

Développement d'un outil de pilotage des patients oncologiques au CHU de Liège

Bodart Gwennaëlle ¹, **Thys Marie** ², **Gangolf Marjorie** ², **Cunin Marie-Pierre** ², **Lambert Florence** ², **Poustovoitov Alexandre** ³, **Beguyn Yves** ¹, **Kolh Philippe** ⁴, **Jacques Jessica** ²

¹ CHU Liège, Institut de Cancérologie Arsène Burny, Avenue de l'hôpital 4000 Liège, 0032(0)4/3667198, gbodart@chuliege.be

² CHU Liège, Service des Informations Médico-Economiques, Avenue de l'hôpital 4000 Liège

³ CHU Liège, Service des Informations Médico-Economiques, Avenue de l'hôpital 4000 Liège

⁴ CHU Liège, Gestion du Système d'Informations, Avenue de l'hôpital 4000 Liège

Résumé. Le nombre de cas de cancer est en augmentation constante, en Belgique comme partout dans le monde. Plusieurs services du CHU de Liège se sont regroupés au sein d'un Institut de Cancérologie, afin d'assurer une gestion intégrée multidisciplinaire de cette masse de patients. Cet article présente comment les données de l'entrepôt de données cliniques institutionnel ont été exploitées afin de développer un outil de pilotage transversal de l'activité des patients oncologiques.

Mots clés: entrepôt de données cliniques, patient oncologique, gestion des institutions de soins.

Introduction

Avec plus de 2500 nouveaux diagnostics de cancer par an, le CHU de Liège est un acteur majeur des soins oncologiques dans la province de Liège. L'hôpital a donc entamé il y a quelques années la construction d'un tout nouveau bâtiment, l'Institut de Cancérologie Arsène Burny. Ce dernier est dédié à accueillir toute l'activité ambulatoire et hospitalière de jour de ces patients atteints d'un cancer. Dans ce cadre, il est vite apparu nécessaire de disposer d'un outil permettant de suivre l'activité de ce nouvel Institut ainsi que ses objectifs stratégiques et économiques.

De nombreux outils de pilotage institutionnels existaient déjà au CHU de Liège, par exemple pour le suivi de l'activité clinique ou de la facturation (Jacques *et al.*, 2020). Cependant, leur approche orientée « discipline » ne permettait pas d'isoler l'activité d'une population bien spécifique de patients, ici les patients oncologiques. En effet, leur parcours est le plus souvent multidisciplinaire au sein de services qui, pour la plupart, ne sont pas uniquement dédiés à l'oncologie. Il a donc fallu, dans un premier temps, créer une base de données recensant les patients oncologiques de l'institution, grâce aux différentes sources d'informations présentes dans l'entrepôt de données cliniques de l'hôpital. Ensuite, un outil de pilotage a été développé, afin de présenter de manière transversale et centralisée différents indicateurs relatifs à cette population de patients oncologiques.

Ce travail a été réalisé dans le cadre du projet Interreg Oncocare, qui, entre 2017 et 2021, a réuni les hôpitaux universitaires d'Aachen, Maastricht et Liège autour de plusieurs actions visant à améliorer la prise en charge et la qualité des soins oncologiques dans l'Euregio, notamment au travers du benchmarking.

Cet article présente la façon dont les patients oncologiques ont été identifiés à partir des données de l'entrepôt de données cliniques, et comment cette information a été utilisée pour développer un outil de pilotage dédié à la gestion d'un centre de soins oncologiques.

Matériel et Méthodes

L'entrepôt de données cliniques institutionnel

Le CHU de Liège, en tant qu'hôpital universitaire, assume des missions d'activité clinique, de recherche, d'enseignement et de service à la communauté. Depuis 2002, le CHU de Liège a déployé un vaste Dossier Patient Informatisé communément adopté par les prestataires de soins de l'institution. Dans les suites de ce déploiement, au cours de ces dernières années, l'hôpital s'est doté d'un entrepôt de données cliniques intégrant une vaste étendue de données administratives et cliniques (Jacques *et al.*, 2020). Cet entrepôt s'est ainsi enrichi de nombreuses données au fil des ans, comme le décrit la Figure 1, avec notamment les données administratives du patient, les données de facturation, les mouvements (admissions et transferts), les données du Résumé Hospitalier Minimum (RHM), les données médicales et infirmières du Dossier Patient Informatisé (DPI), les médicaments et chimiothérapies, ...

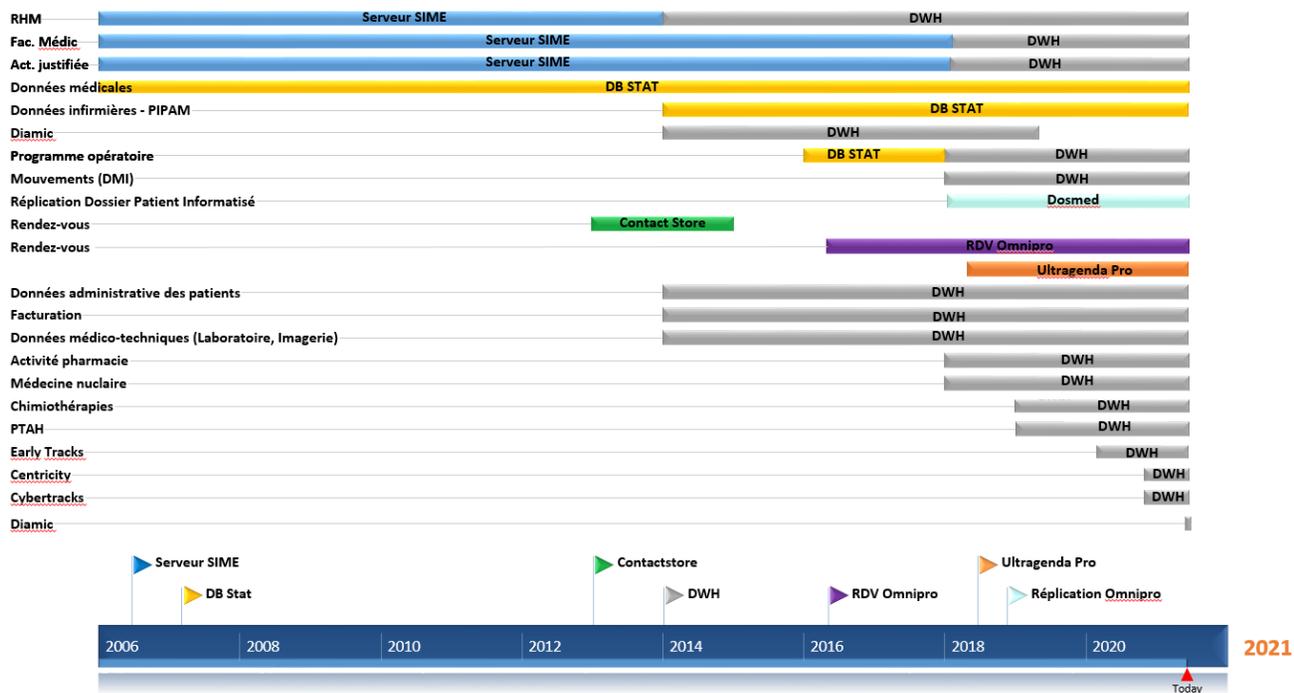


Figure 1 Déploiement de l'entrepôt de données cliniques au CHU de Liège

Grâce à cette richesse d'informations médicales et administratives, de nombreux outils de pilotage ont pu être développés par l'équipe d'analystes du Secteur Exploitation des données du Service des Informations Médico-Economiques. Certains sont transversaux, tels que le suivi de l'activité clinique, les profils de facturation ou l'activité justifiée, et d'autres spécifiques à certains services, comme l'activité de la pharmacie, de la maternité ou de l'imagerie.

Définition du patient oncologique

La difficulté principale rencontrée dans l'établissement d'un outil de pilotage transversal pour l'Institut de Cancérologie était de pouvoir identifier les patients oncologiques. En effet, le domaine des pathologies oncologiques est extrêmement varié, avec plusieurs centaines de types de cancers différents, pouvant affecter n'importe quel organe ou système du corps humain, et avec des parcours de soins très différents d'un cancer à un autre. Il était donc nécessaire de s'accorder sur une définition unique et transversale du patient oncologique, qui est la suivante :

« Tout patient avec un nouveau diagnostic ou une récurrence de cancer est considéré comme patient oncologique pour une période de -2 mois à +5 ans autour de la date d'incidence ou de récurrence ».

Cette fenêtre oncologique de cinq ans correspond à la durée standard de suivi d'un patient atteint d'un cancer. Bien entendu, dans la réalité de terrain, la durée de suivi varie d'un patient à l'autre et d'un type de cancer à l'autre, notamment les cancers chroniques qui nécessitent un suivi à vie. Cependant, dans le cadre d'un outil institutionnel transversal, il n'était pas possible de tenir compte de chaque particularité et exception. Par ailleurs, la fenêtre de deux mois précédant la date de diagnostic permet d'inclure la phase de bilan qui précède le diagnostic définitif, qui se base généralement sur la date de confirmation anatomopathologique et ainsi d'estimer correctement l'activité générée en phase pré-diagnostique.

Le périmètre de diagnostics pris en compte est celui défini par le Registre Belge du Cancer pour l'enregistrement obligatoire des cancers. Ce dernier se base sur les recommandations de l'European Network of Cancer Registries ENCR (Tyczynski *et al.*, 2003) et reprend les tumeurs malignes invasives et in situ, les hémopathies malignes, les tumeurs du système nerveux central, y compris celles avec des comportements bénins et incertains et enfin les tumeurs de faible potentiel malin ou de comportement incertain urothéliales et des ovaires.

Il faut également que le patient ait reçu une prise en charge active de son cancer par un médecin spécialiste du CHU, qu'elle soit diagnostique et/ou thérapeutique. Les patients ayant juste bénéficié d'une analyse de laboratoire, d'un examen technique ou d'une consultation pour second avis ne sont pas repris.

Parmi les informations disponibles dans l'entrepôt de données cliniques institutionnel, trois critères nous semblaient prioritaires :

- L'information devait être structurée pour être facilement exploitable. Il n'était pas possible, à ce stade, d'utiliser du texte libre par exemple
- Les données devaient permettre de fournir le diagnostic précis selon une codification internationale
- Il fallait pouvoir définir une date de diagnostic (incidence) ou du moins s'en approcher le plus possible.

Trois sources d'information nous semblaient pouvoir répondre à ces critères pour identifier les patients oncologiques. Il s'agit 1) des enregistrements cancer du DPI, 2) du Résumé Hospitalier Minimum et 3) des traitements de radiothérapie. Ainsi, une base de données la plus complète possible des patients oncologiques de l'institution a pu être créée, renseignant pour chaque patient la pathologie oncologique dont il est affecté et la fenêtre de temps durant laquelle ce patient est considéré comme oncologique.

Les enregistrements cancer. Tout programme de soins de base en oncologie a l'obligation légale de participer à un enregistrement du cancer (Arrêté Royal du 21 mars 2013). Cet enregistrement est réalisé par des gestionnaires de données spécialement formés par le Registre Belge du Cancer, ce qui garantit une très haute qualité des données. Deux types d'enregistrements existent : les enregistrements « 1^{er} cancer » pour les nouveaux diagnostics, et les « follow-up » pour les récurrences. Au CHU, ces enregistrements sont encodés directement dans le DPI. On y retrouve la date d'incidence, le diagnostic de la tumeur primaire encodé en terminologie ICD-O-3, la date de la Concertation Oncologique Multidisciplinaire (COM) ainsi que quelques infos sur le traitement. Il s'agit de la source la plus complète et la plus fiable pour identifier les patients oncologiques

de l'institution. Néanmoins, certains patients oncologiques échappent à l'enregistrement des cancers, soit car non portés à l'attention des gestionnaires de données (ce qui est fréquent pour les récidives), soit car il s'agit d'un patient référé pour un traitement ponctuel et déjà enregistré dans son hôpital d'origine.

Le Résumé Hospitalier Minimum. Dans le cadre du financement des hôpitaux belges, ces derniers sont soumis à un enregistrement obligatoire de données sur les séjours hospitaliers, incluant notamment le diagnostic responsable de l'hospitalisation, encodé en terminologie ICD9 (jusqu'à 2014) ou ICD10 (à partir de 2015). Les patients oncologiques ont été récupérés via les séjours RHM dont le diagnostic principal est de type oncologique, et en excluant les patients ayant déjà un enregistrement cancer dans les 5 ans précédents. Malheureusement, hormis le diagnostic, aucune autre information ne peut être tirée du RHM concernant le cancer du patient. On ignore donc la date d'incidence et s'il s'agit d'un nouveau diagnostic ou d'une récidive. Dès lors, la date de début du séjour est prise en compte comme approximation de la date d'incidence.

Les traitements de radiothérapie. Le service de radiothérapie du CHU de Liège maintient sa propre base de données des patients ayant reçu un traitement de radiothérapie. Ils y renseignent le diagnostic nécessitant le traitement, en utilisant la terminologie ICD9. La date d'incidence n'est pas communiquée, c'est donc la date de la première consultation en radiothérapie qui est considérée. Les patients ayant déjà un enregistrement cancer ou un séjour RHM dans l'année qui précède le traitement de radiothérapie sont exclus. En effet, on suppose que si le patient reçoit de la radiothérapie plus d'un an après le diagnostic de son cancer ou sa récidive, c'est probablement dans le cadre d'une nouvelle récidive et non pas dans le décours de la ligne précédente.

Classification commune. Les trois sources utilisant des terminologies différentes pour le codage des diagnostics (ICD-O-3 topographie, ICD-O-3 morphologie, ICD9 et ICD10), une classification commune a été établie afin de regrouper dans une même classification de catégories et sous-catégories de cancers les différents codes propres à chaque terminologie. Par exemple, dans la catégorie « Urologie », sous-catégorie « Prostate » sont regroupés les codes suivants : ICD-O-3T C61.9 ; ICD9 185 (invasif) et 2334 (in situ) ; ICD10 C61 (invasif) et D075 (in situ).

Résultats

Grâce à une définition unique institutionnelle du patient oncologique et via les multiples sources disponibles dans l'entrepôt de données clinique du CHU de Liège, une base de données des patients oncologiques a pu être établie et intégrée à son tour à l'entrepôt de données. Comme le montre la Table 1, 80% des patients oncologiques sont identifiés via un enregistrement cancer, ce qui démontre qu'une partie de la population échappe à ces enregistrements spécifiques. Ainsi, l'avantage de travailler avec un entrepôt de données cliniques est de pouvoir assurer une meilleure exhaustivité dans le recrutement d'une population de patients. Ici, le RHM et les traitements de radiothérapie contribuent respectivement à 12,3% et 7,7% des cas inclus dans la base de données des patients oncologiques du CHU de Liège.

Source	% des lignes
Enregistrement cancer nouveau diagnostic	55,8%
Enregistrement cancer Follow-up	24,2%
RHM	12,3%
Radiothérapie	7,7%

Table 1 Contribution de chacune des sources de l'entrepôt de données cliniques dans le recrutement des patients oncologiques

En croisant ce « flag patient oncologique » avec les autres données disponibles, par exemple les données de facturation ou les mouvements, on peut ainsi évaluer l'activité spécifiquement générée par cette catégorie de patients.

A la demande du Conseil de Gouvernance de l'Institut de Cancérologie, un outil de pilotage a été développé pour permettre aux décideurs de quantifier et suivre l'activité des patients oncologiques. Celui-ci s'articule autour de cinq catégories d'indicateurs :

1. Statistiques sur la population des patients oncologiques : suivi du nombre de patients actifs, nouveaux patients, nombre d'enregistrements cancer et de COM
2. Activité clinique : admissions en hospitalisation classique, hospitalisation de jour, consultations, ... chez les patients oncologiques
3. Activité financière : montant des honoraires, des produits pharmaceutiques, ... générés par les patients oncologiques
4. Prestations spécifiques au domaine oncologique : prestations, actes ou procédures spécifiques à certains domaines de l'oncologie (par exemple la pose de port-à-cath ou le nombre de greffe de cellules souches) dont le suivi est pertinent pour les services médicaux
5. Traitements : informations spécifiques concernant certains types de traitement. Actuellement, seul le volet « Radiothérapie » a été développé, et présente l'évolution du nombre de traitements et de séances, le type de traitement et de machine ou encore le délai de mise en traitement.

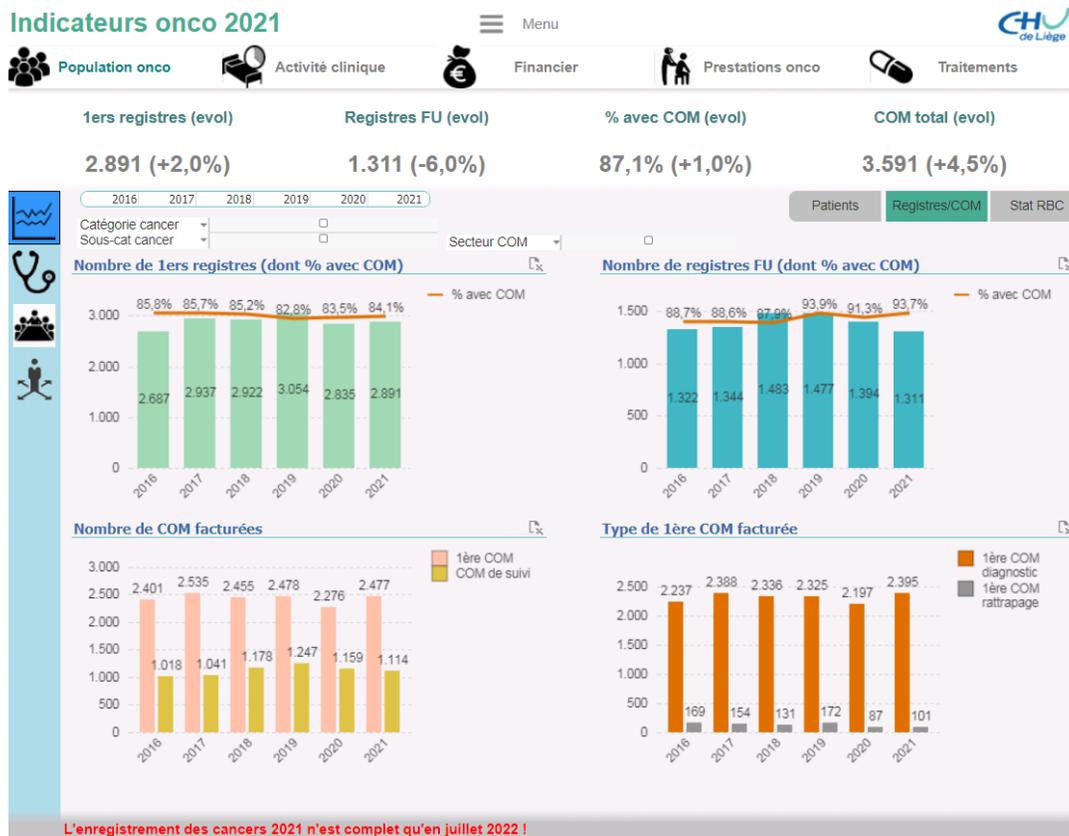


Figure 2 Exemple d'un écran de l'outil de pilotage, présentant les données sur les enregistrements cancer et COM. On peut naviguer d'une catégorie d'indicateurs à l'autre via les onglets supérieurs.

Le principe de base utilisé à travers l'outil est de considérer que toute activité ayant lieu au cours de la fenêtre oncologique du patient est labellisée « onco », quel que soit le type d'activité, le service ou le motif. En effet, il est impossible, surtout à une si vaste échelle, de rattacher une activité hospitalière à une pathologie spécifique. C'est pourquoi on se base sur le statut du patient. Dès lors, il est fort probable d'inclure comme oncologique de l'activité qui n'a en réalité rien avoir avec le cancer du patient. Par exemple, un patient qui a, en plus de son

cancer, une maladie chronique suivie dans l'institution, ou un patient admis aux urgences pour un infarctus trois ans après un cancer en récurrence. Cela entraîne une légère surestimation de l'activité présentée comme oncologique, mais ce biais est supposé constant au fil du temps. La plupart des graphiques et tables de l'outil de pilotage présentent les résultats à la fois en nombre absolu chez les patients oncologiques et en proportion relative par rapport au total des patients de l'institution. C'est surtout l'évolution de cette proportion au fil du temps qu'il est intéressant de regarder.

L'outil est conçu pour offrir une vision globale, à l'échelle de l'institution, mais permet également d'analyser la répartition selon le type de cancer. Dans le cas où un patient souffre de plusieurs cancers, dont les fenêtres oncologiques se chevauchent, il n'est évidemment pas possible de distinguer l'activité générée par le cancer A de celle du cancer B. Dès lors, l'activité est prise en compte dans chacune des catégories de cancer, mais n'est comptabilisée qu'une fois sur le total global.

Les indicateurs d'activité clinique et financière sont déclinables par service. Cela permet d'analyser spécifiquement l'activité oncologique des services rattachés à l'Institut de Cancérologie, mais également d'évaluer l'activité induite dans d'autres services (par exemple en cardiologie en cas de chimiothérapie cardiotoxique), tenant compte d'une surestimation possible. A l'inverse, l'analyse par service révèle aussi des problèmes de sous-estimations, comme par exemple le fait que seules 85% des visites ambulatoires du service d'Oncologie médicale sont attribuables à des patients oncologiques, alors qu'on s'attendrait théoriquement à 100% dans un service dédié uniquement au traitement des cancers. Cela met en évidence les limites de la méthode utilisée : il peut s'agir de patients qui sont au-delà de leur fenêtre oncologique, de patients venus juste pour un second avis, ou encore de patients réellement oubliés dans la base de données.

Des données provenant d'un rapport externe ont également été intégrées dans l'outil de pilotage. Il s'agit d'un rapport transmis par le Registre Belge du Cancer, détaillant les incidences et la survie pour chaque type de cancer des patients du CHU de Liège, de la province de Liège et la Belgique. La présentation de ces données dans l'outil de pilotage permet un « benchmarking » du recrutement des patients et de la qualité des soins au CHU par rapport au reste de la province de Liège et au niveau national.

Le Conseil de Gouvernance de l'Institut de Cancérologie s'inscrit dans une démarche de transparence, et souhaitait que l'outil de pilotage suive cette démarche. Cependant, certaines données sont sensibles, comme les données médico-économiques détaillées par service. C'est pourquoi, une sécurité avec des accès différents selon le niveau d'habilitation des utilisateurs a été mise en place. Les responsables de l'Institut souhaitaient également, toujours dans un souci de transparence, que les statistiques sur les nombres de patients et de COM soient accessibles à toute personne travaillant pour l'Institut, ce qui représente plusieurs centaines de personnes. Il n'était pas possible, pour des questions de licences de logiciel, d'ouvrir un accès à l'outil de pilotage à autant d'utilisateurs. Dès lors, la solution qui a été envisagée est la génération de rapports automatisés envoyés à fréquence régulière. Ces rapports présentent l'avantage de fournir aux utilisateurs une information ciblée (un rapport par catégorie de cancer) dans un format qui leur est familier (excel ou PDF).

Conclusion

La richesse et la diversité de l'entrepôt de données cliniques institutionnel a permis d'établir une base de données des patients oncologiques du CHU de Liège, et de croiser cette base de données avec d'autres sources de données concernant l'activité de l'hôpital au sein d'un outil de pilotage transversal pour l'Institut de Cancérologie. C'est la première fois qu'un outil de pilotage se concentrait sur une population de patients bien spécifique, tout en étant multidisciplinaire.

Cet outil donne une vue macroscopique de l'activité médico-économique des patients oncologiques. Cependant, il comporte des biais que l'utilisateur doit prendre en compte lors de l'analyse des données.

Premièrement, la base de données des patients oncologiques n'est pas exhaustive, malgré une recherche via trois sources de données distinctes, les enregistrements cancers, le RHM et la radiothérapie. En 2020, les données de chimiothérapies ont été ajoutées à l'entrepôt de données cliniques. Malheureusement, il n'y a aucune référence à la pathologie du patient dans cette source de données. Or certaines molécules de chimiothérapies sont également utilisées dans des pathologies non oncologiques. Il n'est donc pas possible actuellement de détecter les patients oncologiques via les données de chimiothérapie. Dans un futur proche, les technologies d'intelligence artificielle, en particulier le Natural Language Processing (NLP), pourraient offrir une solution intéressante, grâce au traitement automatique des données médicales non structurées du DPI tels que les rapports de consultation et d'hospitalisation, les protocoles d'anatomopathologie ou les protocoles d'imagerie (Si et Roberts, 2018 ; Abedian *et al.*, 2022).

Deuxièmement, le fait de travailler sur un domaine aussi large que l'oncologie nous contraint à adopter une définition unique du patient oncologique, sans prendre en compte les particularités propres à chaque type de cancer. Or, il apparaît évident que la fenêtre oncologique de cinq ans n'est pas adéquate pour tous les types de cancers, ni tous les patients. Par exemple, certains cancers chroniques nécessitent un suivi à vie, tandis que d'autres cancers très peu agressifs ne nécessitent presque aucun suivi.

Enfin, la méthode de calcul utilisée pour définir l'activité dite oncologique entraîne également des biais. En effet, en prenant en compte toute l'activité du patient au cours de sa fenêtre oncologique, on risque d'inclure des contacts n'ayant aucun lien avec le cancer du patient.

Pris ensembles, ces biais vont entraîner dans certains cas une surestimation des résultats, par l'inclusion d'activité non concernée, ou une sous-estimation, via l'oubli de certains patients ou l'activité qui se trouve en dehors de la fenêtre oncologique. Néanmoins, ces biais sont supposés constants au fil du temps.

La présentation dans un outil de pilotage centralisé de l'évolution de l'activité des patients oncologiques et de la répartition par domaine apporte une aide aux décideurs dans les choix stratégiques et la gestion de l'Institut de Cancérologie du CHU de Liège, par exemple pour la répartition des ressources humaines transversales dans chaque secteur ou pour estimer l'activité attendue du nouveau bâtiment. L'outil permet également de réaliser certaines études de faisabilité pour les firmes pharmaceutiques, ou encore de faire du benchmarking avec d'autres centres. L'accès est ouvert depuis quelques mois au Conseil de Gouvernance de l'Institut de Cancérologie, constitué essentiellement de médecins et médecins chefs de service. Ils en sont satisfaits et l'ont utilisé à plusieurs reprises pour objectiver certaines questions. L'utilisation est simple et facile à prendre en main, notamment grâce aux différents filtres mis à disposition. Le Conseil de Gouvernance est demandeur de futurs développements dans l'outil, tels que l'ajout des données sur les chimiothérapies, les analyses de laboratoires ou encore les examens d'imagerie.

Les analystes du Service des Informations Médico-Economiques utilisent également régulièrement l'outil de pilotage des patients oncologiques. Cela leur permet de répondre rapidement à certaines demandes sans devoir repartir des données sources. Par ailleurs, l'équipe d'analystes peut également exploiter directement la base de données des patients oncologiques dans le cadre de demandes plus spécifiques, comme la collecte de données pour des études cliniques sur une pathologie bien précise ou encore la mise à disposition de données pour l'extérieur.

En conclusion, l'exploitation d'un entrepôt riche et varié de données cliniques par des équipes de Business Intelligence (BI) et d'analystes qualifiés, en étroite collaboration avec les utilisateurs finaux, a permis de

développer avec succès un outil de pilotage transversal dédié à l'activité des patients oncologiques du CHU de Liège.

Références

Abedian, S., Sholle, E. T., Adekkanattu, P. M., Cusick, M. M., Weiner, S. E., Shoag, J. E., Hu, J. C., & Champion, T. R., Jr (2021). Automated Extraction of Tumor Staging and Diagnosis Information From Surgical Pathology Reports. *JCO clinical cancer informatics*, 5, 1054–1061. <https://doi.org/10.1200/CCI.21.00065>

Jacques J., Thys M., Gangolf M., Cunin M.-P., Lambert F., Vanbrabant T., Poustovoitov A., Waseige E., Laruelle A.-S., Debouny M., Dupont A.-C., Kolh P (2020) Apport d'un Entrepôt de données cliniques dans la gestion d'un hôpital universitaire. 10ème conférence Francophone en Gestion et Ingénierie des Systèmes Hospitaliers, GISEH2020, Oct 2020, Valenciennes, France.

Si, Y., & Roberts, K. (2018). A Frame-Based NLP System for Cancer-Related Information Extraction. *AMIA ... Annual Symposium proceedings. AMIA Symposium*, 2018, 1524–1533.

Tyczynski J., Démaret E., Parkin D. (2003) Standards and guidelines for cancer registration in Europe:the ENCR recommendations. Lyon: IARC.