UNICANCER

ConSoRe

Continuum Soins Recherche

Été 2016

N° 1

édito

L'avènement des nouvelles technologies impacte de plus en plus notre quotidien. Le monde de la santé ne fait pas exception à ce processus. La capacité à analyser des méga-données représente aujourd'hui un enjeu stratégique qui va permettre aux CLLC de confirmer leur rôle de leader en innovation en cancérologie.

Conscients de cette évolution, nous avons lancé il y a 3 ans le programme ConSoRe (Continuum Soin-Recherche). Son objectif initial était de pouvoir exploiter au mieux les données stockées dans nos établissements depuis plus de 10 ans et favoriser la recherche de transfert au sein de nos Centres et entre nos Centres.

Le moteur de recherche développé par les équipes des 4 Centres pilotes (le Centres Georges-François Leclerc de Dijon, le Centre Léon Bérard de Lyon, l'Institut du Cancer de Montpellier et l'Institut Curie à Paris) est devenu un véritable outil de *Data Mining* permettant de trouver facilement des données textuelles ou structurées dans l'ensemble du système d'information hospitalier. L'évolution prévisible de cet outil favorisera, j'en suis sûr, à la fois une prise en charge personnalisée de nos patients et la recherche de transfert.

Les travaux les plus récents des équipes des 4 Centres pilotes, nous permettent d'envisager la structuration et l'analyse automatisée de nos données issues du soin et de la recherche.

Ce programme est l'illustration de la dynamique qui existe au sein de nos établissements et qui représente la force de notre Groupe. ConSoRe constitue la première brique d'une orientation stratégique décidée mutuellement. Le déploiement de ConSoRe au cours du second semestre de cette année, fournira aux CLCC l'outil leur permettant d'entrer dans le monde du *Big Data* en Santé.

Pr Patrice Viens,
Président d'UNICANCER.

Actualités ConSoRe

Conçu comme un outil de fouille de données multi-sources et multi-formats pour la constitution de cohortes, ConSoRe est devenu un véritable outil de Data Mining permettant de trouver facilement des données textuelles ou structurées dans l'ensemble du système d'information hospitalier. ConSoRe va permettre l'agrégation de données venant de différentes sources dans l'objectif de modéliser l'histoire néoplasique du patient.

Cette solution robuste et évolutive devra aider les médecins et les chercheurs à identifier des pathologies ou des situations particulières en vue d'améliorer leurs prises en charge ultérieures ou de valider des hypothèses de recherche.

Après plusieurs mois de développements complémentaires afin d'en faciliter l'usage, nous sommes en passe d'installer ConSoRe dans les 4 Centres pilotes que sont l'Institut Curie, l'ICM, le CGFL et le CLB. Cette installation débutera d'ici la fin juin. A l'issue de celle-ci et d'une période de test qui durera jusqu'à l'automne, nous pourrons progressivement installer ConSoRe dans les autres Centres.

Dans les Centres où il sera installé, il offrira la possibilité d'interroger simultanément plusieurs centaines de milliers de dossiers de patients, soit plusieurs millions de documents.

ConSoRe en chiffes

4 Centres pilotes

	Dossiers patients	*Documents*
Institut Curie, Paris- Saint-Cloud	430 000	10 millions
Institut du Cancer de Montpellier	300 000	3 millions
Centre Georges-François Leclerc, Dijon	120 000	1,8 million
Centre Léon Bérard, Lyon	446 000	9,4 millions

Automne 2016 : début d'installation dans 3 autres Centres

Directeur de la publication :
Patrice Viens, **Président UNICANCER**

Ont participé à ce numéro :
Dr Alain Livartoski alain.livartowski@curie.fr
et **Julien Guerin** julien.guerin@curie.fr , Institut Curie

Coordination :
Dr Pierre Heudel, Département d'oncologie médicale Centre Léon Bérard, Lyon pierreetienne.heudel@lyon.unicancer.fr
Emmanuel Reyrat, Directeur des Systèmes d'Information UNICANCER e-reyrat@unicancer.fr
Christophe Jamain, responsable recherche pré-clinique et fondamentale c-jamain@unicancer.fr
Agathe Lasne, chef de projet communication a-lasne@unicancer.fr

Témoignage

Médecin à l'Institut Curie - Paris dans le Département d'oncologie médicale, le ***Dr Alain Livartowski (à droite)*** *est responsable des cancers thoraciques. Il coordonne le Département d'information médicale, les projets eSanté comme Directeur de la clientèle et le projet Big Data au sein de l'Ensemble hospitalier.*

Julien Guerin (à gauche), *est chef de projet SI à l'Institut Curie en charge des projets d'intégration de données clinico-biologiques et Big Data. Il travaille à l'interface entre l'Hôpital et le Centre de Recherche et pilote notamment le déploiement technique des projets ConSoRe et ConSoRe+ (voir focus ci-dessous).*

Pouvez-vous nous parler des capacités de recherche de ConSoRe et des types de recherches que vous pouvez désormais faire ?

Dès la version v0, il est possible de poser des questions simples ou complexes que nous ne pouvions pas poser auparavant. Par exemple, retrouver des cas de maladies rares et lancer la même interrogation dans les autres centres qui utilisent ConSoRe. Il n'est pas nécessaire que ces données soient codées pour les retrouver car la recherche textuelle est possible. Une autre possibilité ouverte dès la version v0 est de poser des requêtes sur des sources multiples de manière transparente pour l'utilisateur : les comptes rendus médicaux, les données du PMSI, la Fiche Tumeur notamment.

Des services complémentaires sont disponibles comme la constitution de cohortes sur lesquelles il est possible de réaliser des opérations ensemblistes (les unir) ou la mémorisation des requêtes favorites qui peuvent être relancées à tout moment.

Toutefois l'interface utilisateur ne permet pas une exploitation aisée de ces possibilités et cette première version nous a montré la nécessité de modéliser la maladie, améliorer les cartouches sémantiques et ajouter des règles de raisonnement pour diminuer le bruit. En tout cas, cette version v0 a fait la preuve de concept qu'il est possible d'interroger des données structurées ou textuelles entre tous les centres avec des temps de réponse excellents.

En tant que clinicien, comment voyez-vous les impacts possibles de ConSoRe sur la prise en charge des patients et sur la recherche clinique et translationnelle ?

Pour la prise en charge clinique, il sera aisé de retrouver des cas similaires ou de nous fournir un outil pour évaluer nos pratiques cliniques par exemple. En ce qui concerne la recherche translationnelle, la version v0 ne permet pas d'interroger les données issues de la recherche mais cette possibilité est un des axes majeurs du développement que nous aurons à déployer dans les prochaines versions. De toute façon, pour une recherche translationnelle de qualité, nous devons être capables d'interroger les données cliniques et de transmettre des cohortes de patients aux médecins – chercheurs et aux chercheurs.

Un des axes de développement est de se connecter à des outils analytiques open-source dédiés à la recherche translationnelle tel que tranSMART. Une des limites de la version v0 était la difficulté à constituer des cohortes de patients incluables dans une étude. Cette limitation sera levée dès la version v1 avec l'introduction du modèle temporel de données.

Selon vous, est ce que l'intégration de ConSoRe va avoir un grand impact à l'institut Curie et si oui quels seront les domaines possiblement intéressés ?

Un des impacts majeurs sera la possibilité de structurer l'information de façon automatique et d'alimenter des bases de données rétrospectives. Il sera aussi possible d'extraire des données structurées et mises à jour automatiquement. L'impact sur la recherche translationnelle sera majeur car il permettra aux chercheurs d'accéder à des informations anonymisées. Enfin, nous pourrons constituer des projets multi-centriques sur des maladies rares ou des effets inattendus des médicaments. Tout ceci se fera de façon simple et surtout sans recourir à des codages rétrospectifs fastidieux et donc coûteux. Le projet ConSoRe est l'étape nécessaire pour les futurs projets Big Data des centres. Cela va impacter l'évaluation de la qualité des soins, la recherche clinique et translationnelle et ceci doit nous faire repenser sur l'organisation interne.

Focus sur ConSoRe+

Le projet ConSoRe+ est un projet complémentaire à ConSoRe développé à l'Institut Curie avec le soutien important du SiRIC Curie et du projet ICGEx (Institut Curie Génomique d'Excellence (https://siric.curie.fr/page/plateforme-de-sequencage-haut-debit). Il est le fruit des premières constatations réalisées à l'issue de la phase de prototype de ConSoRe. La technologie actuelle de ConSoRe, basée sur un moteur de recherche, présente des limites, par exemple pour poser n'importe quelle question pertinente. Il nous est apparu la nécessité de modéliser les données et l'histoire de la maladie et de tester d'autres technologies issues du Web sémantique pour accroître les possibilités d'interrogation. Notre rencontre avec Jean Delahousse, fondateur de la société Mondeca spécialisée dans la gestion d'ontologies et le web sémantique, a été déterminante et a ouvert la voie à l'utilisation des technologies basées sur l'exploitation sémantique des données. RDF (Resource Description Framework) et le langage OWL (Web Ontology Language) sont des technologies que nous avons retenues pour modéliser la connaissance d'un point de vue sémantique en se plaçant au plus près de l'essence littérale des comptes rendus médicaux et des données biomédicales.

Au travers d'un modèle de connaissance, créant de nombreuses inférences et des relations sémantiques fortes entre les données, nous devrions être en mesure d'interroger le système sous n'importe quel angle, que ce soit celui du patient, de la tumeur ou de l'étude d'un marqueur spécifique par exemple. Une première version de ConSoRe+ sera disponible en 2016 à l'Institut Curie. Si les résultats sont concluants, nous envisagerons les moyens de partager ces technologies avec les autres centres.

UNICANCER

ConSoRe

Continuum Soins Recherche

Hiver 2016

N°2

édito

Tel que décrit dans les orientations de l'étude EVOLPEC-2, les Centres vont devoir s'adapter à une profonde évolution au cours des prochaines années liée au développement des traitements oraux, des thérapies ciblées et à avènement de « l'hospitalisation hors les murs ».

Cette évolution de la prise en charge passe, entre autre, par une caractérisation systématique des tumeurs en vue de donner le bon produit au bon patient, au bon moment. A l'horizon 2025, on estime que près de 10 fois plus de patients auront une tumeur caractérisée par la biologie moléculaire. Ceci nécessitera la mise en place d'outils informatiques performant pour le stockage et le traitement de l'information.

Cette personnalisation de la prise en charge se fera également par le suivi des patients traités en dehors de leur Centre. Celle-ci nécessitera la mise en place de programmes de suivi individualisés, réguliers et pluridisciplinaires.

Le développement de la gestion des toxicités sera également un élément clé de la prise en charge en cancérologie. Anticiper, reconnaitre et gérer efficacement les effets secondaires est indispensable pour atteindre le bénéfice thérapeutique maximal attendu tout en maintenant la qualité de vie des patients. Ceci devra pouvoir se faire via des applications de télé-suivi afin d'assurer une prise en charge des patients même en dehors des murs de nos Centres. En effet, le contrôle de l'observance thérapeutique et des événements indésirables à domicile, le flux d'informations entre la ville et l'hôpital seront des éléments clés pour améliorer la prise en charge dans le cadre d'une médecine préventive et prédictive.

La prise en charge personnalisée et la mise en place d'une médecine à la fois préventive et prédictive sera fortement facilitée par le développement d'outils d'e-santé de suivi et de gestion de la maladie (télémédecine, télé-suivi) et d'outils tels que **ConSoRe** qui permettront d'assurer le continuum de l'information nécessaire à la mise en place de cette médecine 4P.

Pr Pierre Fumoleau,
Directeur général Centre Georges-François Leclerc, CLCC de Dijon

Actualités ConSoRe

L'installation de **ConSoRe** dans les Centres pilotes est maintenant en passe d'être terminée.

Les travaux réalisés ces dernières semaines ont consisté à finaliser les interfaces et à vérifier la cohérence des résultats présentés par **ConSoRe**. Cette période un peu ingrate de « travail de l'ombre » est importante pour optimiser ensuite l'utilisation de **ConSoRe**.

L'outil va maintenant pouvoir être utilisé dans ces Centres afin de répondre à des questions scientifiques et/ou médicales.

Sur cette même période, un travail de préparation de déploiement dans 4 nouveaux Centres a été initié. Celui-ci s'étendra sur le premier trimestre 2017.

Pendant cette période, nous commencerons à travailler avec une nouvelle vague de Centres dans lesquels **ConSoRe** pourra être mis en place avant l'été. Certains se sont déjà manifestés et nous travaillons sur une planification de cette installation.

L'objectif est d'équiper l'ensemble des Centre de **ConSoRe** d'ici la fin 2017 selon le rythme des sollicitations.

ConSoRe EN CHIFFRES

4 nouveaux Centres équipés

	Dossiers patients	*Documents*
ICL, Nancy	200 000	3,5 millions
Institut Paoli-Calmettes, Marseille	250 000	4 millions
Institut Bergonié, Bordeaux	230 000	5,2 millions
Centre Oscar Lambret, Lille	190 000	3 millions

→ **Installation au 1er trimestre 2017**

TEMOIGNAGNE de T. DURAND et C. PEZET

Membres du service informatique du Centre Léon Bérard (Lyon) impliqués dans l'installation de **ConSoRe** *(de gauche à droite* ***Thierry Durand, Amandine Arsaban, Christophe Pezet et Franck Mestre****)*

Pouvez-vous nous détailler les principales étapes de l'installation de ConSoRe dans le Centre Léon Bérard ?

Une des premières étapes a été l'achat et l'installation de 9 serveurs car la puissance et la simplification d'usage de **ConSoRe** passent par beaucoup de traitements invisibles pour les utilisateurs. De manière simultanée, un important travail de développement des interfaces entre le dossier patient informatisé (DPI) du Centre Léon Bérard et **ConSoRe** a été réalisé (les interfaces avec le centre de ressources biologiques, CRB, n'ont pas encore été réalisées).

On a pu estimer que la charge de travail humaine pour atteindre la constitution d'un premier corpus de 1000 patients s'élève à 120 jours homme. Il faudra ajouter 40 jours homme pour la finalisation des tests des interfaces, l'intégration au fil de l'eau puis l'écriture et les tests des interfaces avec le CRB.

En ce qui concerne le temps machine, la création des fichiers d'exports pour les 1000 patients tests s'élève à 5 heures. On estime à environ 100 jours l'envoi des 450 000 DPI, ce qui nécessite donc de construire une architecture de parallélisassions afin de réduire cette durée. Avec **ConSoRe**, le volume des documents est toujours important. Par exemple nous avons 6000000 documents scannés dans les DPI pour lesquels nous devons passer un traitement de reconnaissance optique des caractères (OCR) afin de récupérer du texte des documents venus de l'extérieur et les intégrer dans **ConSoRe**.

Quels sont les compétences humaines nécessaires ?

Coté systèmes d'information, une bonne connaissance du SIH du CLB est indispensable pour identifier les données à fournir. Il faut également des compétences informatiques complexes pour le travail sur les interfaces. A noter que le travail avec un médecin est nécessaire pour la création d'un premier fichier de 100 patients qui servira, lors de l'installation, à appréhender la qualification des résultats de **ConSoRe** et à vérifier que les données générées ou présentées soient correctes.

Quelles sont les difficultés rencontrées ?

En plus de difficultés liées à la charge de travail, il a fallu s'adapter à de nouvelles normes (IHE-PAM/HL7 et CDAR2 principalement) et faire un travail de déphasage entre les normes utilisées au CLB et les normes prévues dans **ConSoRe** (CIM10 vs CIM10-FR PMSI, CIMO3 topo internationale vs CIMO3 topo FNCLCC par exemple). Le suivi des difficultés rencontrées durant cette phase d'installation est assuré par le logiciel REDMINE qui permet une relation étroite et efficace avec SWORD.

FOCUS SUR… I5ADORA, *Le 5ème Élément*

Centre Léon Bérard (Dr P. Heudel et T. Durand) de Lyon, SWORD Insight Sas (JL. Vila et P. Le Calve), Unité de recherche Cancer et environnement (Pr B. Fervers)

Ce projet a pour but d'identifier les seconds cancers primitifs (SCP) et leurs facteurs de risques chez les patients suivis au Centre Léon Bérard en Rhône-Alpes. Alors que les SCP représentent près d'un patient sur cinq, les études épidémiologiques se sont, le plus souvent, limitées aux caractéristiques du premier cancer, aux traitements anti-néoplasiques reçus et au mode de vie du patient (consommation d'alcool ou de tabac, diabète…), qui sont des informations présentes dans tous les dossiers médicaux. La survenue d'un cancer étant dans la majorité des cas plurifactorielle, il semble évident qu'il faille prendre le patient dans son ensemble, que ce soit la prise en charge initiale du premier cancer, son parcours professionnel, ou encore son exposition environnementale géographique… Or, jusqu'alors, compte-tenu du volume de données, du cloisonnement de celles-ci et de leurs contenus hétérogènes, il était impensable de concevoir une solution permettant de répondre à cette question.

Dans le cadre du projet **I5ADORA**, **ConSoRe** est l'outil qui doit permettre la sémantisation des données médicales des 300 000 DPI du Centre Léon Bérard et une identification des patients atteints de seconds cancers primitifs. L'analyse et l'investigation des méga-données vont permettre d'étudier simultanément ces dossiers médicaux informatisés, les bases de données de risques environnementaux, les parcours professionnels des patients. Le croisement des différentes bases de données et une approche inductive, propre aux projets big data, doivent permettre de rechercher des corrélations significatives entre tous ces éléments et de définir éventuellement une stratégie de prévention personnalisée et de dépistage individualisé pour un patient donné.

Pour favoriser la réussite de ce projet, **I5ADORA** associe 3 partenaires : le Centre Léon Bérard, l'unité de recherche « Cancer et environnement » dirigé par le Pr Béatrice Fervers et SWORD Insight, qui apporte son expertise dans le domaine du big data en santé.

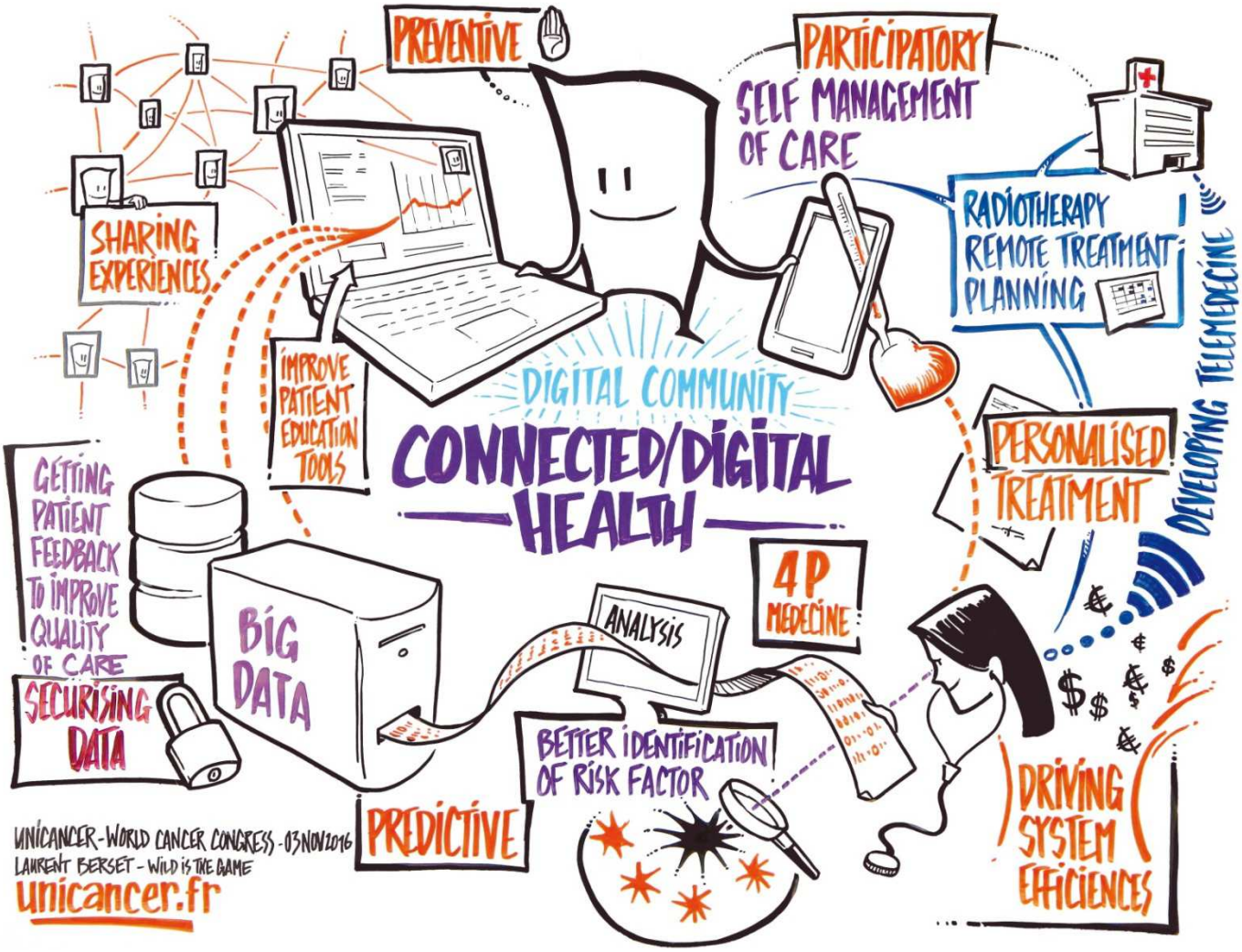
FOCUS SUR… LA SANTÉ CONNECTÉE

Les nouvelles technologies ont impacté profondément notre quotidien. Que ce soit dans nos vies personnelle ou professionnelle. Cela se traduit aussi bien par l'évolution de nos modes de paiements que par notre manière de voyager.

Le monde de la santé ne fait pas exception à cette profonde mutation de la société. Le développement d'outils tel que **ConSoRe** facilite l'accès à de plus grands volumes d'informations et permet d'avoir une vision globale de l'histoire de la maladie. Ceci couplé à des projets tel qu' **I5ADORA** va dans le sens d'une médecine à la fois préventive et prédictive., permettant de personnaliser la prise en charge de chaque patient.

Cette révolution aura sans aucun doute un impact sur la relation des patients et de leur famille avec l'hôpital et les équipes soignantes. D'un côté, il a été démontré l'impact favorable de la M-Santé (applications mobiles) sur le suivi des patients (F. Denis *et al.*, J Clin Oncol 34, 2016). Cependant, face au grand nombre d'informations disponibles il faudra certainement développer des outils pour apporter des réponses claires plus ou moins rapidement à chacune des situations particulières. Ces applications auront nécessairement un impact sur l'organisation de l'hôpital.

Au même titre que les molécules innovantes ou le séquençage moléculaire, ces nouveaux outils feront parties de l'armada thérapeutique de la médecine 4P.



Le Congrès mondial contre le cancer de l'UICC - Union for International Cancer Control - s'est tenu du 31 octobre au 3 octobre 2016 au Palais des Congrès de Paris. UNICANCER a animé un stand au sein du village d'exposition. Les visiteurs ont été invités à répondre à la question : « A quoi ressemblera, selon vous, la prise en charge du cancer, dans 10, 15, 20 ans ? ».

A partir des réponses, l'artiste Laurent Berset, a construit en direct, en s'inspirant des contributions de chacun, des fresques graphiques. Parmi les thèmes abordés, la e-santé, une des grands tendances de l'étude prospective d'UNICANCER, EVOLPEC, a été illustrée.

Directeur de la publication :
Pascale Flamant,
Déléguée Générale UNICANCER

Ont participé à ce numéro :
Pr Pierre Fumoleau (CGFL, Dijon)
Dr Pierre Heudel (CLB, Lyon)
Equipe informatique du Centre Léon Bérard (Lyon)
Christophe Jamain (R&D UNICANCER, Paris)

Coordination :
Dr Pierre Heudel, Département d'oncologie médicale Centre Léon Bérard, Lyon pierreetienne.heudel@lyon.unicancer.fr
Christophe Jamain, responsable recherche préclinique et fondamentale c-jamain@unicancer.fr
Agathe Lasne, chef de projet communication a-lasne@unicancer.fr
Emmanuel Reyrat, Directeur des Systèmes d'Information UNICANCER e-reyrat@unicancer.fr

UNICANCER

ConSoRe

Continuum Soins Recherche

JANVIER 2018

N°3

édito

L'intelligence artificielle est aujourd'hui sur toutes les lèvres. Dans le monde de la santé, on parle de son inexplicable efficacité dans l'aide à la décision, des bouleversements des prises en charge diagnostiques qu'elle induira, mais également des nouvelles questions d'ordres juridiques et éthiques qu'elle soulève... L'installation de cette première version de ConSoRe (l'installation au Centre Léon Bérard, Lyon, a eu lieu en novembre dernier), prouve, s'il était besoin, que l'innovation fait partie intégrante de l'ADN des CLCC et que cette révolution numérique est un enjeu stratégique pour nous tous. ConSoRe devra nous permettre d'exploiter les gisements de données stockées dans nos dossiers médicaux et de réfléchir à de nouveaux domaines de recherche pour la lutte contre le cancer. Les premières requêtes déjà réalisées ont montré la puissance de cet outil mais ont également démontré la nécessité de poursuivre son développement. A ce titre, il faut encourager l'étroite collaboration des différentes équipes médicales et informatiques impliquées qui correspond au sens même de la dimension participative du big data.

A titre d'exemple, le projet ConSoRe+ de l'Institut Curie, présenté dans le premier numéro de cette newsletter, démontre l'importance de cet aspect collaboratif, où l'intérêt d'une évolution individuelle bénéficie à toute la communauté. De même, le projet ConSoRe-R, développé par l'IPC, vise à créer du savoir en alimentant un lac de données pouvant être exploité par les chercheurs. Cette amélioration de ConSoRe sera prochainement implémentée dans tous les Centres utilisant ConSoRe.

Je suis d'ailleurs certain que l'énergie instillée par l'installation de ConSoRe dans les prochains CLCC permettra d'envisager d'autres axes d'évolution bénéficiant à tous.

Pr Jean-Yves Blay
Directeur général du Centre Léon Bérard, CLCC de Lyon

Actualités ConSoRe

L'installation de ConSoRe dans les Centres se poursuit activement. Le cap des 10 CLCC ayant terminé l'installation (Curie, Lille, Lyon, Marseille, Montpellier) ou étant en cours de finalisation de la mise en œuvre de ConSoRe (Bordeaux, Caen, Clermont-Ferrand, Dijon, Nancy) a été passé.

D'autres Centres ont d'ores et déjà planifié leur installation pour le courant de l'année 2018.

Les premiers retours des Centres utilisateurs font état d'un réel intérêt quant à l'utilisation de ConSoRe :

« ConSoRe est déjà au centre de la stratégie « Datas » de certains Centres et est considéré comme une des briques des fondations du BigData dans les Centres »

« La plupart des projets Data *que les Centres lancent aujourd'hui utilisent ou utiliseront ConSoRe (ou des éléments de ConSoRe) comme fournisseur de données. »*

« C'est LE moyen de valoriser plusieurs millions de compte-rendu médicaux présents dans les DPI »

Pour répondre aux attentes, ConSoRe doit évoluer sans cesse et intégrer de nouvelles sources et fonctionnalités. Un groupe de travail, aujourd'hui constitué d'utilisateurs de l'application, est en charge d'élaborer le plan de développement. Bien entendu, à l'avenir, ce groupe a vocation à être élargi aux futurs utilisateurs.

ConSoRe EN CHIFFRES

Médecins utilisateurs formés :

Institut Curie (Paris / Saint-Cloud)	75
Centre Oscar Lambret (Lille)	12
Centre Léon Bérard (Lyon)	20
Institut Paoli-Calmettes (Marseille)	30
Institut du Cancer de Montpellier	10

CAS D'USAGE N°1

Pr Patricia Niccoli, Institut Paoli-Calmetes – Marseille, département d'oncologie médicale - Référente Nationale du réseau des tumeurs neuroendocrines RENATEN - Centre d'excellence Européen pour les tumeurs neuroendocrines IPCNet center

Les tumeurs neuroendocrines font partie des cancers rares qui peuvent atteindre tous les organes, pour lesquels les classifications sont multiples et la caractérisation histologique est très spécifique. La recherche par texte disponible via ConSoRe est particulièrement adaptée et permet d'approcher l'exhaustivité.

Catégorie d'usage	**Question posée**	**Gain lié à l'usage de ConSoRe**	**Compléments de travail suite à la récupération des résultats de ConSoRe**
Recherche clinique pour évaluation des pratiques à usage de publication	Exemple : Décembre 2017 Participation à une étude Nationale sur la prise en charge des tumeurs neuroendocrines rectales.	Evaluation des pratiques Objectif d'exhaustivité.	Identification des dossiers et des données nécessaires à l'étude. Retour au dossier (identifier et sélection des dossiers éligibles => cohorte) Saisie des données dans le CRF de l'étude.
Essais cliniques	Recherche de patients porteurs de tumeurs neuroendocrines de toute origine en échec de traitement conventionnel pour inclusion dans protocole d'immunothérapie (essai Industriel et programme AcSé)	Identification des patients éligibles sur le plan des critères histologiques et en échec thérapeutiques après 2 lignes de traitement conventionnel.	Export des patients et convocation en consultation pour confirmer l'éligibilité et l'inclusion potentielle.
Recherche fondamentale	Identification de patients porteurs de tumeurs neuroendocrines pancréatiques grade 1 -2 opérées pour recherche de marqueurs moléculaires prédictifs et cultures cellulaires primaires en lien avec la biothèque.	Sélection des patients pour extraction des échantillons.	Récupération des échantillons identifiés via ConSoRe auprès du CRB

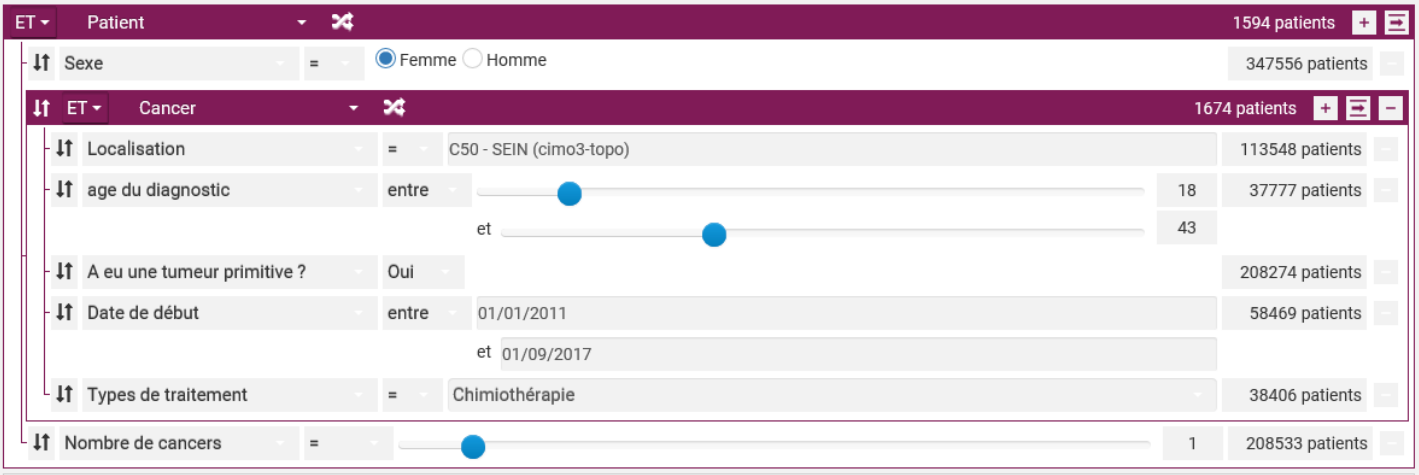
CAS D'USAGE N°2

Médecin chercheur à l'Institut Curie, Paris, le Dr Anne-Sophie HAMY-PETIT est impliquée dans les thématiques de recherches sur la fertilité après cancer du sein.
A la fin de sa thèse de sciences en bioinformatique, elle rejoindra le département des datas de l'Institut Curie.

ConSoRe a été utilisé dans le cadre d'une recherche médicale concernant la préservation de la fertilité dans le parcours de soin des patientes atteintes de cancer du sein : l'objectif est de reprendre de manière rétrospective les dossiers de femmes ayant été traitées pour un cancer du sein à l'Institut Curie et d'évaluer les résultats des procédures de préservation, les délais de traitement, les grossesses obtenues, qu'elles soient spontanées ou non et la survenue de rechutes et les décès.

En partenariat avec le département d'informatique et des data et grâce à ConSoRe, nous avons pu générer une cohorte de patientes correspondant à nos critères de sélection :

- Patientes qui avaient eu cancer du sein C50,
- Découvert entre le 1er janvier 2011 et le 1er septembre 2017,
- Avec un âge au diagnostic entre 18 et 43 ans,
- Non métastatiques d'emblée,
- Traitées avec de la chimiothérapie.



A partir de cette cohorte contrôle de 1594 patientes, la recherche des patientes ayant pu bénéficier d'une technique de préservation de la fertilité a été effectuée avec les mots clefs suivants « oncofertilité », « MIV », « cryopréservation ov* », « cryoconservation ov* », « préservation + fertilité », « ovocytes congelé », « embryon congelé », « fragment ov* », ainsi que les noms des praticiens auxquels sont adressés les patientes lors du circuit d'oncofertilité.
Sur 1594 dossiers, 531 comportaient au moins un mot clé. Parmi eux, 267 (50.3%) des patientes avaient bénéficié d'une préservation de la fertilité ; et 264 (49.7%) n'en avaient pas bénéficié.
L'utilisation de ConSoRe a non seulement grandement facilité la construction de la cohorte ; mais nous a permis de répondre à une question « non structurée » avec un taux très satisfaisait d'identification des procédures de préservation de la fertilité.

Directeur de la publication : Pascale Flamant, Déléguée Générale Unicancer
Coordination : Direction de la Communication, Direction des Systèmes d'information, et Direction de la Recherche Unicancer

NEWSLETTER N° 04
MAI 2019

CONSORE
continuum soins recherche
unicancer

édito

RENFORCER L'EXPERTISE POUR UNE MEILLEURE EFFICACITE DANS LA LUTTE CONTRE LE CANCER

Pr. Khaled Meflah, Unicancer

Renforcer l'expertise, pour une meilleure efficacité dans la lutte contre le cancer, reste un objectif inscrit résolument dans les missions des CLCC.
C'est par le développement d'une recherche active, qui doit s'appuyer sur des masses de données structurées, que nous parviendrons à promouvoir une amélioration permanente de la qualité et de la pertinence des soins. Le remarquable essor de l'intelligence artificielle, que nous vivons, fait évoluer en permanence les concepts et impose, par là même, une nécessité permanente d'adapter nos pratiques. Le développement de Consore et son évolution s'inscrivent dans cette perspective, et les équipes médicales et informatiques se doivent d'intégrer cette notion pour permettre aux CLCC de maintenir leur capacité à assurer la promotion de l'innovation dans la prise en charge des malades.

Dans les prochains mois, 11 CLCC disposeront de Consore. Ils participeront à l'évolution de cet outil de requête et seront en mesure de développer de nouveaux projets de recherche qui pourront s'appuyer sur les importantes bases de données constituées par les CLCC.

Une nouvelle gouvernance se met en place pour accompagner le déploiement de ce projet. L'équipe Unicancer en charge de Consore s'est étoffée par le recrutement d'une jeune cheffe de projet, Alexa KAPRAL, Ingénieure en génie biologique et bio-informatique et passionnée de nouvelles technologies. De plus, mon arrivée comme référent médical va permettre de renforcer le pôle scientifique de ce projet en complément du groupe des médecins référents des Centres, toujours très actifs autour du Dr Alain Livartowski qui en assure la coordination. Toute l'équipe doit avoir une attention particulière sur l'ergonomie de l'outil et en permettre une utilisation facilitée. La participation déterminée des CLCC est indispensable à la réussite d'un projet, qui devrait permettre à nos établissements de répondre à l'attente des patients en proposant une prise en charge innovante des cancers.

A propos de Consore

http://consore.unicancer.fr

La quantité de données accumulées chaque jour dans les Centres hospitaliers recèle des informations essentielles pour accroître notre connaissance du cancer : fréquence des effets secondaires, efficacité des médicaments en vie réelle, résistance de certains patients aux traitements, etc.

Les technologies du big data offrent aujourd'hui l'opportunité d'exploiter ces données pour **faire avancer la recherche et améliorer les soins en cancérologie**.

Mais le défi est de taille. Pour le relever, Unicancer construit un moteur de recherche puissant, sorte de **Google des données de santé**, capable de retrouver des informations disséminées dans le texte de centaines de milliers de dossiers des patients des Centres de lutte contre le cancer. C'est l'outil **Consore**, pour **Continuum soins-recherche**.

Cet outil, utilisant des technologies d'analyse de type **IA et Machine Learning** est en mesure, entre autres, de reconstituer visuellement l'histoire du Patient, de rechercher des patients similaires et devrait intégrer, dans un proche avenir, des données complémentaires de génétique.

Nouveautés

En 2019 Consore fait le choix de l'IA pour son moteur d'analyse des textes

Pour reconstruire l'histoire de la maladie cancéreuse de chaque patient, Consore analyse sémantiquement chacun des documents présents dans le dossier médical. Le challenge est de taille puisqu'il faut comprendre le vocabulaire médical, et le contexte dans lequel les différents concepts sont identifiés (gestion des formulations négatives ou hypothétiques ou caractérisation des antécédents).

Thomas Balezeau, Ingénieur data – Institut Curie

Jusqu'à aujourd'hui, Consore utilisait déjà deux techniques de traitement automatique du langage : l'extraction de concepts médicaux et la classification de phrases.

Afin d'amplifier la détection des concepts, nous avons enrichi le vocabulaire de l'outil avec des milliers de synonymes récupérés dans des bases de connaissances externes. Ceci a permis de récupérer plus de 30% de concepts médicaux supplémentaires.

A cela s'ajoute l'utilisation d'algorithmes de classification qui permettent :
- d'écarter les phrases négatives et hypothétiques et de pondérer les concepts qui y sont identifiés,
- de catégoriser les cancers et leurs évolutions (tumeur primaire, récidive locale ou métastatique).

Cet exercice de classification est désormais bien appréhendé par les méthodes d'intelligence artificielle et leur apport est incontestable.

Alors que le travail de traitement du texte se basait jusqu'à aujourd'hui sur l'édification de règles de détection (ou cartouches sémantiques), l'intégration de méthodes de machine learning a accru les capacités de détection de Consore

Pour atteindre une qualité de prédiction suffisante, les algorithmes de *machine learning* ont besoin d'assimiler un grand nombre d'exemples lors de l'apprentissage supervisé. **L'expertise humaine reste donc primordiale.**

Plusieurs médecins, Attachés de Recherche Clinique (ARC) et informaticiens de la communauté Consore ont annoté manuellement des milliers de phrases issues des dossiers patients. C'est grâce à cette contribution essentielle et à l'innovation technologique que Consore peut poursuivre son évolution et offrir des perspectives très intéressantes pour l'avenir en tirant partie de l'intelligence artificielle.

Etat des lieux

Consore en chiffres

270 médecins utilisateurs

5 travaux présentés ou publiés

35 000 000 documents analysés

1 700 000 patients intégrés

6 centres équipés

8 centres en cours d'équipement

Nouveauté

Consore se spécialise dans l'hématologie

Les parcours hématologique classiques étant différents des patients pris en charge pour des tumeurs solides, il convenait également de marquer cette différence dans Consore.

Un groupe de travail composé d'hématologues s'est donc penché sur ce sujet à travers 3 axes: le diagnostic, la représentation et l'axe thérapeutique

Dr. Anne Marie Stoppa, Institut Paoli Calmettes

- LE DIAGNOSTIC des pathologies hématologiques a été amélioré à travers l'établissement de synonymes identifiables dans les comptes rendus textuels.
- LA REPRESENTATION visuelle actuelle des parcours classiques des tumeurs solides (primitif/ récidive locale/métastase) a été adapté avec l'ajout du terme **progression** afin de mieux représenter l'histoire des maladies hématologiques :

Les hémopathies malignes aigues-leucémies aigues (LA), la plupart des myelodysplasies (MDS) et lymphomes non Hodgkinien (LNH) ou Hodgkinien (LH) de haut grade sont souvent curables mais survivent rarement après 2 lignes ; les hémopathies dites indolentes – myélomes (MM), leucémies lymphoïdes chronique (LLC), lymphomes de bas grade (LBG) - obtiennent des rémission, subissent de constantes récidives n'empêchant pas une survie longue.

Les syndromes myeloprolifératifs chroniques (SMP) sont plus rares et se rapprochent des pathologies indolentes chronique (LLC),

L'AXE THERAPEUTIQUE s'est focalisé sur :

- l'identification des chimiothérapies orales grâce à des algorithmes d'intelligence artificielle,

- l'identification des activités de thérapie cellulaire (Greffes autologues, allo géniques) qui concerne au moins 30 % des patients affectés d'hémopathies malignes classiquement en 1ere ou 2eme rémission.

Cas d'usage

Utilisation de Consore dans l'identification de patients cibles

Le cas d'usage illustre la recherche de leucémies aigues myéloïdes secondaires chez les patients traités pour myélome dont la survie est désormais de 10 à 15 ans au prix de récidives et de multiples lignes le plus souvent continues à partir de la 1ere rechute.

Les SPM -second Primary Malignancies- sont un des sujets de santé publique particulièrement surveillé post autorisation de mise sur le marché des drogues .
Par exemple Revlimid* -une drogue parmi les mieux tolérées et parmi les plus efficaces- est disponible en monotherapie dés la rechute depuis 20 ans et en 1ere ligne chez les sujets agés ne recevant pas d'autogreffe depuis 10 ans . Sa disponibilité chez les sujets plus jeunes à été retardée du fait de constatations initiales d'excès de second cancers en 1ere ligne post autogreffe ; il a fallu attendre 5 ans de plus pour vérifier que cet excès de SPM n'affectait pas le bénéfice de survie globale et obtenir l'AMM dans cette situation.

Nous avons choisi d'utiliser Consore pour repérer les LAM particulièrement redoutables dans cette population de sujets agés .
Consore a identifiée 23 LAM secondaires chez environ 2700 patients atteints de myélome. Il apparait que ces leucémies secondaires ne sont pas nécessairement associées à la multiplicité des lignes et des combinaisons. Elles surviennent le plus souvent après 2 lignes, et en majorité chez des patients ayant reçu une ou plusieurs autogreffes conditionnée par Melphalan haute dose .

Perspectives

Une simplification de l'outil pour une utilisation augmentée

Outre les améliorations techniques et fonctionnelles déjà effectuées, 2019 doit être l'année de l'amélioration de l'utilisation de Consore. Et pour qu'il soit plus utilisé, l'usage de Consore doit être simplifié.

Emmanuel Reyrat, Directeur des Systèmes d'information - Unicancer

- **Amélioration de la facilité d'usage de Consore par les médecins :**

Une première série d'améliorations permettant une aide à la génération de requête vient d'être livrée et permet, à partir d'un terme saisi, de proposer de multiples critères directs ou associés.

Nous souhaitons aller encore plus loin et proposer à l'utilisateur un système à double vitesse :

- **UTILISATION SIMPLE** : l'utilisateur propose une phrase dans la barre de recherche et le système analyse ces termes et génère une requêtes complète intégrant les critères demandés.

- **UTILISATION AVANCEE :** Dans ce cas, l'utilisateur peut construire, comme aujourd'hui, une requête complexe à partir d'une liste des critères de recherche.

- **Aide à la recherche**

Un système de « boîte aux lettres » SAV permettra à un utilisateur « novice » d'exposer sa recherche. Un référent Consore lui proposera alors une requête répondant au mieux à son besoin. Et lui expliquera comment, à l'avenir, générer ce type de requête pour le rendre autonome.

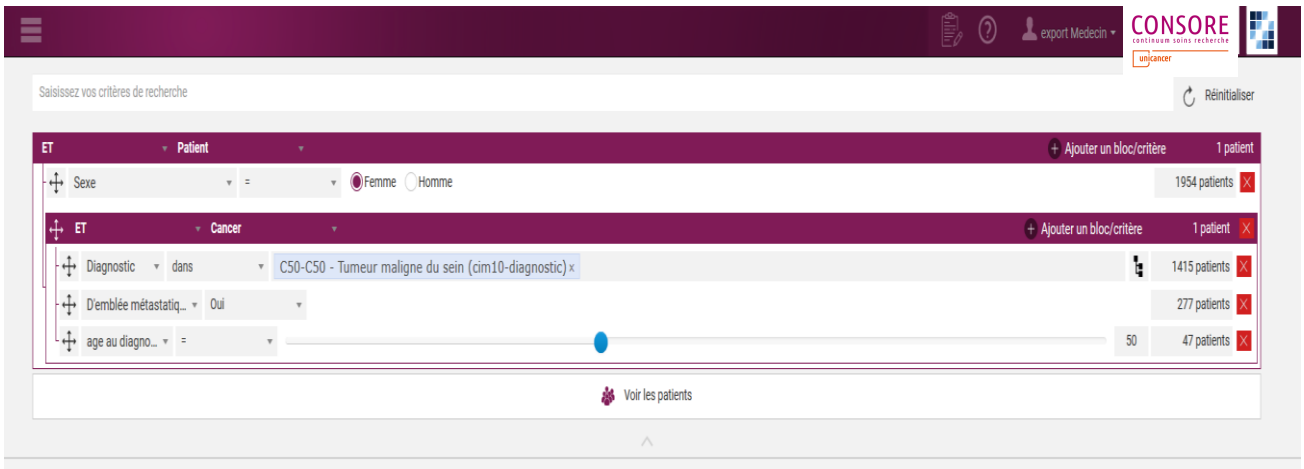
- **Amélioration des fonctionnalités d'usage de Consore**

Outre ces améliorations d'ergonomie et d'aide à l'usage, des travaux importants sont en cours sur :

- L'intégration
  - ✓ de données d'imagerie
  - ✓ de données d'essais cliniques
  - ✓ de données génétiques
- L'amélioration de la fonction d'export de données
- L'analyse de la réponse aux traitement (pour déterminer d'éventuelles résistances)

ILLUSTRATION PRATIQUE

« femme ayant eu un cancer du sein métastatique à 50 ans » devient :



Actualités

Publications autour de Consore

POSTER – JFHOD 2019

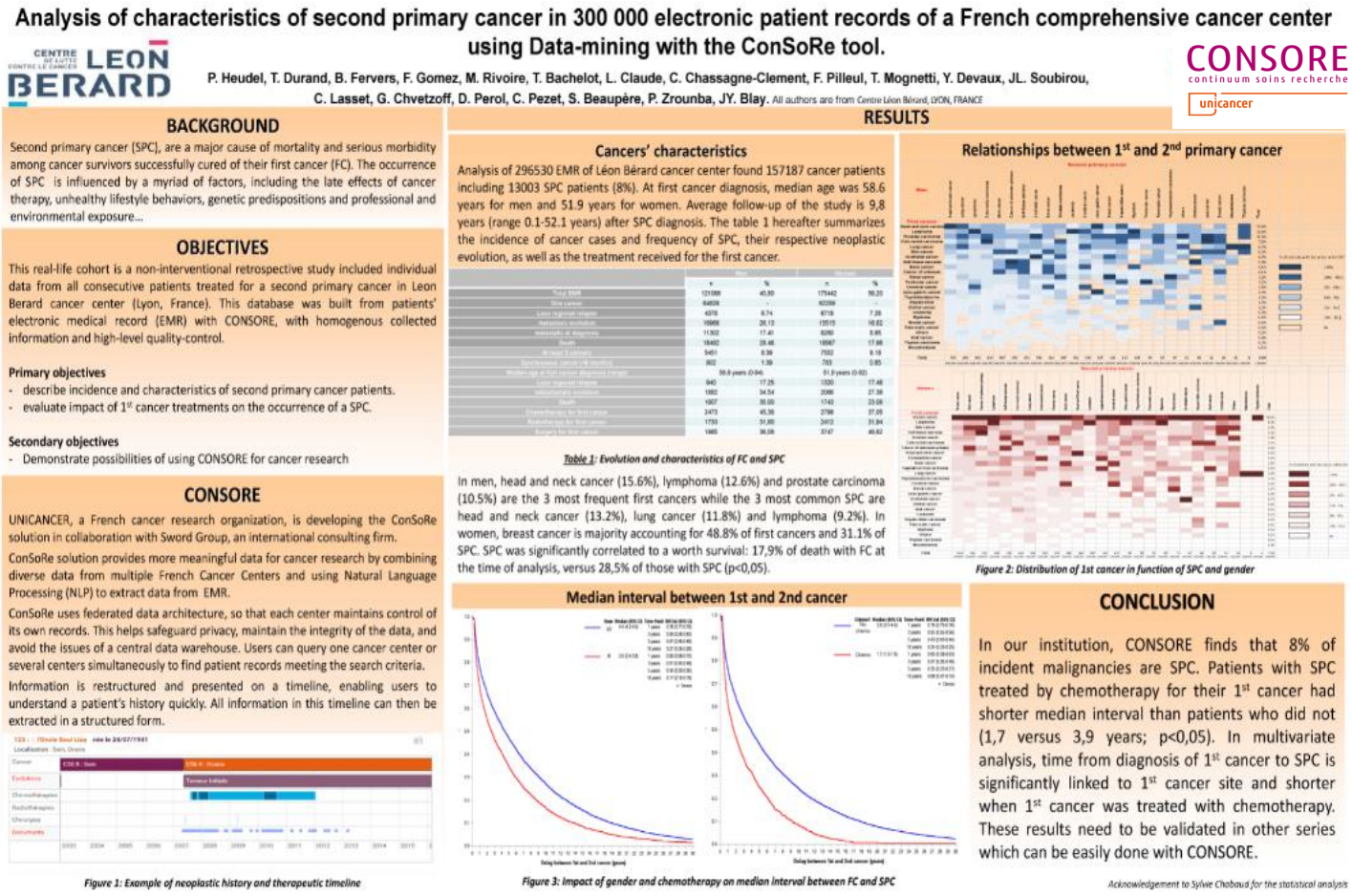
Les résultats de la résection endoscopique dans les cancers colorectaux superficiels sont-ils comparables à ceux des séries intéressant principalement les adénomes? Reste-t-il une place pour la résection piecemeal dans cette indication ?

- **Auteurs :** J.P. Ratone, E. Bories, C. Zemmour, F. Caillol, C. Pesenti, S. Hoibian, S. Godat, C. de Chaisemartin, B. Lelong, H. Meillat, F. Poizat, M. Giovannini
- **Lien :** https://www.snfge.org/resumes-2019/000143

POSTER – ESMO 2018

Data-mining of 110 172 electronic patient records with the ConSoRe tool: An analysis of second primary cancer in a comprehensive cancer center

- **Auteurs :** P Heudel T Durand B Fervers F Gomez M Rivoire T Bachelot L Claude C Chassagne-Clement F Pilleul T Mognetti Y Devaux J-L Soubirou C Lasset D Perol G Chvetzoff C Pezet S Beaupere P Zrounba J-Y Blay
- **Lien url :** https://academic.oup.com/annonc/article/29/suppl_8/mdy294.009/5141180?searchresult=1
- **Lien DOI :** https://doi.org/10.1093/annonc/mdy294.009



Analysis of characteristics of second primary cancer in 300 000 electronic patient records of a French comprehensive cancer center using Data-mining with the ConSoRe tool.

P. Heudel, T. Durand, B. Fervers, F. Gomez, M. Rivoire, T. Bachelot, L. Claude, C. Chassagne-Clement, F. Pilleul, T. Mognetti, Y. Devaux, JL. Soubirou, C. Lasset, G. Chvetzoff, D. Perol, C. Pezet, S. Beaupère, P. Zrounba, JY. Blay. All authors are from Centre Léon Bérard, LYON, FRANCE

BACKGROUND

Second primary cancer (SPC), are a major cause of mortality and serious morbidity among cancer survivors successfully cured of their first cancer (FC). The occurrence of SPC is influenced by a myriad of factors, including the late effects of cancer therapy, unhealthy lifestyle behaviors, genetic predispositions and professional and environmental exposure...

OBJECTIVES

This real-life cohort is a non-interventional retrospective study included individual data from all consecutive patients treated for a second primary cancer in Leon Berard cancer center (Lyon, France). This database was built from patients' electronic medical record (EMR) with CONSORE, with homogenous collected information and high-level quality-control.

Primary objectives
- describe incidence and characteristics of second primary cancer patients.
- evaluate impact of 1st cancer treatments on the occurrence of a SPC.

Secondary objectives
- Demonstrate possibilities of using CONSORE for cancer research

CONSORE

UNICANCER, a French cancer research organization, is developing the ConSoRe solution in collaboration with Sword Group, an international consulting firm.

ConSoRe solution provides more meaningful data for cancer research by combining diverse data from multiple French Cancer Centers and using Natural Language Processing (NLP) to extract data from EMR.

ConSoRe uses federated data architecture, so that each center maintains control of its own records. This helps safeguard privacy, maintain the integrity of the data, and avoid the issues of a central data warehouse. Users can query one cancer center or several centers simultaneously to find patient records meeting the search criteria.

Information is restructured and presented on a timeline, enabling users to understand a patient's history quickly. All information in this timeline can then be extracted in a structured form.

Figure 1: Example of neoplastic history and therapeutic timeline

RESULTS

Cancers' characteristics

Analysis of 296530 EMR of Léon Bérard cancer center found 157187 cancer patients including 13003 SPC patients (8%). At first cancer diagnosis, median age was 58.6 years for men and 51.9 years for women. Average follow-up of the study is 9,8 years (range 0.1-52.1 years) after SPC diagnosis. The table 1 hereafter summarizes the incidence of cancer cases and frequency of SPC, their respective neoplastic evolution, as well as the treatment received for the first cancer.

Table 1: Evolution and characteristics of FC and SPC

In men, head and neck cancer (15.6%), lymphoma (12.6%) and prostate carcinoma (10.5%) are the 3 most frequent first cancers while the 3 most common SPC are head and neck cancer (13.2%), lung cancer (11.8%) and lymphoma (9.2%). In women, breast cancer is majority accounting for 48.8% of first cancers and 31.1% of SPC. SPC was significantly correlated to a worth survival: 17,9% of death with FC at the time of analysis, versus 28,5% of those with SPC (p<0,05).

Relationships between 1st and 2nd primary cancer

Figure 2: Distribution of 1st cancer in function of SPC and gender

Median interval between 1st and 2nd cancer

Figure 3: Impact of gender and chemotherapy on median interval between FC and SPC

CONCLUSION

In our institution, CONSORE finds that 8% of incident malignancies are SPC. Patients with SPC treated by chemotherapy for their 1st cancer had shorter median interval than patients who did not (1,7 versus 3,9 years; p<0,05). In multivariate analysis, time from diagnosis of 1st cancer to SPC is significantly linked to 1st cancer site and shorter when 1st cancer was treated with chemotherapy. These results need to be validated in other series which can be easily done with CONSORE.

Acknowledgement to Sylvie Chabaud for the statistical analysis

Publication Unicancer

Directeur de la publication :
Patrice VIENS
Président d'Unicancer

Rédacteur en chef :
Emmanuel Reyrat
Directeur des Systèmes d'information

Participation à la rédaction
Unicancer – Institut Curie – Centre Léon Bérard
Institut Paoli-Calmettes

NEWSLETTER N° 05
Octobre 2019

CONSORE
continuum soins recherche
unicancer

édito

CONSORE : UNE « PLATEFORME DE DONNEES » DE PLUS EN PLUS PUISSANTE, EN PLEINE EXPANSION

Dr. Alain Livartowski, Institut Curie

L'année 2020 sera cruciale pour le projet Consore. Cruciale car l'outil est de plus en plus puissant tout en conservant des performances exceptionnelles en temps de réponse. De nouvelles fonctionnalités et une interface graphique améliorée permettent de mieux préciser les requêtes. Pour autant, de nombreux travaux vont être lancés et sont encore devant nous : améliorer l'export des données, mesurer la qualité des données, intégrer de nouvelles sources : la biologie médicale, la radiothérapie, l'imagerie médicale, la pathologie, la biologie moléculaire et demain la génomique.

Au-delà de l'outil, il nous faut réussir plusieurs challenges. **S'élargir à d'autres Centres signifie intégrer de nouvelles équipes dans le projet.** Des médecins référents ont été désignés mais il est nécessaire que la communauté des DIM et des responsables de la recherche clinique participent. Les informaticiens des centres sont concernés au premier chef et des échanges entre nouveaux venus et plus anciens doivent permettre à ces centres de s'intégrer plus vite et rendre service à l'ensemble de la communauté médicale.

Des études difficiles à réaliser, demandant trop de temps ou de moyens, sont devenues possibles grâce à Consore mais ceci n'est pas suffisamment connu. Il faut également démontrer que cela aboutit à plus de publications et à une meilleure évaluation de nos pratiques.

Consore est considéré par tous comme l'incontournable « plateforme de données du cancer » et ceci nous ouvre des perspectives au-delà des CLCC. Il devient possible de participer à des projets nationaux ambitieux comme le Health Data Hub ou coopérer avec d'autres entrepôts : l'entrepôt des données de santé de l'AP-HP ou le projet HUGO du Grand-Ouest.
L'aventure ne fait que commencer : ceci est possible grâce à une vision portée par Unicancer et par les nouveaux moyens humains dévolus au projet.

A propos de Consore

http://consore.unicancer.fr
Contact : info-consore@unicancer.fr

La quantité de données accumulées chaque jour dans les Centres hospitaliers recèle des informations essentielles pour accroître notre connaissance du cancer : fréquence des effets secondaires, efficacité des médicaments en vie réelle, résistance de certains patients aux traitements, etc.

Les technologies du big data offrent aujourd'hui l'opportunité d'exploiter ces données pour **faire avancer la recherche et améliorer les soins en cancérologie**.

Mais le défi est de taille. Pour le relever, Unicancer construit un moteur de recherche puissant, sorte de **Google des données de santé**, capable de retrouver des informations disséminées dans le texte de centaines de milliers de dossiers des patients des Centres de lutte contre le cancer. C'est l'outil **Consore**, pour **Continuum soins-recherche**.

Cet outil, utilisant des technologies d'analyse de type **IA et Machine Learning** est en mesure, entre autres, de reconstituer visuellement l'histoire du Patient, de rechercher des patients similaires et devrait intégrer, dans un proche avenir, des données complémentaires de génétique.

Focus sur

La protection des données : un axe majeur pour la pérennité de Consore

Ce qui est formidable dans ce projet, c'est le souci qui pousse toutes les équipes à respecter ces principes, tout en se « lâchant » dans l'intégration de fonctions réellement complexes, qui permettront elles seules de nourrir l'appétit des professionnels de santé.

Que les acteurs de chaque centre embarquent de même, à leur niveau, le principe de responsabilisation.

Astrid Lang, DPO – Institut Curie

Consore fédère des données de plusieurs centres autour d'un outillage optimisé pour les professionnels en cancérologie dans leurs réflexions de prise en charge de leurs patients. Il s'agit aussi d'une solution très pointue pour la recherche. En effet, elle interprète, grâce à des règles de structuration « métiers », la connaissance médicale en tenant compte des spécificités de parcours et des disciplines, et des méthodes d'analyse de nos professionnels.

De part le volume conséquent de données de Consore et l'utilisation croissante et variée qui en est faite, les concepteurs et ingénieurs data, doivent veiller à garantir aux utilisateurs une aide au respect de la règlementation qui s'est fortement étoffée ces derniers temps.

Le principe du « Privacy by design » doit garantir une protection optimale des données personnelles dès la conception et lors de chaque intégration d'une nouvelle technologie.

Quelques points à ne pas négliger dans ce cadre :

Le volume et la pertinence des informations, ainsi que la richesse des fonctions proposées, imposent de veiller à disposer d'une forte protection et confidentialité des données stockées et restituées, qui proviennent de multiples sources (beaucoup sites, maintes origines de données)

- Pour que les données puissent être exposées sans risques, quelles que soient leurs utilisations, il faut a minima :
  - Une authentification personnelle pour l'accès au système, avec une revue régulière des habilitations données,
  - Une étanchéité par centre pour disposer des seules données autorisées,
  - Une finesse de paramétrage des fonctions autorisées par profil,
  - Une traçabilité des actions et des auteurs.
- Des extractions de données de Consore ne doivent se faire qu'à des fins d'utilisation légitimes, argumentées, et avec une traçabilité du périmètre atteint.
- Les statistiques, ou les dénombrements de patients répondant aux mêmes critères dans d'autres centres, doivent être proposés avec un anonymat absolu.

Le besoin de perfectionnement de chaque algorithme et de garantie de résultat intègre et optimal est primordial. Afin d'éprouver les fonctions avant de les mettre à disposition des professionnels de santé, les données de « vie réelle » ont toute leur signification pour les développeurs.

Pour réduire les risques de rupture de la protection des données patient, il faut donc, dans ces phases d'apprentissage, ou dans les tests de pertinence de leur réalisation, à chaque fois :

- S'approcher de l'anonymat des informations fournies,
- Procéder, pour chaque test à réaliser, à la minimisation des données.

Il ne faut pas oublier les droits du patient, en particulier la prise en compte dans Consore de son opposition à la recherche, et donc le cas échéant une mise à jour du système.

La protection des données personnelles en ce sens doit donc être incorporée à chaque nouvelle fonctionnalité, même si elle n'est pas toujours prévue dans les technologies utilisées.

dossier

La stratégie DATA dans les Centre De Lutte Contre le Cancer

Cette stratégie suit la vision d'innovation de l'IPC qui consiste à "dresser une armée en ordre de marche, bien déterminée à faire reculer le cancer"

Manuel Tejeda, Institut Paoli-Calmettes

La stratégie DATA et Innovation de l'IPC s'appuie sur 5 piliers :

1. La collecte et la valorisation du patrimoine de données IPC (DATA LAKE), à savoir les données médicales (300 000 patients et 4 Millions de CR dans Consore) et données de vie des patients recueillies via IPC Connect qui compte 10 000 patients connectés et suivis en 18 mois, Prom's, et IOT (objets connectés).

2. La création d'un pôle Data composé d'une équipe de 3 personnes et d'un budget de 4 Millions d'Euros.

3. L'innovation. Le LIDAO (Laboratoire d'Innovation Digitale Appliquée à l'Oncologie) a été créé avec une gouvernance assurée par une équipe pluridisciplinaire de médecins, chirurgiens, recherche clinique et recherche translationnelle, bio-informaticiens, data-experts... L'idée est d'identifier les projets digitaux innovants de l'institut ou proposés par des partenaires. Trois partenariats d'intelligence artificielle ont déjà été signé.

4. La construction d'un bâtiment hébergeant startups et partenaires sur des projets identifiés par le LIDAO autour de l'IA, le Big Data, l'IOT, la réalité virtuelle, la blockchain, l'open data…

5. Une ambition européenne : un partenariat a déjà été tissé avec l'université d'Oxford. Ces coopérations internationales feront dialoguer les différents systèmes. Une équipe de l'IPC participe ainsi au projet OSIRIS qui a publié un modèle de donnée partageable via les standards internationaux FHIR.

Cette stratégie est un cycle vertueux d'innovations : les IOT vont générer des DATA qui vont à leur tour être la source de projets innovants au service de la vocation d'innovation de l'IPC pour faire reculer le cancer.

Le Centre Léon Bérard souhaite s'affirmer comme un acteur majeur de l'intelligence artificielle au niveau national. Nous y voyons un outil capable de révolutionner à la fois le dépistage, le suivi des patients et la recherche en cancérologie. C'est une opportunité unique d'associer médecins et ingénieurs dans une optique commune.

Hugo Crochet - Centre Léon Bérard

Nous menons actuellement de nombreux projets d'intelligence artificielle, que ce soit en collaboration avec des entreprises innovantes ou des laboratoires de recherche. **Ces partenariats riches permettent de mettre en commun les données de l'hôpital et les compétences en data science de nos partenaires.** Des compétences internes sont en cours de développement et permettront de continuer à enrichir ces projets.

Préambule indispensable au développement de projets d'intelligence artificielle, nous investissons beaucoup dans la construction de nos données.

Notre stratégie Data repose sur les axes suivants :

- Augmenter le périmètre de données accessibles, notamment dans l'outil Consore, pour pouvoir lier des sources de données différentes et complémentaires
- Améliorer la qualité (exactitude, structuration, fiabilité) des données pour faciliter leur utilisation rapide
- Favoriser l'accessibilité des données, au sein de l'établissement mais aussi au niveau national en participant à des projets comme Substra ou le Health Data Hub
- Transformer les contraintes liées au RGPD en force de progrès en les intégrant au plus tôt dans les projets

dossier

Ce sont aujourd'hui plus de 80 utilisateurs enregistrés dans le Centre et au-delà de 460 requêtes réalisées en moyenne chaque mois. C'est pourquoi, Consore est maintenant au cœur de la stratégie « Data » portée par notre Direction des Données.

Julien Guérin – Institut Curie

La capacité remarquable de Consore à interpréter les dossiers médicaux pour reconstruire l'histoire tumorale de chaque patient permet d'accélérer la recherche d'informations pertinentes et la constitution de cohortes.

Notre réactivité a ainsi été décuplée pour répondre aux demandes récurrentes en données de nos interlocuteurs (médecins, chercheurs mais également nos Directions). De plus, le système nous permet d'évaluer rapidement les données disponibles afin de pouvoir contribuer à de nombreux appels à projets, notamment autour du développement de l'Intelligence Artificielle.

Mais, Consore ne peut adresser tous les besoins. De ce fait, nous avons construit autour de la plateforme une offre de services « Data » beaucoup plus large permettant de répondre aux différents enjeux :

- Un entrepôt de données « expert » agrégeant des données supplémentaires et dédié à la réalisation d'extractions personnalisées et d'études thématiques,
- Une base cartographiant l'ensemble de nos images disponibles et permettant leur mise à disposition dans le cadre de projets de recherche,
- L'utilisation d'un outil de collecte de données permettant d'héberger les données d'études rétrospectives (REDCap),
- Une plateforme de Data Science, en cours de développement, qui offre des espaces de travail sécurisés aux chercheurs et bénéficiant d'outils d'analyse (R, Python, Jupyter, cBioPortal) et qui repose sur l'usage d'un cluster de calcul.

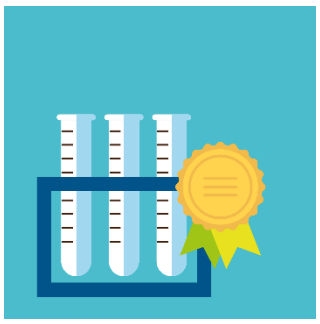
Alors qu'un grand nombre de ces services est utilisé en complément, ou intègre des données de Consore, nous devons porter une attention accrue à la qualité des données produites. Bien que celle-ci se soit nettement améliorée, elle reste toujours perfectible.

Bien évaluer la qualité des données et l'améliorer est un leitmotiv pour garantir la confiance de nos utilisateurs.

Nous avons, par exemple, mis en place un processus automatisé de mesures comparatives avec des registres ou des bases de données de référence (Base Sein, ESME).

Les travaux à réaliser sont encore multiples, souvent complexes, mais la mise en œuvre de notre stratégie « Data » se poursuit pour créer les outils qui accompagneront la recherche en cancérologie de demain.

Publication



Analysis of prognostic factors on overall survival in elderly women treated for early breast cancer using data mining and machine learning

P Heudel, D Hooijenga, R Phan, V Augusto, X Xie, C Terret, C Faure, S Racadot, O Tredan, T Bachelot

Annals of Oncology, Volume 30, Issue Supplement_5, October 2019, mdz257.022,

https://doi.org/10.1093/annonc/mdz257.022

Centre Jean PERRIN
unicancer Clermont Auvergne Métropole

interview

La parole aux centres

Rencontre avec l'équipe de coordination

Pr Xavier Durando, coordinateur Médical, Catherine Abrial DPO, Dr Juliette Afonso DIM, Alexia Giro, ARC, Anthony Raymond, chef de projets service informatique

Comment s'est déroulée la mise en œuvre de Consore au sein de votre centre ?

« Le Pr X Durando, responsable du département d'Oncologie a été nommé coordinateur médical . Une fois la plateforme ConSore opérationnelle techniquement, sous l'impulsion du DPO et du service informatique, le consensus a été de limiter l'accès aux médecins salariés du Centre Jean Perrin, au personnel de la recherche clinique (ARC et chefs de projets) et aux ingénieurs biologistes. Le choix a été fait de ne pas inclure les internes du fait de leur turn over trop fréquent. Ceux-ci peuvent néanmoins sous la responsabilité d'un médecin utiliser Consore. »

Quel est leur premier ressenti ?

« D'abord la satisfaction d'être autonome pour interroger un tel volume d'informations avec la possibilité de recherche en full text. Consore est particulièrement pertinent sur la recherche de pathologies rares. »

Qui sont vos utilisateurs de Consore et quels en sont leurs usages?

« Les premiers utilisateurs sont les ARC, l'objectif étant d'augmenter le taux d'inclusion dans les essais (screening). Ainsi, chaque ARC gère en moyenne une dizaine de protocoles et en amont de la consultation, il pré-sélectionne via Consore, les patients potentiellement incluables par les médecins. »

Quelles évolutions fonctionnelles envisageriez-vous pour Consore ?

« Lors de l'export des résultats, avoir à notre disposition davantage de critères que l'on pourrait exporter et au moins les critères de sélection (ex : médicaments sélectionnés). »

Le centre Jean Perrin en chiffres : utilisation de Consore

Démarrage officiel
12/06/2019

115 médecins utilisateurs
16 chercheurs utilisateurs

353 682 patients intégrés

1 968 744 compte-rendu

Cas d'usage

Etude multicentrique : Consore au service des réponses exceptionnelles

Consore permet d'identifier un grand nombre de patients de manière efficace, rapide et reproductible.

Brice Chanez,
Institut Paoli-Calmettes

L'institut Paoli Calmette a été sollicité pour participer à l'étude EXPRESS, dont l'objectif est de décrypter les mécanismes moléculaires des tumeurs de patients présentant une réponse tumorale exceptionnelle aux thérapeutiques ciblées (TC). Pour cela il est nécessaire d'identifier les patients sous TC et longs répondeurs porteurs d'un cancer du poumon, sein, rein, ovaire, mélanome, colon ; avec du matériel tumoral disponible pour analyse génomique par la suite.

Grace à Consore, la recherche de ces patients a pu être simplifié et accélérée en interrogeant la base de données avec des mots clés présélectionnés.

Les critères de recherche initiaux ont consisté à lister les mots clé les plus couramment utilisés par les médecins dans les courriers médicaux ou compte rendu d'examen d'imagerie pour définir une excellente réponse tumorale : «réponse complète», «très bonne réponse», «longue réponse» entre autre. Nous avons couplé ces mots clé avec les thérapies ciblées concernées par l'étude ce qui permettait d'obtenir un très grand nombre de résultats. Pour filtrer, nous avons défini les patients qui étaient en cours de traitement par de la thérapie et en utilisant les options de filtre sur la survie, nous avons fixé le curseur sur deux ans**.**

Une fois la cohorte de patients identifiée, Consore a permis de sélectionner les patients susceptibles d'être retenu en fournissant une information sur la présence de matériels disponible et le type de conservation pour chaque patient.

EVENEMENTS

Consore est présent sur le **stand n°7** lors de la Convention nationale des Centres.
Venez nombreux nous retrouver pour échanger et découvrir l'outil et ses multiples fonctionnalités.

Réunion d'un groupe d'experts (généticiens, bio-informaticiens, oncologues, onco-généticiens) dans l'objectif de **définir les données génétiques pertinentes à intégrer dans Consore**, dans le cadre de l'évolution des connaissances dans le domaine de la génétique qui génèrera un ensemble important de données indispensables pour la mise en œuvre d'une médecine de précision pertinente.

Temps d'information, d'échange et de partage de connaissances entre les binômes de référents Consore (médical & data) des Centres, dans l'objectif **d'impulser une dynamique Consore et d'impliquer l'ensemble des Centres dans la réussite de ce projet.**

Publication Unicancer

Directeur de la publication :
Jean-Yves Blay
Président d'Unicancer

Rédacteur en chef :
Emmanuel Reyrat
Directeur des Systèmes d'information

Participation à la rédaction
Unicancer – Institut Curie – Centre Léon Bérard
Institut Paoli-Calmettes – Centre Jean Perrin