en effet, bien cet acte volontaire et l'apport du patient qui va générer des données, même si, dans le cas de Facebook, Twitter ou des forums et blogs de discussion, les patients ne sont pas toujours pleinement conscients de l'utilisation possible de ces données par les opérateurs.

Facebook, 1^{re} communauté virtuelle au monde de 1,5 milliard d'individus permet aux patients de s'organiser autour de plusieurs forums par aires thérapeutiques, en plateformes communautaires d'échanges entre patients (ex. : « Diabète côté Femmes » – plus de 31 000 membres). Sur

une période de deux ans, 2000 communautés ou sujets de santé ont été répertoriés sur Twitter, le caractère totalement public des tweets ne semblant pas être un frein à l'expression, puisque ce nombre est en constante progression.

Les forums et blogs, pour certains à l'origine d'initiatives individuelles, sont également de nouveaux canaux d'expression, de niveaux divers, plus ou moins structurés, où la part de témoignages est prépondérante et l'information empruntée à diverses sources (scientifiques, pratiques, personnelles) mêlant plusieurs postures énonciatives.

Enfin les très nombreuses plateformes et communautés virtuelles de patients, avec ou sans composante commerciale, associative ou privée, pour certaines fortement structurées, qui transforment les histoires personnelles en de vastes bases de données générées à partir des données que fournissent les patients, permettent des analyses de natures variées, par exemple démographiques, épidémiologiques, sociologiques ou cliniques (Fig. 1).

Traitement des données

Dès lors, une fois le support sélectionné, fonction de l'objet de la recherche, se pose la question des méthodes de collecte, d'extraction et d'analyse de la donnée à mobiliser. Certaines de ces méthodes font appel à des analyses qualitatives de contenu sur les données textuelles, et les difficultés sont multiples :

- l'accès et la récupération des données (aspects techniques liés au volume de données, à l'accès possible ou non à l'ensemble des données brutes ; aspects légaux du *data mining*, *text mining*, du copyright, de l'utilisation des « bots » d'aspiration des données) ;
- la reconstruction des blocs d'informations et des fils de discussion (l'information dans un message peut faire référence au titre du fil de discussion voire au titre du forum pour lesquels il est parfois nécessaire d'analyser ces différents niveaux d'information conjointement) ou choix d'un traitement de l'information déstructurée en « sacs de termes », « nuages de points ou de mots » ;
- le choix du type d'analyse : recherche du sens par l'analyse de milliers de co-occurrences de termes ou traitement automatique du langage (TAL) avec une analyse morphosyntaxique, lexicale et sémantique (repérage des syntagmes nominaux, verbaux, prépositionnels et adjectivaux et leurs relations ; repérage des négations, de la temporalité et du niveau de certitude : utilisation du conditionnel, notion d'hypothèse, de suspicion, etc.) les techniques de TAL étant très dépendantes de la langue et nécessitant pour leur évaluation de grandes quantités de corpus annotés manuellement ;
- la constitution d'un thésaurus ou glossaire patient relié à une terminologie de référence (une conversation de type « j'ai passé une nuit blanche » peut ainsi alors être raccrochée au terme médical d'« insomnie » et être agrégée et traitée), les patients n'utilisant pas le dictionnaire MedDRA – ou d'autres terminologies médicales normalisées – mais des termes courants moins complexes. La création d'un glossaire ne suffit parfois pas, et c'est la structure de la phrase entière qui permet parfois de comprendre le sens (par exemple : « je n'ai pas fermé l'œil de la nuit »).

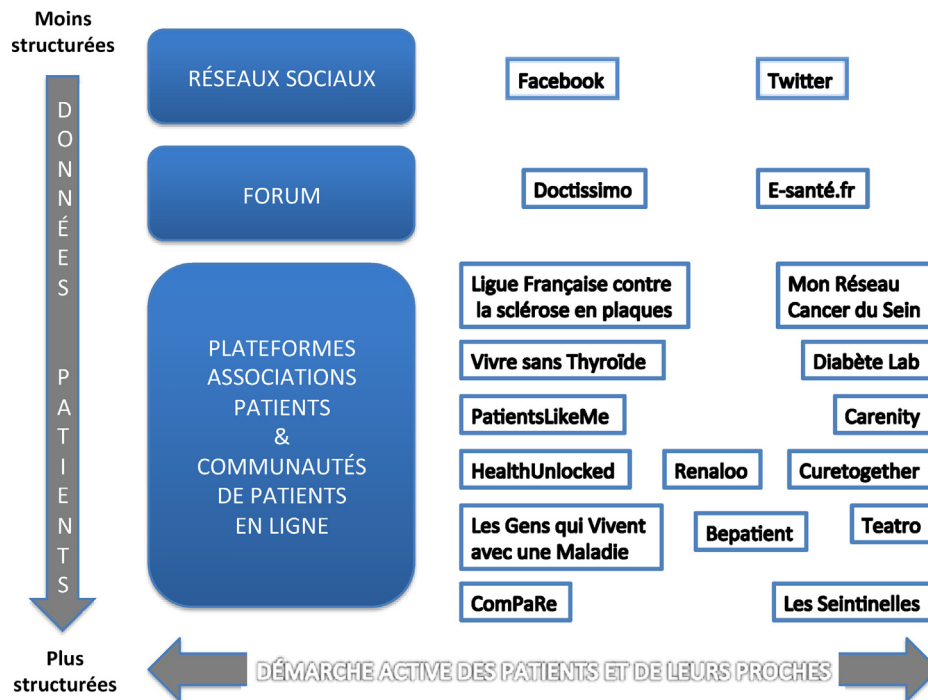


Figure 1. Paysage actuel des communautés virtuelles de patients.

Des méthodes automatisées, encore exploratoires en France, sont en cours d'expérimentation, tel le projet de plateforme pharmaco-épidémiologique d'évaluation systématisée du médicament en population « Drugs-Safe » soutenu par l'Agence nationale de sécurité des médicaments et des produits de santé (ANSM) [5] dont l'un des axes de recherche est de détecter dans les forums de discussion des mésusages du médicament, ce qui nécessite de faire une représentation de l'usage normal de celui-ci et d'offrir des outils de visualisation statistiques des usages « autres » connaissant les indications, les contre-indications, les effets indésirables connus de chaque molécule. Citons également le projet « DIPEX d'Oxford », plateforme à la fois associative et académique qui a ainsi généré 50 publications grâce au soutien et à un accompagnement méthodologique universitaire [6].

Types de recherches possibles

Dans le domaine de la recherche intéressant les produits de santé et leur environnement, les données ainsi collectées et retraitées peuvent avoir potentiellement un intérêt à toutes les étapes du continuum de la recherche, dès les phases de génération d'hypothèses et de preuve de concept jusqu'aux questions auxquelles le patient est confronté face à sa pathologie et aux traitements qui lui sont prescrits, et enfin à l'analyse en situation réelle d'utilisation (Fig. 2) :

- l'identification des besoins non satisfaits ;
- le « design » et la réalisation des essais ;
- la sélection et la construction des outils d'évaluation ;
- la compréhension des usages des produits de santé des patients en temps réel ;
- les études cliniques centrées sur les données du patient ;

- le suivi des patients en vie réelle ;
- l'éducation et le soutien aux patients ;
- ...

L'analyse de la littérature et les témoignages recueillis confirment dans tous les cas que ces supports présentent des atouts, pour répondre à certaines questions posées dans le domaine de l'information autour des produits de santé, de l'approche sociétale ou de l'organisation des soins mais sont également sources de contraintes à leur mise en œuvre, à leur analyse et à l'exploitation de leurs résultats (d'ordre juridique, méthodologique, etc.).

Intérêt et principaux apports des communautés de patients en ligne et en réseau pour la recherche

Parmi les domaines de recherche possible, la table ronde a choisi de s'intéresser plus particulièrement à deux activités phares : la recherche clinique et les études réalisées en conditions réelles dites aussi « de vraie vie », dans le contexte de la recherche sur les produits de santé.

Les patients sur Internet sont également une formidable source de données permettant de passer à une autre échelle dans les domaines des vigilances : « cyber pharmacovigilance » (ou e-pharmacovigilance), e-addictovigilance, par l'étude de blogs, forums ou communautés en ligne, mais également en « pharmacologie sociale » de par le rôle socioculturel de ces communautés dans l'éducation thérapeutique des patients.

On recense à ce jour 80 publications de travaux issus des données colligées au sein des communautés tels Patients-LikeMe [7] sans compter les autres dont la communauté

Activités de recherche possibles

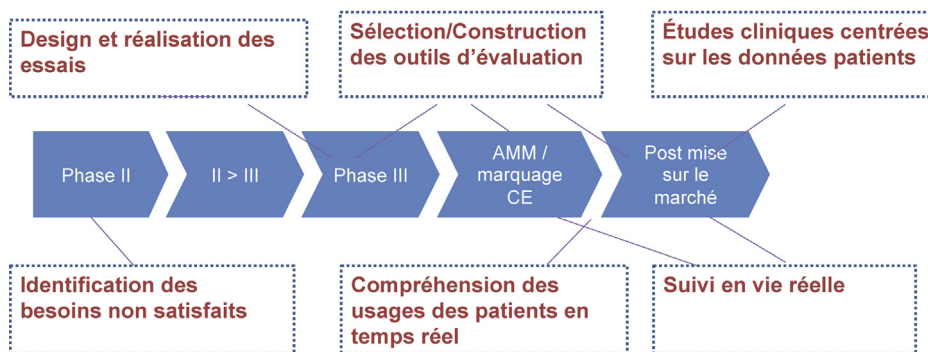


Figure 2. Communauté virtuelle de patients : champs de recherches possibles.

Carency constituée depuis 5 ans en France, dont les publications sous forme de « posters » ou communications diverses se multiplient dans le cadre de congrès scientifiques.

Apport en recherche clinique et dans les études en conditions réelles

Les apports spécifiques suivants ont ainsi été identifiés sans que cette liste soit exhaustive.

Avantages des communautés de patients en ligne en matière de recherche clinique

Voici les avantages en matière de recherche clinique :

- l'exploration des besoins non couverts et des attentes des patients. L'expression des patients sur leurs attentes et les espoirs thérapeutiques est une source extrêmement riche permettant d'appréhender au plus près les besoins médicaux non ou mal satisfaits. Les exemples dans la littérature sont nombreux [8,9]. L'identification des besoins non satisfaits sur le plan thérapeutique pourrait bénéficier de ces méthodes, en étant au plus près également des communautés de patients et de leurs aidants, susceptibles d'exprimer les observations remontées par les patients (*patient reported outcome* dans la littérature anglo-saxonne) ;
- l'apport à la conception des protocoles et aux outils associés à la recherche. La littérature abonde de publications sur la construction, l'amélioration ou l'adaptation d'outils existants utilisés en soins courants mais plus rarement en recherche [10,11] grâce à l'utilisation de plateformes structurées comme PatientsLikeMe ;
- la possibilité de conforter les protocoles et les *endpoints* en relation avec la problématique centrée sur le patient (notion de *patient centricity*) ;
- les améliorations et l'aide à la faisabilité d'une étude ainsi que l'aide au recrutement et à l'observance au protocole. Une écoute active des blogs de patients ou une sollicitation par l'intermédiaire des médias sociaux et des communautés numériques de patients pourraient permettre une meilleure adéquation entre la faisabilité des études et la réalité de leur acceptation. Une étude réalisée aux États-Unis a montré que peu d'industriels des produits de santé recherchent actuellement un retour

d'expérience ou explorent les aspects de faisabilité et recevabilité des protocoles auprès des participants par ces moyens, alors que certaines plateformes proposent des solutions en *crowdsourcing* [12,13] ;

- la réalisation d'études cliniques dans des pathologies rares d'évolution rapide [14] ;
- le choix du territoire pour la conduite de l'essai.

Six grands domaines d'apport pour les études en conditions réelles

De par la nature même des données spontanément fournies par les patients, il est facile de comprendre que les travaux sur la compréhension des parcours et des usages des patients en temps réel trouvent un terrain d'exploration privilégié pour les chercheurs avec ces communautés de patients. L'expression libre du patient permet également d'aborder des thèmes sociologiques, peu investigués par les moyens classiquement utilisés en pharmaco-épidémiologie.

Les 6 domaines d'apport identifiés sont :

- l'amélioration de la compréhension du parcours patient [15] ;
- l'analyse du ressenti patient sur sa maladie (y compris les pathologies médicamenteuses) et sur la qualité de vie [16,17] ;
- les données épidémiologiques [18] ;
- l'évaluation des bénéfices des approches thérapeutiques (efficacité ressentie) [19] ;
- l'observance au traitement, la fréquence et la caractérisation des effets indésirables, l'émergence d'effets ressentis inhabituels, les motifs de passage d'un médicament à un autre [20,21] ;
- l'efficacité de programme d'éducation thérapeutique [22].

Pour ces deux domaines de recherche (recherche clinique et études en conditions réelles) il est possible de créer une communauté en ligne spécifique et dédiée pour l'étude.

Patients sur Internet : source de données

Les patients utilisent maintenant l'ensemble des possibilités d'Internet (forums patients, sites, bases, Facebook, Twitter...) pour chercher des informations sur le médicament, et notamment sur l'effectivité et/ou les effets

indésirables. Les consultations des pages Internet et les échanges entre internautes sur le médicament sont potentiellement des sources nouvelles d'information pour l'étude de l'usage des médicaments.

Plusieurs groupes de chercheurs ont déjà mis en évidence l'importance de ces ressources pour évaluer autrement le médicament en situation réelle d'utilisation. Complément aux données traditionnellement collectées en pharmacovigilance, l'analyse des conversations et des messages véhiculés dans les fils de discussions sont autant de sources pour évaluer le bon usage et le mésusage du médicament, les détournements d'indications, et rechercher de possibles signaux d'alerte.

Ainsi dans le domaine de la pharmacovigilance et de l'addictovigilance il est possible de détecter et de confirmer un signal, de caractériser les patients, les « événements » indésirables ou les effets recherchés, les modalités d'usage lors de détournement d'indications, mais aussi d'appréhender la gestion des « événements » indésirables par les patients, l'évolution de leur perception du risque, de comparer les « événements » indésirables notifiés par les patients avec celles des professionnels de santé [23–27]. Un travail récent a également caractérisé les aspects délétères connus des forums où s'expriment des « conseils » inappropriés, en particulier en termes d'auto-prescription et la nécessité d'alerter les patients sur le possible manque de fiabilité de ces sources en termes de recommandations médicales [28].

Néanmoins, le volume et la vélocité des données générées à partir des réseaux sociaux fournissent potentiellement des opportunités pour faire avancer la pharmacovigilance. Cependant, un certain nombre de défis doivent être résolus, notamment techniques, réglementaires et éthiques. En effet, la question critique à laquelle une réponse doit être trouvée est quelle est la valeur ajoutée des réseaux sociaux par rapport aux processus actuels de pharmacovigilance et que doit-on mettre en œuvre pour apprécier cette valeur.

Science de l'étude des interactions entre le médicament et la société, la pharmacologie sociale ou sociétale s'intéresse à l'influence des facteurs sur l'utilisation des substances pharmacologiques, indépendamment des raisons purement cliniques ou rationnelles. Les analyses qualitatives du discours des patients, sur le vécu, les comportements d'usage, le bien-être, le mode de vie des patients offrent des possibilités intéressantes dans les thématiques de pharmacologie sociale et peuvent remplacer des méthodes d'entretiens dirigés ou semi-dirigés contraignants, longs à développer et à administrer et coûteux [29].

D'autres thèmes de recherche en pharmacologie sociale, comme l'étude des modifications des schémas éducatifs et l'intérêt de valider la problématique d'un rôle éducatif de ces plateformes de patients sont d'autres champs de travail possibles [30].

Intérêts des communautés et réseaux en ligne de patients

Au-delà de la typologie des communautés de patients et réseaux en ligne et de leur composante académique, associative ou commerciale, la démarche suivante a consisté à

l'identification de leurs atouts principaux, regroupés selon deux attributs : « l'approche patients » et les « intérêts méthodologiques » pour la recherche.

Intérêts liés à « l'approche patients »

C'est parce que le patient est au cœur de ces communautés quel que soit l'outil et sa typologie et que l'approche patients est par essence privilégiée, que ces outils permettent d'acquérir une connaissance précieuse sur les comportements, le verbatim, la représentation, la proximité, les usages, des éléments sur le vécu et le bien-être des patients difficilement observables par d'autres méthodes. En ce sens, ces supports sont accessibles 24h/24 par les patients, donc très souples et non-opérateurs dépendants. Les données sont le reflet très exact de leurs préoccupations quotidiennes qui ne sont pas nécessairement ou exclusivement de nature médicale mais peuvent être également de nature domestique, professionnelle, relationnelle, voire intime. Dans certains sites ou communautés, l'anonymat, parfois relatif, ou l'utilisation d'un pseudonyme encouragent les patients à livrer leurs interrogations entre pairs avec une spontanéité et une sincérité d'expression dont ils ne font pas forcément preuve auprès de leur médecin traitant ou des soignants.

L'approche « patients » a le mérite de permettre l'information et l'implication des patients lorsqu'il est sollicité pour une recherche (dans le cas des communautés associatives, le patient a le sentiment de se prendre en main et d'être acteur de sa maladie et de son traitement). Cette implication se traduit par une mobilisation accrue, la motivation du sujet étudié étant souvent un facteur clef de succès dans la conduite d'une recherche. Acteur et au contact de la recherche, adossé à un savoir basé sur l'expérience, le patient acquiert des connaissances et peut devenir un « expert profane ».

La complémentarité avec les données existantes classiques recueillies, l'étendue et la variété des champs offerts : pathologie, perception du traitement, vécu, qualité de vie... sont autant d'opportunités pour enrichir la connaissance sur les produits de santé, en particulier après leur commercialisation.

Intérêts méthodologiques

Les intérêts méthodologiques sont essentiellement de deux natures : ceux liés à la population joignable et ceux liés au traitement de l'information.

La taille de la population jointe et le volume potentiel de données disponibles sont très supérieurs à ce qui est habituellement possible dans un cadre d'étude classique (450 000 patients pour PatientslikeMe, 250 000 pour Carenity, 3 431 170 192 posts traités par Treato). Le corollaire est une forme de représentativité géographique et sociodémographique plus importante, même si cette représentativité pose certaines limites qui doivent être prises en compte au moment de l'énoncé des résultats.

Les communautés virtuelles de patients centrées sur une pathologie ou un nombre restreint de pathologies et la structuration de certaines communautés permettent un accès précieux à une population ciblée par pathologie et/ou à des patients qui ne sont pas dans les canaux de recherche et de soins traditionnels.

La rapidité de l'obtention de l'information est un avantage incontestable de ces supports pour le traitement de l'information et la génération des données. Les possibilités de modifier la requête, d'amender le protocole de façon itérative, de suivi longitudinal des paramètres (chaînage), offrent une flexibilité difficilement égalable par les études conventionnelles.

Limites et recommandations

Limites : l'attractivité de ces nouveaux outils et sources de données est cependant aujourd'hui largement freinée par des limites méthodologiques, d'une part, et des freins réglementaires et éthiques au sens large, d'autre part. C'est à partir de ce constat en deux volets que sont ensuite formulées les recommandations et préconisations de la table ronde.

Limites méthodologiques

Les limites méthodologiques liées à cette activité de recherche sont nombreuses.

En revers de l'avantage mentionné plus haut de spontanéité et sincérité des données livrées par les patients anonymement, sous un « pseudo » ou alias la plupart du temps, la première limite tient au caractère déclaratif des données. À l'extrême, les déclarations peuvent être totalement inventées (*On the Internet, nobody knows you are a dog – the New Yorker 1993*), ce qui dans un environnement très sensible à la qualité, au contrôle des données, à la vérification des sources, peut apparaître comme rédhibitoire. Il serait donc utile de connaître la réalité de ces « faux » patients dans les différents supports, le problème n'étant probablement pas le même en fonction du nombre de patients étudiés si on s'adresse à des communautés virtuelles de patients structurées et administrées ou si on s'adresse à des blogs et forums ou des réseaux sociaux.

Une autre limite du caractère déclaratif des données, est l'absence de validation médicale, ainsi il est difficile par exemple de caractériser un « événement » indésirable en effet indésirable. Néanmoins, certaines données subjectives recueillies actuellement dans le cadre des études cliniques conventionnelles sont exposées aux mêmes difficultés.

Les difficultés d'interprétation et de classification des « verbatim » et parfois l'obstacle de la langue avec un outil multilingue, déjà évoqués plus haut, représentent un enjeu majeur et de taille pour l'utilisation de ces données dans le cadre des travaux de recherche.

En dépit de la large taille des panels de population des communautés, les biais de représentativité sont potentiellement multiples : le niveau d'appétence pour l'informatique, le web et les réseaux sociaux va varier souvent fortement en fonction de l'âge, de caractéristiques socioculturelles et professionnelles, du style de vie, d'habitat urbains/ruraux etc. Une connaissance plus fine des populations des communautés étudiées est nécessaire afin de pouvoir corriger éventuellement ces biais.

L'influence de l'environnement et en particulier des médias [31], le niveau de notoriété d'une question dans une communauté, ne doivent pas être négligés, d'une part, pour prendre en compte ce facteur dans une analyse s'intéressant

à l'incidence d'un événement, d'autre part, pour identifier les mots clés pertinents, et notamment détecter, comprendre et traiter l'amplification d'un signal.

Dans le cas des forums, le parasitage par d'autres sujets de discussion que le sujet d'intérêt initial conduit à un rétrécissement souvent drastique des données exploitables, une fois traités les milliers de fils de discussion au bout du compte vides ou hors sujet.

Une des autres principales limites méthodologiques de l'utilisation des communautés de patients en ligne tient enfin à la très grande hétérogénéité de la structuration des supports qui est liée aux objectifs différents poursuivis : sites, forums, plateformes commerciales, communautés associatives dédiées et tournées vers le patient à titre principal ou accessoire, ou indifférent.

Un des freins majeurs répertorié est également l'hétérogénéité et l'absence d'interopérabilité des lexiques sémantiques et glossaires utilisés pour les terminologies médicales (*unified medical language system [UMLS]*, *systematic nomenclature medical clinical terms [SNOMED]*, serveur multi-terminologies en santé [SMTS]).

Enfin la faible visibilité des communautés et réseaux par les chercheurs, notamment les communautés dédiées patients très fermées et la problématique de l'accès payant de certaines plateformes commerciales parmi les plus structurées est à regretter du point de vue de la recherche.

Aspects éthiques et réglementaires

Aux limites méthodologiques s'ajoutent des questions éthiques et déontologiques ainsi qu'une quasi-absence ou à l'inverse une complexité de réglementation sur certains aspects, liés certainement au caractère novateur du sujet.

La sécurité, l'anonymisation et la confidentialité des données fournies par les patients sont-elles toujours garanties ? Comment garantir que le patient est conscient et consent à l'utilisation de ses données et qu'il y a une transparence de leur exploitation ?

La problématique de l'accessibilité des données et de la propriété des données, qui rappelons-le sont dans le cas présent des données personnelles publiques, confère un niveau d'incertitude pour les chercheurs. Cette problématique est proche de celle rencontrée pour les *big data*, *open data*, entrepôts de données et à l'utilisation des bases de données hospitalières à des fins de recherche mais cette incertitude est renforcée en raison du caractère spontané voire candide et naïf, des déclarations, verbatim et informations apportés par les patients. La question peut se résumer en : « à qui appartiennent les données des patients ? », ou plus précisément dans ce cas « à qui appartiennent les informations que les patients postent sur Internet ? ».

Des outils puissants existent déjà et les technologies progressent très rapidement. Ainsi les robots informatiques permettent d'« aspirer » des forums et blogs entiers ou des centaines de fils de discussion pour les étudier. Il est donc urgent de clarifier les questions éthiques que se posent les chercheurs. Très récemment, deux projets français, utilisant ces technologies, ont fait l'objet d'annonces. Financé par les pouvoirs publics, il s'agit pour le premier (projet Vigi4Med) de créer les moyens de retrouver, filtrer et analyser les commentaires des patients sur Internet pour certains médicaments, et de réaliser une évaluation rétrospective et

prospective des informations [32]. Pourront également être mesurés les bénéfices attendus et les coûts associés à une recherche proactive des signaux sur Internet. Le deuxième, le projet ADR-PRISM, conduit par un consortium public privé et financé par un fonds interministériel unique poursuit un objectif similaire dans le domaine de la pharmacovigilance [33].

In fine, s'agissant du champ de la recherche clinique ou des études en conditions réelles d'utilisation, se pose la question de la reconnaissance et la recevabilité des données par les autorités compétentes dès lors que les conditions de l'authentification des données et de la remontée aux dossiers médicaux sources des patients ne sont pas réalisables, l'anonymat étant en général la règle de ces sites.

En regard de ces principales limites méthodologiques, éthiques et réglementaires identifiées, la table ronde s'est attachée à proposer un certain nombre de recommandations.

Recommandations et préconisations

Les recommandations et préconisations envisagées par les membres de la table ronde sont, en ce qui concerne les recommandations méthodologiques, surtout destinées à l'intention des chercheurs et professionnels du secteur, et pour les recommandations réglementaires, adressées au premier chef, en direction des autorités sanitaires et des décideurs publics.

Recommandations méthodologiques

En premier lieu, il s'agit de lister et de décrire les ressources disponibles. Cette description devrait pouvoir permettre de qualifier les différentes ressources en fonction de leurs thématiques et objectifs affichés, mais également dans un objectif de recherche, de connaître le nombre d'adhérents, de visiteurs, de communautés, les caractéristiques des données disponibles et les possibilités d'accès aux chercheurs externes, les structures des bases et les langages informatiques utilisées, les expertises internes en statistiques, informatique. . .

Les conditions nécessaires à la réalisation d'un tel effort ne peuvent être réunies que grâce à un support des pouvoirs publics et/ou par la mise en commun de ressources de plusieurs parties prenantes, qu'elles soient d'origine publique ou privée. En raison de l'évolutivité continue des sites Internet qui sont ouverts aux patients et aux aidants, des mises à jour régulières de ce répertoire public devraient être envisagées très en amont.

La table ronde recommande la mise en place d'un référencement des cohortes issues des communautés et réseaux en ligne (« web cohortes »), dans le « Portail Épidémiologie France » au même titre que des cohortes constituées de façon plus classique par des chercheurs, afin de leur donner une meilleure visibilité et de renforcer leur intérêt.

Pour continuer à progresser dans l'utilisation de ces outils et répondre à la difficulté identifiée d'interprétation et de classification des verbatim et d'obstacle des langues, un travail indispensable de création de glossaires/lexiques en « langage patient » et d'alignement de chacun de ces termes vers les concepts d'une terminologie de référence devra

être menée et partagée par les professionnels et chercheurs. Il serait souhaitable que cette terminologie de référence soit intégrée dans un serveur multi-terminologique, tels que l'UMLS (avec les limites connues vis-à-vis du français) ou en français *united medical lexicon for French* (UMLF) [34] ou *French multi-terminology indexer* (FMTI) [35].

Sur le plan méthodologique et pour traiter à la fois l'hétérogénéité des approches et des supports, les professionnels et chercheurs devront enfin élaborer des « bonnes pratiques » harmonisées de recueil et des méthodes d'analyses des données issues des communautés et réseaux de patients en ligne (approches quantitatives et qualitatives). Celles-ci pourraient être formalisées en un guide méthodologique ou « cahier des charges ». À ce titre il sera très intéressant de suivre les progrès du projet « Vigi4Med » qui a l'ambition de produire des recommandations concernant la meilleure façon d'utiliser à des fins de recherche, les commentaires des patients dans les forums.

Recommandations et préconisations sur le plan réglementaire et éthique/déontologique

Le respect des règles de droit des patients se prêtant à des recherches, doit être garanti. Là aussi la rédaction d'un guide des bonnes pratiques de l'utilisation pour des activités de recherche des données personnelles de e-santé, destiné aux communautés virtuelles de patients, permettrait de compléter le cadre réglementaire actuel : information sur la structure porteuse de la communauté, usage des données de santé collectées, recommandations aux usagers pour éviter les risques inhérents à la communication des données personnelles sur des espaces publics, à l'accès par des tiers, périmètre et temporalité du consentement, incitation au partage de la démarche de la recherche et accès aux résultats obtenus [36].

Plus spécifiquement, en matière de recherche biomédicale, une réflexion conduite avec la conférence nationale des CPP permettrait également de développer les conditions de réalisation de ces études, et de partager avec cet acteur la connaissance des différents types de recherches conduites au sein des communautés virtuelles de patients.

En matière de pharmacovigilance, la notification d'un effet indésirable aux autorités de santé, requiert classiquement de connaître une source identifiable (le notificateur), un patient identifiable, le nom du produit suspecté, la nature de l'effet indésirable. Dès lors que des équipes de chercheurs de tous bords vont mettre en œuvre des moyens importants de *data mining* de plateformes comportant plusieurs centaines/milliers de patients, de défilement de milliers de conversations dans les blogs ou forum, la probabilité de découverte d'informations de type effets indésirables susceptibles de grossir le corpus de connaissances de pharmacovigilance sur un produit de santé, voire de générer de potentiels signaux d'alerte devrait augmenter sensiblement. Il est souhaitable qu'une réflexion soit entamée avec l'ensemble des parties prenantes (ANSM, les centres de pharmacovigilance, les associations de patients, les chercheurs et pharmacovigilants industriels) sur la conduite à tenir face à ces situations.

L'établissement des conditions de recevabilité des PRO/PRI dans les études à visée réglementaire, par les

autorités de santé françaises et européennes constitueront également une étape indispensable, afin de promouvoir la réalisation d'études de recherche biomédicale au sein de la communauté scientifique et médicale. À ce jour, il n'a été retrouvé qu'un seul témoignage de réflexion engagée par un panel constitué d'un représentant de la Food and Drug Administration (FDA), un représentant de communauté virtuelle de patients et des industriels [37].

Conclusion

La révolution digitale et son extrême mouvance lancent un défi sans précédent à la communauté des chercheurs dans le domaine des produits de santé, en démultipliant la quantité d'informations disponibles. Ces informations, ces données patients, peuvent utilement soutenir des travaux de recherche clinique, épidémiologique, pharmaco-épidémiologiques et pharmaco-sociétaux. Parmi les nombreux modèles de plateformes numériques existantes, certains se sont davantage dotés de moyens pour réaliser ces recherches, produisant des résultats de qualité suffisante pour permettre des publications dans des revues scientifiques de renom. La démarche active, l'appartenance à une communauté, la constitution de savoirs au contact de ses pairs, remettent au centre du dispositif de la recherche le patient qui doit pouvoir être assuré de la qualité et de la pertinence des recherches utilisant ses données personnelles. La modernisation de la recherche dans les produits de santé est un enjeu critique tant il est aujourd'hui difficile, complexe et coûteux de continuer d'améliorer le progrès thérapeutique. Les quelques pistes dégagées par la table ronde sont autant d'incitations à poursuivre la réflexion pour que ces recherches se développent de façon éthique et responsable et pour favoriser l'initiative tant des associations de patients et des académiques que des acteurs économiques indispensables à la création de valeur et à l'innovation.

Déclaration de liens d'intérêts

M. Michael Chekroun, co-auteur de l'article, est président et fondateur de la société Carentity.

Références

- [1] Chatelier G, Varlet V, Blachier-Poisson C, les participants à la table ronde n° 6 de Giens XXXI. « Big data » et « open data » : Quel accès pour la recherche ? *Therapie* 2016;71:97–105.
- [2] Institut Curie. Décrypter les perceptions des français sur le cancer; 2015 http://www.unicancer.fr/sites/default/files/DP_Observatoire_Cancer_Institut_Curie_Viavoice_2015.pdf [consulté le 29 novembre 2016 (11 pp.)].
- [3] IPSOS Observer. Usages et attentes des français à l'égard du digital en matière d'information sur leur santé; 2016 http://www.ipsos.fr/sites/default/files/doc_associe/rapport_ipsos_pour_msd_france_-_digital_et_sante.pdf [consulté le 29 novembre 2016 (31 pp.)].
- [4] Baldwin M, Spong A, Doward L, Gnanasakthy A. Patient-reported outcomes, patient-reported information, from randomized controlled trials to the social web and beyond. *Patient* 2011;4(1):11–7.
- [5] ANSM. L'ANSM soutient deux plateformes en épidémiologie des produits de santé – Point d'information; 2014 <http://ansm.sante.fr/S-informer/Points-d-information-Points-d-information/L-ANSM-soutient-deux-plateformes-en-epidemiologie-des-produits-de-sante-Point-d-information> [consulté le 29 octobre 2016].
- [6] <http://www.healthtalk.org/> [consulté le 29 novembre 2016].
- [7] PatientsLikeMe. A poster bibliography of PatientsLikeMe; 2016 <https://patientslikeme.posters.s3.amazonaws.com/PatientsLikeMe%20Posterbook.pdf> [consulté le 29 novembre 2016].
- [8] Meunier B, Jourde-Chiche N, Mancini J, Chekroun M, Retornaz F, Chiche L. Characteristics and information searched for by French patients with systemic lupus erythematosus: a web-community data-driven online survey. *Lupus* 2016;25(4):370–5.
- [9] Simacek K, Raja P, Chiauzzi E, Eek D, Halling K. What do ovarian cancer patients expect from treatment?: perspectives from an online patient community. *Cancer Nurs* 2016 <http://www.ncbi.nlm.nih.gov/pubmed/27454765> [(Epub ahead of print) consulté le 29 novembre 2016].
- [10] Kear T, Harrington M, Bhattacharya A. Partnering with patients using social media to develop a hypertension management instrument. *J Am Soc Hypertens* 2015;9(9):725–34.
- [11] Tran VT, Harrington M, Montori VM, Barnes C, Wicks P, Ravaud P. Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform. *BMC Med* 2014;12:109.
- [12] Industry usage of social and digital media communities in clinical research. A Tufts center for the study of drug development white paper; 2014 http://csdd.tufts.edu/files/uploads/TCSDD_Social_Media_Final.pdf [consulté le 29 novembre 2016 (19 pp.)].
- [13] Okun S, Wicks P, Scott M, Gilbert J. Patient-informed clinical trials: cross-sectional survey on a patient powered research network, PatientsLikeMe, or information association; 2014 http://news.patientslikeme.com/sites/patientslikeme.newshq.businesswire.com/files/doc_library/file/PostersLikeMe_Sept_2014.pdf [consulté le 18 octobre 2016 (20 pp.)].
- [14] A study of NP001 in subjects with amyotrophic lateral sclerosis. <https://clinicaltrials.gov/ct2/show/NCT01281631> [consulté le 29 novembre 2016].
- [15] Chiauzzi E, DasMahapatra P, Cochin E, Bunce M, Khoury R, Dave P. Factors in patient empowerment: a survey of an online patient research network. *Patient* 2016;9(6):511–23.
- [16] McCarrier K, Bull S, Fleming S, Simacek K, Wicks P, Cella D, et al. Concept elicitation within patient-powered research networks: a feasibility study in chronic lymphocytic leukemia. *Value Health* 2016;19(1):42–52.
- [17] De la Loge C, Dimova S, Mueller K, Phillips G, Durgin T, Wicks P, et al. PatientsLikeMe® online epilepsy community: patient characteristics and predictors of poor health-related quality of life. *Epilepsy Behav* 2016;63:20–8.
- [18] Blin R. Peut-on déterminer la prévalence d'une maladie en fonction de la fréquence de son nom sur le web ? Le cas du cancer au Japon; 2016 <https://hal.archives-ouvertes.fr/hal-01214311v2> [consulté le 29 novembre 2016].
- [19] Drahos J, Cerrato D, Sherman J, Wang V, Bradley M, Mack S, et al. Patient insights and voice on major depressive disorder treatment efficacy and symptom perception (PIVOT) study: initial findings from the PatientsLikeMe online community. In: NNDC 2015 annual conference. National Network of Depression Centers (NNDC); 2015.
- [20] Wicks P, Brandes D, Park J, Liakhovitski D, Koudinova T, Sasane R. Preferred features of oral treatments and predictors of non-adherence: two web-based choice experiments in multiple sclerosis patients. *Interact J Med Res* 2015;4(1):e6.
- [21] Blaser D, Wang V, Wicks P. A novel tool for monitoring patient experiences with nintedanib and pirfenidone in idiopathic

- pulmonary fibrosis: preliminary data and insights. In: 111th Int. Conf. of the American Thoracic Society (ATS). ATS; 2015.
- [22] Attai D, Cowher M, Al-Hamadani M, Schoger J, Staley A, Landercasper J. Twitter social media is an effective tool for breast cancer patient education and support: patient-reported outcomes by survey. *J Med Internet Res* 2015;17(7):e188.
- [23] Powel G, Seifert H, Reblin T, Burstein P, Blowers J, Menius JA, et al. Social media listening for routine post-marketing safety surveillance. *Drug Saf* 2016;39(5):443–54.
- [24] Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. *Drug Saf* 2014;37:343–50.
- [25] Bagheri H, Lacroix I, Guitton E, Damase-Michel C, Montastruc JL. Cyber pharmacovigilance: what is the usefulness of the social networks in pharmacovigilance? *Therapie* 2016;71(2):235–9.
- [26] Schifano F, Deluca P, Baldacchino A, Peltoniemi T, Scherbaum N, Torrens M, et al. Drugs on the web; the Psychonaut 2002 EU project. *Prog Neuropsychopharmacol Biol Psychiatry* 2006;30(4):640–6.
- [27] Kheloufi F, Default A, Blin O, Micallef J. Investigating patient narratives posted on Internet and their informativeness level for pharmacovigilance purpose: the example of comments about statins. *Therapie* 2016, <http://dx.doi.org/10.1016/j.therap.2016.10.004> [Epub ahead of print].
- [28] Palosse-Cantaloube L, Lacroix I, Rousseau V, Bagheri H, Montastruc JL, Damase-Michel C. Analysis of chats on French internet forums about drugs and pregnancy. *Pharmacoepidemiol Drug Saf* 2014;23(12):1330–3.
- [29] Moncrieff J, Cohen D, Mason JP. The subjective experience of taking antipsychotic medication: a content analysis of Internet data. *Acta Psychiatr Scand* 2009;120:102–11.
- [30] Goulinet G. Rôle socio-culturel des communautés virtuelles de patients dans le suivi des maladies chroniques. Vers un nouveau modèle d'éducation thérapeutique ? Observatoire des mondes numériques en sciences humaines; 2013 http://www.omnsh.org/sites/default/files/geraldine_goulinet-vf.pdf [consulté le 29 novembre 2016 (121 pp.)].
- [31] Abou Taam M, Rossard C, Cantaloube L, Bouscaren N, Roche G, Pochard L, et al. Analysis of patients' narratives posted on social media websites on benfluorex's (Mediator[®]) withdrawal in France. *J Clin Pharm Ther* 2014;39(1):53–5.
- [32] Vigi4MED. Vigilance dans les forums sur les médicaments. <https://iscod.emse.fr/content/vigi4med-vigilance-dans-les-forums-sur-le-me%CC%81dicament> [consulté le 29 novembre 2016].
- [33] Projet ADR-Prism; 2016 http://www.acadpharm.org/dos_public/SCHUK_presentation_ADR-PRISM_ANP_V0.1.pdf [consulté le 29 novembre 2016].
- [34] Zweigenbaum P, Baud R, Burgun A, Namer F, Jarrousse E, Grabar N, et al. UMLF: a unified medical lexicon for French. *Int J Med Inform* 2005;74(2–4):119–24.
- [35] Pereira S, Névéal A, Kerdelhué G, Serrot E, Joubert M, Darmoni SJ. Using multi-terminology indexing for the assignment of MeSH descriptors to health resources in a French online catalogue. *AMIA Annu Symp Proc* 2008:586–90.
- [36] Pour un patient acteur de la qualité de son parcours de santé. Le numérique en santé Collectif inter associatif sur la santé; 2016 http://www.leciss.org/sites/default/files/160912_Note-Ciss-E-sante-Bon-Usage.pdf [consulté le 29 novembre 2016 (44 pp.)].
- [37] Rothman M, Gnanaskathy A, Wicks P, Papadopoulos E. Can we use social media to support validity of patient-reported outcome instruments in medical product development? *Value Health* 2015;18(1):1–4.