HAL
open science

Construction d'indicateurs de prise en charge en cancérologie à partir de l'entrepôt de données du CHU de Bordeaux. Application au cancer du poumon dans le contexte COVID-19

Benjamin Naffrechoux

To cite this version:

Benjamin Naffrechoux. Construction d'indicateurs de prise en charge en cancérologie à partir de l'entrepôt de données du CHU de Bordeaux. Application au cancer du poumon dans le contexte COVID-19. Sciences du Vivant [q-bio]. 2020. dumas-03108289

HAL Id: dumas-03108289

https://dumas.ccsd.cnrs.fr/dumas-03108289v1

Submitted on 13 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

U.F.R DES SCIENCES MEDICALES

Année 2020 Thèse n° 3275

THESE POUR L'OBTENTION DU

DIPLOME D'ETAT de DOCTEUR EN MEDECINE

Présentée et soutenue publiquement

Par Benjamin NAFFRECHOUX

Né le 24/11/1990 à Chambéry

Le 14/12/2020

Construction d'indicateurs de prise en charge en cancérologie à partir de l'entrepôt de données du CHU de Bordeaux

Application au cancer du poumon dans le contexte COVID-19

Sous la direction de :

Docteur Vianney Jouhet et Docteur Françoise Colombani

Membres du jury :

Professeur Rodolphe Thiebaut	Président
Docteur Sylvie Maurice	Rapportrice
Professeur Jacques Jougon	Juge
Docteur Frantz Thiessard	Juge
Docteur Renaud Trouette	Juge

Remerciements

A mes directeurs de thèse, Docteurs Vianney Jouhet et Françoise Colombani, pour m'avoir fait confiance pour ce travail ainsi que pour votre accompagnement et votre soutien sans faille tout au long de l'année. Ce fut un réel plaisir d'apprendre aux côtés de personnes avec de telles qualités tant humaines que professionnelles.

A l'ensemble de l'équipe de l'UIAM du CHU de Bordeaux, pour leur bonne humeur, leur aide et leur expertise dans ce domaine obscur qu'est l'informatique médical. La boucle est bouclée !

Au Docteur Sylvie Maurice, pour sa gentillesse, sa relecture minutieuse et ses nombreux conseils prodigués en un temps record.

Au Professeur Rodolphe Thiebaut ainsi qu'à l'ensemble des membres du jury pour l'honneur que vous me faites en ayant accepté d'évaluer mon travail.

A ma famille, toujours présente dans les bons comme dans les mauvais moments. Papa, maman, petite sœur, Papi, Grand-mère, Grand-Père, merci d'être là, vous êtes un soutien et une force incroyable. Ça y est Thomas, à mon tour d'être thèsé ! Et puis ça sera toi Marion, admiratif que tu trouves le courage de te lancer dans une thèse de science, mais en attendant, des travaux à finir pour un bel emménagement avec Greg !

A Fanny, pour avoir embelli et pimenté ces années à Bordeaux. Merci pour tous ces moments heureux, il y en aura beaucoup d'autres j'en suis sûr ! « L'amour ce n'est pas se regarder l'un l'autre, mais c'est regarder ensemble dans la même direction. » Antoine de Saint-Exupéry

A mes amis d'enfance et des années collège/lycée, Thibaut, Micka, Marion, Mériéla, Silouane, Flo, Aurélien, Gaël, Tom, et tous les autres.

A mes amis et collègues qui m'ont accompagné durant mon internat de médecine, Louis, Kevin, Marine, Mehdi & Mehdi, Lamine et tous ceux que j'oublie mais qui se reconnaitront.

Préambule

Au fil de mon parcours d'internat, j'ai découvert l'univers riche et varié de la santé publique. Un univers mêlant épidémiologie, politiques de santé, activités de prévention […], et articulé autour d'une variable indissociable de toutes ces activités : la donnée de santé. Après être passé dans des terrains de stage aussi différents qu'enrichissants au début de mon internat, j'ai logiquement orienté ma formation autour de deux domaines qui me tiennent à cœur : l'informatique médicale, mêlant intégration et réutilisation de ces données de santé si précieuses, et la cancérologie, fardeau de notre société pour lequel il reste tant de choses à faire pour améliorer le bien-être et la qualité de vie de la population.

Mon travail de thèse s'inscrit dans cette continuité, en mettant à profit mes compétences en informatique médicale acquises pendant mon année de Master (Spécialité Science de l'Information et Technologies Informatiques en Santé) pour un projet ambitieux de construction d'indicateurs de prise en charge en cancérologie dans le contexte de la pandémie Covid-19, en collaboration avec le Centre de Coordination en Cancérologie (3C) et l'unité d'Informatique et d'Archivistique médical (IAM) du service d'information médicale du CHU de Bordeaux.

Ce travail présente tout d'abord le contexte scientifique qui l'a motivé, puis les méthodes employées pour parvenir à décrire les parcours de soins de patients atteints de cancer du poumon grâce à des indicateurs construits de façon automatisée via l'entrepôt de données de santé du CHU, ainsi que les résultats de ces indicateurs qui sont discutés.

Table des matières

1 Introduction 1
1.1 Organisation de l'offre de soins en cancérologie en France 2
1.1.1 Schéma général 2
1.1.2 Centres de coordination en cancérologie (3C) 3
1.2 Le contexte de la pandémie de CoviD-19 3
1.2.1 Conséquences de l'épidémie sur la société et le système de santé 3
1.2.2 Réorganisation des soins en cancérologie en France 4
1.2.1 Le besoin d'évaluer les conséquences de la réorganisation 5
1.3 Projet proposé dans le cadre de cette thèse 5
1.3.1 Réutilisation secondaire des données de santé du CHU de Bordeaux 6
1.3.2 L'exemple du cancer du poumon 8
1.4 Objectif 9
2 Matériel 10
2.1 Source des données et terminologies 10
2.2 Outils informatiques 10
3 Méthodes 11
3.1 Identification des patients atteints de cancer primitif du poumon 11
3.1.1 Population cible 12
3.1.2 Règles de sélection des patients pris en charge pour cancer du poumon 12
3.1.3 Règles d'exclusion des patients atteints de cancer primitif multiple 13
3.1.4 Règles d'attribution d'une date de début de prise en charge 13
3.2 Construction des indicateurs 15
3.2.1 Identification de deux cohortes d'analyse 16
3.2.2 Stade TNM au diagnostic 17
3.2.3 Délai d'accès à la proposition thérapeutique en RCP 18
3.2.4 Délai d'accès au premier traitement 18

3.3 Analyses 19
4 Résultats 20
4.1 Patients identifiés et caractéristiques 20
4.2 Caractéristiques des patients par cohorte 25
4.3 Indicateurs 28
4.3.1 Stade TNM au diagnostic 28
4.3.2 Délai d'accès à la proposition thérapeutique en RCP 32
4.3.3 Délai d'accès au premier traitement 34
5 Discussion 36
5.1 Résultats principaux 36
5.2 Fiabilité de nos indicateurs et comparaison avec la littérature 37
5.3 Limites du travail 39
5.3.1 Identification des patients atteints de cancer du poumon 39
5.3.2 Méthode de construction des indicateurs 39
5.4 Perspectives 41
5.4.1 Evaluer et améliorer l'identification des patients 41
5.4.2 Reconstruire le parcours patient à l'échelle d'un établissement 42
6 Conclusion 43
7 Bibliographie 44

Table des tableaux

Tableau 1 : Cancer du poumon non à petites cellules, fréquence et survie par stade 9

Tableau 3 : Codes CIM10 (35) et ADICAP (37) de cancer primitif infiltrant des bronches ou du poumon 12

Tableau 4 : Codes « cancer » utilisés pour attribuer une date de début de prise en charge (correspondances entre les terminologies CIM10 (35) et ADICAP (37)).....14

Tableau 5 : Codes CIM10 (35) et CCAM (36) identifiant le traitement anticancéreux initial.19

Tableau 5: Caractéristiques des patients identifiés, par année de première prise en charge au CHU de Bordeaux 24

Tableau 7 : Caractéristiques des patients pris en charge au CHU de Bordeaux, par cohorte d'analyse 26

Tableau 8 : Caractéristiques de la RCP d'intérêt et du premier traitement réalisés au CHU de Bordeaux, par cohorte d'analyse, au 27/10/2020.....27

Tableau 9 : Délai en jours d'accès à la proposition thérapeutique en RCP pour les patients pris en charge au CHU de Bordeaux, par cohorte d'analyse.....32

Tableau 10 : Délai en jours d'accès au premier traitement pour les patients pris en charge au CHU de Bordeaux, par cohorte d'analyse 34

Table des figures

Figure 1 : Parcours de soins type d'un patient atteint de cancer en France.....1

Figure 2 : Organisation de la cancérologie en France 2

Figure 3 : Sources de données intégrées dans l'entrepôt des données de santé (EDS) du CHU de Bordeaux 7

Figure 4 : Etapes d'identification des cas de cancer primitif des bronches et du poumon dans l'entrepôt de données de santé du CHU Bordeaux 11

Figure 5 : Définition des indicateurs (stade TNM, délais de prise en charge) en vie réelle et à l'échelle de l'entrepôt de données.....15

Figure 6 : Diagramme de flux des patients identifiés *via* l'entrepôt de données de santé du CHU de Bordeaux 20

Figure 7 : Diagramme de Venn des patients inclus selon les critères et sources de données de l'entrepôt de données de santé du CHU de Bordeaux 21

Figure 8 : Nombre de patients atteints de cancer infiltrant du poumon identifiés via l'entrepôt de données de santé du CHU de Bordeaux 22

Figure 9 : Diagramme d'Euler des cohortes identifiées parmi l'ensemble des 2753 patients inclus, au CHU de Bordeaux.....25

Figure 10 : Evolution du stade TNM au diagnostic, des cas de cancer du poumon identifiés via l'EDS pour la cohorte de patients « avec RCP » 29

Figure 11 : Evolution du stade TNM au diagnostic, pour la cohorte de patients « avec RCP pré-thérapeutique ». 31

Figure 12 : Délai d'accès à la proposition thérapeutique en RCP, pour la cohorte de patients « avec RCP » 33

Figure 13 : Délai d'accès à la proposition thérapeutique en RCP, pour la cohorte de patients « avec RCP pré-thérapeutique » 33

Figure 14 : Délai d'accès au premier traitement, pour la cohorte de patients « avec RCP » ...35

Figure 15 : Délai d'accès au premier traitement, pour la cohorte de patients « avec RCP pré-thérapeutique » 35

Abréviations

ADICAP Association pour le développement de l'informatique en cytologie et anatomo-pathologie

BDD Base de données

CCAM Classification Commune des Actes Médicaux

3C Centres de coordination en cancérologie

CHU Centre Hospitalier Universitaire

CIM-10 Classification Internationale des Maladies, 10ème révision

CIRC Centre International de Recherche sur le Cancer

CRAP Compte-Rendu d'Anatomo-cyto-Pathologie

DAD Diagnostic Associé Documentaire

DAS Diagnostic Associé Significatif

DP Diagnostic Principal

DPI Dossier Patient Informatisé

DR Diagnostic Relié

EDS Entrepôt de Données de Santé

HCSP Haut Conseil de Santé Publique

HDJ Hôpital De Jour

IDEC Infirmiers coordinateurs

INCa Institut National du Cancer

IPAQSS Indicateurs pour l'Amélioration de la Qualité et de la Sécurité des Soins

ORSAN Dispositif d'organisation de la réponse du système de santé en situations sanitaires exceptionnelles

ORSAN REB Dispositif d'organisation de la réponse du système de santé en situations sanitaires exceptionnelles « Risques Epidémiologiques et Biologiques »

PMSI Programme de Médicalisation du Système d'Information

PPS Plan Personnalisé de Soins

RCP	Réunions de Concertation Pluridisciplinaires
RRC	Réseau Régional de Cancérologie
RUM	Résumé d'Unité Médicale
SIH	Système d'Information Hospitalier
TAS	Temps d'Accompagnement Soignant
TNM	Classification Tumeur-Node-Métastase

1 Introduction

Le cancer constitue la deuxième cause de mortalité dans le monde, soit un décès sur six (1). En France, cette pathologie est même la première cause de décès chez les hommes et la première cause de décès prématurés (avant 65 ans) tout sexe confondu (2). Ainsi, la lutte contre le cancer fait l'objet d'une mobilisation nationale, traduite dans les Plans Cancer depuis 2003. L'un des objectifs majeurs, présent dans chacun de ces plans, est de garantir à tous les patients une prise en charge sécurisée et de qualité. Pour répondre à cet objectif de qualité de prise en charge, le plan cancer 2014-2019 (3) prévoit, dans son objectif 2, d'améliorer le parcours de soins du patient dès le diagnostic initial de la maladie, grâce à plusieurs actions fortes :

- Réduire les délais de prise en charge et d'accès aux examens
- Adapter la prise en charge des personnes âgées, des enfants et des jeunes adultes
- Poursuivre et faire évoluer la structuration de l'offre de soins, mise en place via les circulaires de 2005 relative à l'organisation de la cancérologie (4) et de 2007 relative aux réseaux de cancérologie (5).

La Figure 1 décrit les étapes du parcours de soins type d'un patient atteint de cancer en France, en cinq temps : le diagnostic de la maladie, l'annonce de ce diagnostic, un temps d'accompagnement, une prise en charge thérapeutique adaptée et personnalisée et enfin la prise en charge de l'après-cancer. Pour assurer la continuité de cette prise en charge, un lien ville-hôpital est mis en place dès le début du parcours du patient.

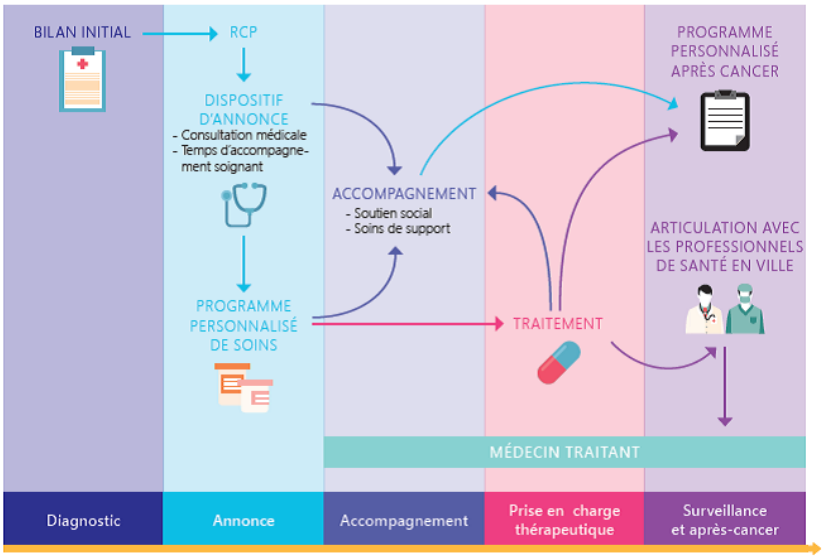


Figure 1 : Parcours de soins type d'un patient atteint de cancer en France
(source : Réseau régional de cancérologie du Centre-Val de Loire)

1.1 Organisation de l'offre de soins en cancérologie en France

1.1.1 Schéma général

La structuration de l'offre de soins en cancérologie en France s'articule autour d'un dispositif d'autorisation de traitement du cancer (Figure 2), mis en place depuis 2009 afin de garantir à tous les patients, quelle que soit leur localisation géographique, des soins de qualité et un accès à l'innovation. Ces autorisations sont délivrées par les Agences régionales de santé (ARS) aux établissements publics et privés de santé pour une durée de cinq ans, sous réserve de respecter trois mesures socles définies en 2007 :

- Six conditions transversales de qualité (mise en place du dispositif d'annonce, organisation des réunions de concertation pluridisciplinaires (RCP), remise aux patients de plans personnalisés de soins (PPS), suivi des référentiels de bonnes pratiques, facilitation de l'accès pour la personne malade aux soins oncologiques de support ainsi qu'aux traitements innovants et aux essais cliniques).
- Critères d'agrément définis par l'Institut National du Cancer (INCa) pour les principales thérapeutiques du cancer (chirurgie, radiothérapie, chimiothérapie)
- Un seuil d'activité minimale, distinct pour chacune des trois pratiques thérapeutiques

Les établissements doivent être également membre d'un réseau régional de cancérologie (RRC), dont le rôle principal est de coordonner les acteurs ayant une activité en cancérologie et d'assurer ainsi la continuité des soins au patient. Au niveau local, cette coordination des soins s'appuie sur les centres de coordination en cancérologie (3C).

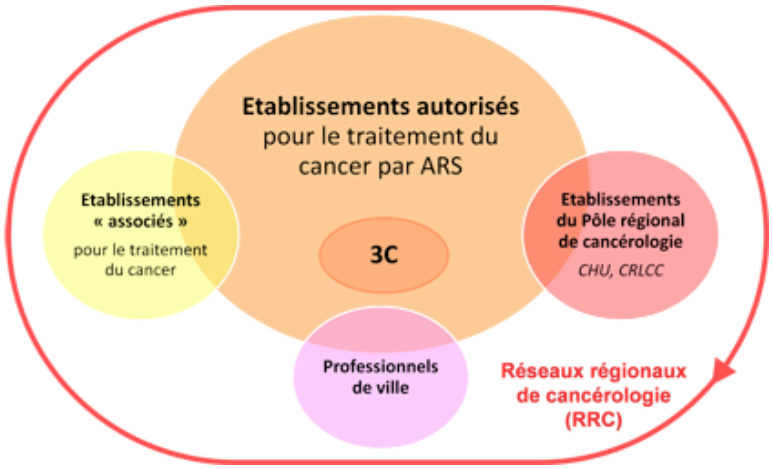


Figure 2 : Organisation de la cancérologie en France (source : 3C CHU de Bordeaux, (4))

1.1.2 Centres de coordination en cancérologie (3C)

Les 3C, véritables cellules transversales de qualité implantées au sein d'un ou plusieurs établissements de santé, ont pour mission de structurer et d'évaluer la cancérologie. Ils font partie intégrante des projets d'établissement et assurent des missions obligatoires centrées sur la qualité, la coordination et la continuité du processus de prise en charge du patient, le respect de ses droits et son association aux décisions qui le concernent ainsi que sur l'évaluation de la qualité des soins et des pratiques professionnelles (4), *via* la mise en œuvre opérationnelle et le respect des conditions transversales de qualité exigibles dans le cadre des autorisations de traitement du cancer. Pour cela, les 3C évaluent l'activité des établissements qui en sont membres à l'aide de rapports d'activité et d'une démarche qualité reposant sur des méthodes d'évaluation des pratiques professionnelles (exemples : méthode du patient traceur, audits), ou encore de mesure d'indicateurs pour l'Amélioration de la Qualité et de la Sécurité des Soins (IPAQSS). Parmi eux, l'indicateur IPAQSS-RCP évalue l'exhaustivité et les modalités d'organisation de la RCP lors de la prise en charge initiale d'un patient atteint de cancer. Ce sont donc les 3C qui garantissent, à l'échelle d'un établissement, l'accompagnement de la bonne mise en œuvre d'une prise en charge de qualité pour chaque patient atteint de cancer, en s'assurant que les étapes cruciales du parcours de soins soient respectées en toutes circonstances.

1.2 Le contexte de la pandémie de CoviD-19

1.2.1 Conséquences de l'épidémie sur la société et le système de santé

La pandémie Covid-19, provoquée par le coronavirus SARS-CoV-2, a débuté entre octobre et décembre 2019 à Wuhan en Chine avant de se propager dans le monde. En France, les premiers cas sont apparus en janvier 2020, entrainant la mise en place de mesures nationales de prévention exceptionnelles, basées sur le dispositif d' « organisation de la réponse du système de santé en situations sanitaires exceptionnelles » ORSAN et le nouveau volet introduit en 2019 ORSAN REB (« risques épidémiologiques et biologiques »), pour lutter contre la propagation du virus : promotion des gestes barrières, premier confinement de la population du 17 Mars au 10 Mai 2020 et réorientation des activités sanitaires vers la gestion des cas CoviD-19.

L'offre de soins s'est alors réorganisée dans l'ensemble des secteurs sanitaires (ambulatoire, hospitalier et médico-social), bouleversant les parcours de soins des patients atteints de pathologies hors CoviD-19 et notamment des patients atteints de cancer : report des interventions chirurgicales (6,7), redéploiement des ressources humaines et matérielles dans les services CoviD-19 (8),…

1.2.2 Réorganisation des soins en cancérologie en France

En France, la réorganisation de l'offre de soins en cancérologie s'est appuyée sur de multiples recommandations à la fois nationales (avis du Haut Conseil de Santé Publique (HCSP) (9), préconisations de l'INCa (10), recommandations françaises publiées dans le Lancet (11)) régionales et locales.

Au niveau régional, le RRC de Nouvelle-Aquitaine a publié sur son site Internet plusieurs recommandations pour la réorganisation des soins dans le contexte du Covid-19 (12).

Au niveau local, au CHU de Bordeaux, de nombreuses réorganisations de terrain ont été mises en œuvre au niveau des services de chirurgie (principalement des déprogrammations de blocs opératoires dues à la réorganisation des services de réanimation et de soins continus post-chirurgicaux), des services médicaux prenant en charge des patients atteints de cancer dont l'oncologie médicale, la radiothérapie, l'hématologie, et autres services dits « d' organes » (pneumologie, dermatologie, digestif...). Les principaux types de réorganisations au niveau des secteurs médicaux étaient : la sanctuarisation des secteurs « sans CoviD-19 » avec notamment la limitation des accompagnants et des visites des accompagnants ; la limitation des consultations en présentiel pour les suivis ; la délocalisation des prises en charge en hôpital de jour (HDJ) pour limiter le contact avec les secteurs conventionnels ; la vérification de la température et un questionnaire CoviD-19 des patients à J-1 avant les hospitalisations.

En outre, des modifications de protocoles de traitement ont été discutées au sein des RCP : certains patients ont été traités par stéréotaxie plutôt qu'en chirurgie ; espacement des cures d'immunothérapie ou de chimiothérapie. La réorganisation des RCP en cancérologie s'est appuyée notamment sur le travail du 3C, dont les deux objectifs prioritaires ont été de (a) sécuriser les équipes médicales en séance RCP pour éviter leur exposition au SARS-CoV2 et ainsi assurer la continuité des soins, (b) sécuriser les phases cruciales du parcours de soins en cancérologie, notamment le maintien des RCP (passage en visio-conférence pour 75% des RCP du CHU et 25% en présentiel réduit) et le maintien de la continuité du parcours (temps d'accompagnement soignant (TAS) par l'équipe infirmiers du 3C, maintien du dispositif des infirmiers de coordination en cancérologie (IDEC) et intervention des psychologues auprès des patients mais également des équipes soignantes.

1.2.1 Le besoin d'évaluer les conséquences de la réorganisation

Comme nous l'avons vu, la structuration de l'offre de soins en cancérologie a pour objectif de garantir aux patients une prise en charge sécurisée et de qualité. Dans ce contexte exceptionnel, il est possible que la réorganisation rapide des structures sanitaires, notamment hospitalières, ait pu mettre en défaut cet objectif de qualité et entrainé une perte de chance pour les patients atteints de cancer (13–20). Il semble alors important de développer des outils pour décrire les parcours de soins et mesurer les éventuelles conséquences, en particulier en termes de délais, de cette réorganisation sur la prise en charge des patients atteints de cancer.

1.3 Projet proposé dans le cadre de cette thèse

Pour ce projet, nous avons fait deux choix : (a) utiliser la masse de données présente dans l'entrepôt de données de santé (EDS) du CHU de Bordeaux, avec l'hypothèse sous-jacente que la réutilisation secondaire de ces données nous permettra de décrire les parcours de soins en cancérologie et mesurer leurs éventuelles variations durant la pandémie Covid-19, (b) cibler les patients atteints de cancer du poumon, du fait de la volonté médicale de cette filière de spécialité de travailler sur cette localisation dont le pronostic est influencé par les délais de pris en charge (16).

1.3.1 Réutilisation secondaire des données de santé du CHU de Bordeaux

L' « utilisation secondaire de données » représente l'usage de données dans un cadre différent de celui pour lesquelles elles ont été acquises ou produites. Ainsi, toute utilisation des données de santé d'un patient en dehors du cadre du soin (par exemple à des fins de recherche, de surveillance épidémiologique, de tarification,…) correspond à une réutilisation secondaire de ses données (21,22). A l'échelle d'un établissement de santé, ces données sont contenues dans le système d'information hospitalier (SIH).

a) Le système d'information hospitalier : lieu de stockage des données

Un SIH est un système informatique destiné à faciliter la gestion de l'ensemble des informations médicales et administratives d'un hôpital. Son architecture est découpée en différentes couches permettant d'afficher et de stocker les données de santé d'un patient. Les données stockées au sein d'un SIH sont hétérogènes. On en distingue classiquement trois grands types :

- Les données non structurées, comme les données textuelles (comptes rendus médicaux) ou les images (exemple : radiographies). Ce sont les données de santé les plus fréquentes dans le système d'information d'un établissement de santé (22), mais également les plus difficiles à exploiter,
- Les données semi-structurées, notamment les formulaires du dossier patient informatisé (DPI). Ces données sont également difficiles à réutiliser car elles ne sont pas standardisées : d'un service à l'autre, un même concept peut être défini différemment (exemples : sexe, genre) et représenté différemment (exemples : H/F, 0/1),
- Les données structurées et codées selon un référentiel (terminologie, ontologie, …) international, national ou local permettant de leur apporter une sémantique et ainsi faciliter leur réutilisation. Par exemple, au sein du programme de médicalisation du système d'information (PMSI), les diagnostics des patients sont codés selon la terminologie CIM-10 (Classification Internationale des Maladies, 10ème révision) et les actes médicaux sont codés avec la CCAM (Classification Commune des Actes Médicaux).

Les données sont également dispersées entre les différentes applications (logiciels métiers) du SIH. Au CHU de Bordeaux, nous pouvons par exemple citer DxCare®, dédié au dossier patient informatisé (DPI), Synergy® pour les données de biologie, Xplore® pour les données d'imagerie, ou encore Pharma® pour la pharmacie hospitalière. Cette dispersion des données est un frein à une réutilisation secondaire optimale de celles-ci.

b) L'entrepôt de données de santé : une architecture adaptée à la recherche

L'architecture d'un SIH est optimisée pour l'utilisation primaire des données de santé, c'est-à-dire dans le cadre du soin. A l'inverse, l'architecture d'un entrepôt de données (EDS) (*datawarehouse*) a été pensée de façon à faciliter la réutilisation des données acquises à chaque prise en charge d'un patient, notamment en utilisant une seule base de données, permettant ainsi d'éviter la dispersion des données. Au CHU de Bordeaux, c'est l'unité Informatique et archivistique médicale (IAM) du Service d'information médicale (SIM) qui est en charge de la mise en œuvre opérationnelle de l'entrepôt de données de santé (EDS) de l'établissement. Celui-ci se base sur la solution *open source* i2b2, construit autour d'un modèle de stockage dit « en étoile » (23). La Figure 3 représente les différentes sources de données du SIH intégrées dans l'EDS du CHU de Bordeaux. Nous retrouvons par exemple les données du PMSI (diagnostics CIM10 et actes CCAM), les données d'anatomo-pathologie (provenant de Diamic® et de la tumorothèque) ou encore les données issues des dossiers de spécialités de DxCare® (notamment les formulaires). L'EDS du CHU de Bordeaux intègre tout l'historique des données de l'ensemble des patients venus depuis 2010. Cela correspondant à plus de 1,5 millions de patients, près de 14 millions de venues et 1,9 milliards d'observations.

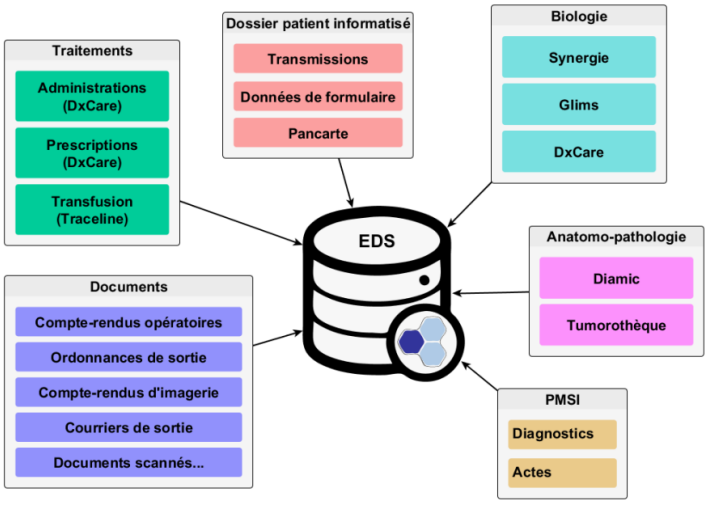


Figure 3 : Sources de données intégrées dans l'entrepôt des données de santé (EDS) du CHU de Bordeaux (source : IAM, CHU Bordeaux)

1.3.2 L'exemple du cancer du poumon

Nous avons fait le choix d'étudier une localisation cancéreuse, le cancer du poumon, pour plusieurs raisons :

- La fréquence de ce cancer au sein de la population générale. En effet, le cancer du poumon est le deuxième cancer le plus fréquent chez l'homme (parmi les tumeurs solides) et le troisième chez la femme (2). Il est également au premier rang des décès par cancer chez l'homme et au second rang chez la femme. En 2018, nous estimons à 46 363 le nombre de nouveaux cas de cancer du poumon en France métropolitaine, pour 33 117 décès (2). Une prise en charge optimale des patients atteints par cette pathologie apparaît donc comme un enjeu majeur de santé publique.
- La complexité des données utilisées. En effet, il existe différents freins à la réutilisation des données de vie réelle d'un patient à partir d'un EDS, comme la gestion de la temporalité de ces données ou encore la difficulté de rattacher une donnée (codée ou non) à une pathologie. Par exemple, lorsqu'un patient présente plusieurs tumeurs malignes, il peut être difficile de rattacher de façon certaine la chimiothérapie réalisée au temps t à l'une ou l'autre des tumeurs. En ciblant l'étude d'un seul cancer, le cancer du poumon dans notre cas, ces freins deviennent moins susceptibles de compromettre la faisabilité du projet.
- Enfin, la présence de facteurs pronostics identifiés comme prioritaires de la prise en charge des patients atteints de cancer du poumon par les cliniciens du CHU de Bordeaux et la littérature : le stade TNM au diagnostic (24,25) et les délais de prise en charge (16,26–32).

Le stade TNM au diagnostic correspond au stade calculé avant toute intervention thérapeutique, après un bilan d'extension complet réalisé à l'aide d'examens d'imagerie, d'une fibroscopie bronchique et éventuellement d'une confirmation histologique (par ponctions et/ou biopsies). Lorsqu'une confirmation histologique est réalisée, on parle de stade pathologique, sinon de stade clinique. On parle également de stade combiné lorsque des informations à la fois histologiques et cliniques sont utilisées pour le calcul. C'est un indicateur fortement corrélé à la survie des patients atteints de cancer du poumon, comme le montre le Tableau 1 à propos des cancers du poumon non à petites cellules.

Tableau 1 : Cancer du poumon non à petites cellules, fréquence et survie par stade (source : INCa, 2010 (24))

Diagnostic au stade :	**Fréquence**	**Survie relative à 5 ans**
Cancer localisé (Stades I et II)	15 à 30%	52,6%
Cancer localement avancé (Stade III)	20%	23,7%
Cancer métastatique (Stade IV)	40 à 55%	3,8%

Le délai de prise en charge des patients atteints de cancer du poumon est également fortement corrélé à la survie de ces patients (16,26–32). Il est de plus susceptible d'avoir été allongé par la réorganisation des soins lors de la pandémie de Covid-19. Il est classique de détailler ce délai en sous-délais, par exemple le délai d'accès à la proposition thérapeutique en RCP ainsi que le délai d'accès au premier traitement. Bien que ces sous-délais tendent à mesurer les mêmes évènements dans le parcours de soins d'un patient, leurs définitions peuvent varier selon les études (33,34). Ainsi, nous proposerons en méthode une définition propre à ce projet pour chacun des deux sous-délais.

1.4 Objectif

L'objectif principal de ce projet de thèse était de construire des indicateurs de prise en charge des patients atteints de cancer du poumon à partir de l'EDS du CHU de Bordeaux, pour à la fois pouvoir décrire les parcours de soins des patients pendant la période de crise sanitaire et établir un suivi dans le temps de ces indicateurs. Cet objectif sera poursuivi en deux étapes :

a) Identifier les patients atteints de cancer du poumon et pris en charge au CHU
b) Définir, développer et mettre en œuvre des indicateurs fiables et reproductibles

2 Matériel

2.1 Source des données et terminologies

Les données provenaient de l'EDS du CHU de Bordeaux, stockées dans une base de données (BDD) Oracle®. Ce dernier regroupe les données de l'ensemble des établissements rattachés au CHU de Bordeaux, soit le groupe hospitalier Pellegrin, le groupe hospitalier Sud (hôpital Haut-Lévêque et hôpital Xavier-Arnozan) et le groupe hospitalier Saint-André.

Pour ce projet, nous avons essentiellement utilisé des données structurées au sein de l'EDS, soit codées à l'aide de terminologies médicales, soit semi-structurées par des champs de formulaires (par exemple le formulaire des RCP). Les terminologies utilisées étaient la Classification Internationale des Maladies dans sa version 10 (CIM-10 (35)), la Classification Commune des Actes Médicaux (CCAM (36)) ainsi que la terminologie de l'Association pour le Développement de l'Informatique en Cytologie et Anatomo-Pathologie (ADICAP (37)) permettant d'attribuer un code résumant l'information d'un compte-rendu anatomo-cyto-pathologique (CRAP). Pour chacune de ces terminologies, des exemples de codes associés à leur libellé sont présentés en annexe (annexes 2, 3 et 4).

2.2 Outils informatiques

Les requêtes de la base de données de l'EDS et les analyses ont été réalisées à l'aide du logiciel R version 4.0.1. Le modèle relationnel de la BDD de l'EDS est présenté en annexe (annexe 1).

3 Méthodes

3.1 Identification des patients atteints de cancer primitif du poumon

Pour construire nos indicateurs de prise en charge, il était nécessaire d'identifier les patients (ou cas) atteints de cancer primitif invasif des bronches et du poumon et de leur attribuer une date de *primo* prise en charge. Pour cela, nous avons utilisé les données structurées codées présentes dans l'ensemble des dossiers patients enregistrés dans l'EDS du CHU de Bordeaux puis nous avons développé un algorithme basé sur des règles de décision. Les étapes de cet algorithme sont détaillées ci-dessous et résumées par la Figure 4.

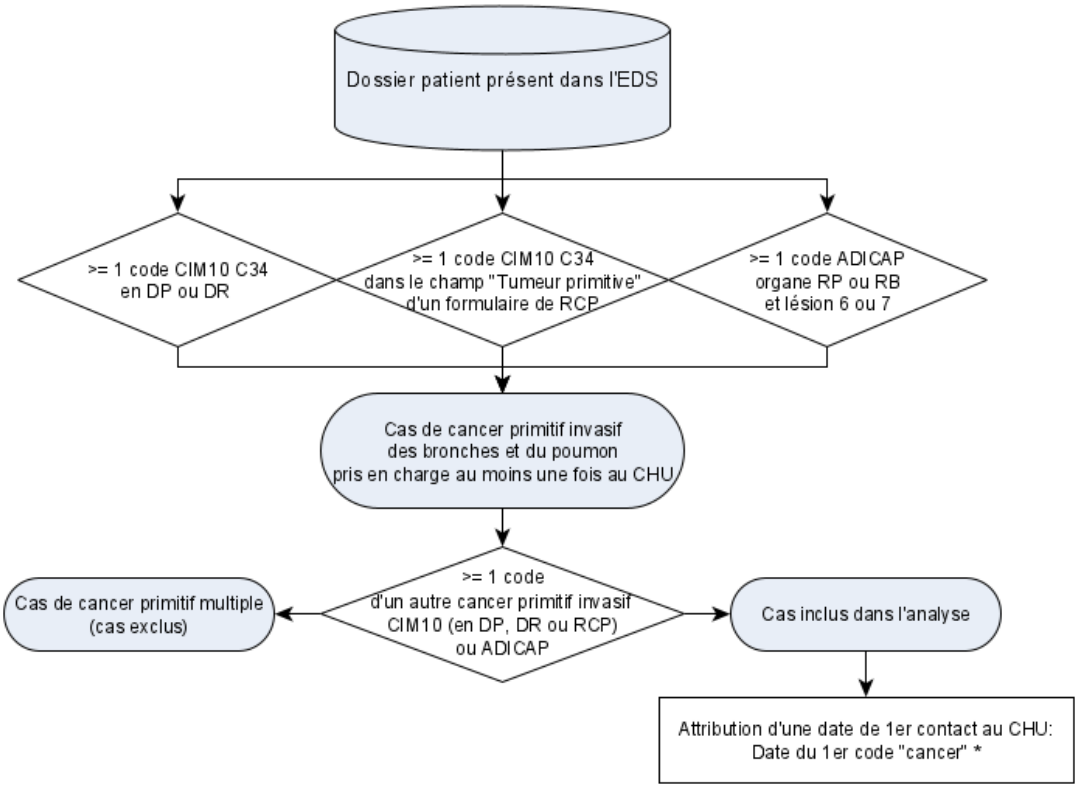


Figure 4 : Etapes d'identification des cas de cancer primitif des bronches et du poumon dans l'entrepôt de données de santé du CHU Bordeaux.

*date de la 1ère donnée codée relative à un cancer selon la méthode détaillée dans le paragraphe 3.1.4

Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

3.1.1 Population cible

Les patients que l'on cherchait à identifier correspondaient aux patients atteints d'une tumeur maligne primitive infiltrante des bronches ou du poumon et pris en charge au CHU de Bordeaux dans le cadre d'un premier cancer. A l'échelle du CHU, cela consistait à identifier à l'aide du SIH la première prise en charge du patient relative à un diagnostic de cancer du poumon.

Nous avons fait le choix d'inclure uniquement les tumeurs infiltrantes et d'exclure par conséquent les tumeurs *in situ* (qui ne relèvent pas de la même prise en charge). En outre, nous avons choisi de ne pas inclure les patients atteints de plusieurs tumeurs malignes primitives infiltrantes (tumeurs primitives multiples), pour deux raisons : (a) dans le système d'information hospitalier, il était parfois difficile de faire correspondre de façon exacte les données de prise en charge (RCP, actes thérapeutiques,…) à l'une ou l'autre des tumeurs du patient ; (b) il a été montré dans la littérature que ces patients avaient une surveillance plus active que les patients nouvellement diagnostiqués (38,39), ce qui aurait pu modifier l'interprétation de nos résultats en incluant dans notre population des patients aux parcours de soins différents. Pour suivre l'évolution de nos indicateurs dans le temps, nous avons fixé une période d'analyse : celle-ci s'étendait du 01 janvier 2017 au 31 juillet 2020, c'est-à-dire que les patients inclus étaient les patients pour lesquels la date de *primo* prise en charge au CHU pour cancer du poumon était comprise dans cette période d'analyse.

3.1.2 Règles de sélection des patients pris en charge pour cancer du poumon

Soit un dossier patient. Pour que le patient soit identifié par l'algorithme, son dossier devait contenir au moins un code CIM10 ou ADICAP relatif à un cancer primitif infiltrant des bronches ou du poumon (Tableau 2). De plus, les codes CIM10 devaient être en position de diagnostic principal (DP), de diagnostic relié (DR) ou dans le champ « tumeur primitive » d'un formulaire de réunion de concertation pluridisciplinaire (RCP) de cancérologie.

Tableau 2 : Codes CIM10 (35) et ADICAP (37) de cancer primitif infiltrant des bronches ou du poumon

Libellés	**Code CIM10**	**Code ADICAP**	
		Code organe	**Code tumeur**
Cancer primitif invasif des bronches et du poumon	C34	RB RP	.6.. ou .7..

Pour un code ADICAP, un « . » renvoie à n'importe quel caractère alphanumérique

3.1.3 Règles d'exclusion des patients atteints de cancer primitif multiple

Le Centre international de recherche sur le cancer (CIRC) définit les cancers primitifs multiples comme deux ou plus cancers primitifs survenant chez un même individu, indépendamment du temps, un cancer primitif n'étant ni une extension, ni une récurrence ni une métastase (40). A l'échelle du CHU de Bordeaux, notre définition correspondait à des patients pris en charge au CHU pour au moins deux tumeurs malignes primitives infiltrantes de localisation différente, sans intervalle de temps minimum entre les prises en charge.

Pour exclure ces patients, nous avons utilisé des règles similaires à l'étape précédente (paragraphe 3.1.2) pour fouiller le dossier des patients identifiés lors de cette première étape. Ainsi, si un patient possédait dans son dossier un code CIM10 de tumeur maligne primitive infiltrante différent de C34 (à l'exception des codes C76.1 (« cancer primitif du thorax, SAI »), C77 à C79 (« métastases ») et C80 (« cancer primitif de siège non précisé »)) en DP, DR ou dans le champ « tumeur primitive » d'un formulaire de RCP, ou un code ADICAP de tumeur maligne primitive invasive avec un code organe différent de RB ou RP (« bronches et poumon »), RZ (« appareil respiratoire »), XT (« région thoracique »), XX ou XZ (« siège non précisé ») et SG (« ganglions lymphatiques »), alors celui-ci était exclu de notre population d'étude.

3.1.4 Règles d'attribution d'une date de début de prise en charge

Pour estimer la date de début de prise en charge au CHU, nous avons utilisé la date de la première donnée codée présente dans le dossier patient et dont le libellé renvoie à la notion de cancer. Ces codes étaient en premier lieu les codes CIM10 et ADICAP relatifs à un cancer primitif invasif du poumon (soit C34 en CIM10, RP et RB associés à un code tumeur 6 ou 7 en ADICAP), mais également l'ensemble des codes de cancer primitif dont le siège est mal ou non précisé, de métastases, d'évolution imprévisible ou inconnue, ou encore les codes indiquant un cancer dont la nature primitive ou secondaire est incertaine (Tableau 3).

Pour les données du PMSI (c'est-à-dire les DP, DR, DAS et DAD), la date attribuée correspondait à la date de début de résumé d'unité médicale (RUM). Pour les RCP, celle-ci était la date exacte de la tenue de la réunion RCP, et pour les données des autres formulaires et de prélèvement, la date de début de prise en charge correspondait à la date d'enregistrement de la donnée dans l'entrepôt, soit la date de remplissage du formulaire par les secrétaires des services concernés ou la date de rédaction du compte-rendu anatomo-pathologique.

Tableau 3 : Codes « cancer » utilisés pour attribuer une date de début de prise en charge (correspondances entre les terminologies CIM10 (35) et ADICAP (37))

Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

Libellés	**Codes CIM10**	**Codes ADICAP**	
		Code organe	**Code tumeur**
Cancer primitif invasif des bronches et du poumon	C34	RB, RP	.6.. ou .7..
Cancer primitif invasif du thorax, SAI (sans autre indication)	C76.1	RZ, XT	
Cancer primitif invasif de siège non précisé	C80	XX, XZ	
Cancer primitif invasif des ganglions lymphatiques		SG	
Extension secondaire ou métastases	C77-C79	Tous	.E.. ou .M..
Evolution imprévisible ou inconnue (poumon, appareil respiratoire)	D38.1, D38.6		
Nature primitive ou secondaire incertaine		Tous	.9..
Mise en observation pour suspicion de tumeur maligne	Z03.1		

Pour un code ADICAP, un « . » renvoie à n'importe quel caractère alphanumérique

Au final, les patients inclus dans notre analyse étaient ceux pour lesquels la date estimée de *primo* prise en charge au CHU était comprise entre le 01/01/2017 et le 31/07/2020.

3.2 Construction des indicateurs

Les choix retenus pour construire nos indicateurs étaient guidés par notre objectif de mesurer les conséquences éventuelles sur les parcours patient de la pandémie de Covid-19 et reposaient sur trois points clés : (a) développer une méthode simple et reproductible, (b) évaluer nos résultats au cours du temps, (c) tenir compte des données récentes de 2020. Pour cela, nous avons identifié et défini à l'échelle de l'entrepôt de données (EDS) les étapes clés nécessaires à l'estimation du stade TNM au diagnostic, du délai d'accès à la première proposition thérapeutique en RCP et du délai d'accès au premier traitement anti-cancéreux (Figure 5).

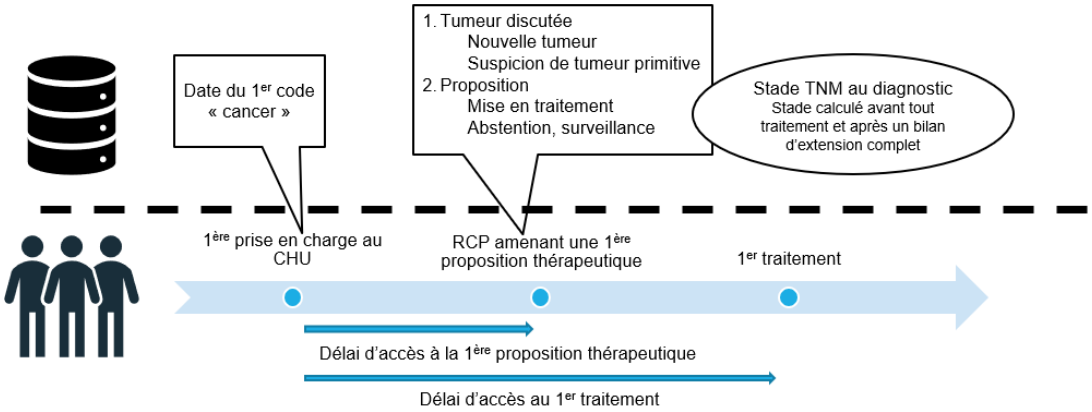


Figure 5 : Définition des indicateurs (stade TNM, délais de prise en charge) en vie réelle et à l'échelle de l'entrepôt de données
la RCP correspond ici à la RCP dite « pré-thérapeutique », sa définition à l'échelle de l'EDS peut varier (paragraphe 3.2.1)
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

3.2.1 Identification de deux cohortes d'analyse

Les indicateurs que l'on a cherché à construire correspondaient à une étape bien précise du parcours patient : la prise en charge initiale de son cancer. Il était alors nécessaire d'identifier à partir de l'EDS une donnée fiable permettant de se repérer dans le temps et d'estimer l'étape du parcours patient que l'on analysait. Pour cela, nous avons fait l'hypothèse que cette donnée fiable correspondait en vie réelle à la RCP discutant d'un nouveau cancer et aboutissant à une proposition thérapeutique.

Pour identifier cette RCP, il a ensuite fallu la définir au niveau de l'EDS, en gardant à l'esprit que nos indicateurs devaient être simples et reproductibles. Nous avons alors choisi d'utiliser deux définitions, une définition simple et une définition plus complexe :

a) Cette RCP correspondait à la première RCP du patient réalisée au CHU dans le cadre de son cancer.
b) Cette RCP correspondait à la première RCP du patient dont le champ du formulaire <Proposition de la RCP> prenait pour valeur « Mise en traitement », « Surveillance active » ou « Abstention thérapeutique » et dont le champ <Tumeur discutée> prenait pour valeur « Nouvelle tumeur » ou « Suspicion de tumeur primitive ». Par abus de langage, nous avons nommé la RCP respectant ces critères « *RCP pré-thérapeutique* »

Au final, à l'échelle du patient, cela permettait de construire deux cohortes au sein de notre population d'analyse : (a) une cohorte de patients avec au moins une RCP réalisée au CHU, appelée « *cohorte avec RCP* », et (b) une cohorte de patients avec au moins une RCP dite « *pré-thérapeutique* » réalisée au CHU, appelée « *cohorte avec RCP pré-thérapeutique* », sous-cohorte de la cohorte (a).

Ainsi, nos trois indicateurs ont été construits pour chacune de ces deux cohortes.

3.2.2 Stade TNM au diagnostic

Dans l'EDS du CHU, les seules données structurées (donc facilement exploitables) indiquant le stade de la maladie étaient les champs pT, pN, cT, cN et M des formulaires de RCP (p pour pathologique, c pour clinique, et M pour indiquer la présence ou non de métastases à distance). Nous avons alors reconstruit le stade TNM complet à l'aide de ces champs, en associant les informations du stade clinique (cT, cN) et du stade pathologique (pT, pN), soit un stade TNM combiné. Lorsque les deux informations étaient présentes, le stade pathologique primait sur le stade clinique.

Pour compléter les stades manquants, nous avons utilisé d'autres données :

- Le champ « Extension » des RCP, lorsque celui-ci prenait la valeur « A distance », renseignait un stade IV (métastatique).
- Les données textuelles des RCP. A l'aide d'une expression régulière[1], nous avons pu extraire les stades TNM écrits dans le corps de texte des RCP et ainsi compléter les cas pour lesquels le stade n'était pas renseigné dans les champs précédemment cités.
- Enfin, les codes CIM10 et ADICAP de métastase (sauf C77.1 qui code un envahissement ganglionnaire intrathoracique). Lorsqu'un de ces codes étaient présents dans le dossier patient, nous avons attribué un stade IV à toutes les RCP du patient postérieures à la date d'enregistrement du code dans l'entrepôt (soit la date du RUM dans lequel ce code a été produit).

Ainsi, ces différentes données nous ont permis d'attribuer un stade TNM complet à chacune des RCP de nos patients. Lorsque, pour une même RCP, plusieurs stades étaient attribués, nous avons utilisé la donnée la plus fiable de notre point de vue, soit par ordre décroissant de fiabilité : les champs <T, N et M>, le champ « Extension », les codes CIM10 et enfin les extractions faites à partir des données textuelles des RCP. Au final, un patient possédait entre 0 et n stades TNM, et pour chacun de ces stades, la date de calcul correspondait à la date de la RCP dont étaient extraites les informations nécessaires à ce calcul.

[1] Méthode de traitement automatique de la langue, basée sur des règles permettant d'extraire une chaine de caractères d'un texte

Puis, pour identifier le stade TNM au diagnostic parmi ces stades construits à un temps différent dans le parcours patient, nous avons utilisé la RCP définie au paragraphe 3.2.1. Cette RCP, et donc le stade défini comme étant le stade au diagnostic à l'échelle de l'EDS, était différent selon la cohorte d'analyse : soit le stade attribué à la première RCP du patient pour la cohorte (a), soit le stade attribué à la « *RCP pré-thérapeutique* » pour la cohorte (b). Ainsi, chaque patient possédait, pour les deux cohortes, entre 0 et un stade TNM défini comme étant le stade TNM au diagnostic à l'échelle de l'EDS du CHU de Bordeaux. Nous avons alors classé ces stades dans une des catégories du Tableau 1 (localisé, localement avancé, métastatique) ou dans la catégorie « stade inconnu » puis nous avons calculé une proportion pour chacune de ces classes (« classe x » dans l'équation 1) et chaque cohorte selon l'Équation 1 :

$$P = \frac{Nombre\ de\ patients\ avec\ un\ stade\ de\ la\ classe\ x}{Total\ des\ patients\ de\ la\ cohorte}$$

Équation 1 : Formule de l'indicateur « Stade au diagnostic ».
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

3.2.3 Délai d'accès à la proposition thérapeutique en RCP

A l'échelle de l'EDS, nous avons défini le délai d'accès à la proposition thérapeutique en RCP comme étant le délai entre la date de la première prise en charge du patient au CHU dans le cadre de son cancer (ou t0) et la date de la RCP aboutissant à une première proposition thérapeutique (ou t1). Pour la cohorte de patients « avec RCP », t1 correspondait à la date de la première RCP du patient au CHU, et pour la cohorte de patients « *avec RCP pré-thérapeutique* », t1 correspondait à la date de la première RCP « *pré-thérapeutique* ».

Ainsi, pour chacune des cohortes, nous avons calculé un délai moyen et médian en jours d'accès à la proposition thérapeutique en RCP.

3.2.4 Délai d'accès au premier traitement

Enfin, nous avons défini le délai d'accès au premier traitement comme étant le délai entre la date de première prise en charge du patient au CHU dans le cadre de son cancer (t0) et la date du premier traitement anti-cancéreux, ou traitement initial (t1), qui correspondait à une chirurgie curative de l'appareil respiratoire, une radiothérapie ou une chimiothérapie anticancéreuse. Ce délai a également été calculé pour chacune des deux cohortes, et était exprimé sous la forme d'un délai moyen et médian en jours.

Pour identifier le traitement initial au sein de l'EDS, nous avons utilisé les données codées du dossier patient, selon les terminologies CIM10 ou CCAM (Tableau 4). La chirurgie curative était identifiée par les actes de la classification CCAM commençant par GFFA et par le code GFFC002, la radiothérapie par les codes CCAM commençant par ZZNL et AANL et le code CIM10 Z5101 situé en DP avec un DR relatif à un cancer (codes CIM10 C34, C76.1, C77-C79 et C80) ou situé dans un formulaire du dossier patient, et la chimiothérapie par le code CIM10 Z511 situé en DP avec un DR relatif à un cancer (codes CIM10 C34, C76.1, C77-C79 et C80) ou situé dans un formulaire du dossier patient.

La date t1 correspondait alors à la date de début du RUM si le traitement était identifié par un code CIM10 en DP, à la date d'enregistrement du formulaire dans l'EDS pour les codes CIM10 extraits d'un champ de formulaire, ou à la date de réalisation de l'acte pour les traitements identifiés par un code CCAM.

Tableau 4 : Codes CIM10 (35) et CCAM (36) identifiant le traitement anticancéreux initial
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

Libellés	**Codes CIM10**	**Codes CCAM**
Chirurgie curative de l'appareil respiratoire		GFFA… GFFC002
Séance de radiothérapie	Z51.01	ZZNL… AANL…
Chimiothérapie pour tumeur	Z51.1	

Pour un code CCAM, un « . » renvoie à n'importe quel caractère alphanumérique

3.3 Analyses

Pour répondre à notre objectif principal, nous avons construit les indicateurs par mois et par années, de janvier 2017 à juillet 2020, de façon à suivre et pouvoir comparer les résultats d'une année sur l'autre. De plus, du fait de la complexité des données et du nombre important de facteurs d'ajustements potentiels, nous avons fait le choix de produire une analyse purement descriptive, sans tests statistiques.

4 Résultats

4.1 Patients identifiés et caractéristiques

Au 27/10/2020, date de l'analyse, 9 870 patients atteints de cancer primitif infiltrant du poumon ont été identifiés à partir de l'ensemble des données de l'EDS (Figure 6). Parmi eux, 2 340 (23,7 %) ont été identifiés par notre algorithme comme des patients atteints de tumeurs malignes primitives multiples et ont été exclus de notre échantillon. Au total dans l'EDS, 7 530 patients étaient atteints d'un cancer unique du poumon et parmi eux 2 753 avaient une date de début de prise en charge au CHU de Bordeaux entre le 01/01/2017 et le 31/07/2020.

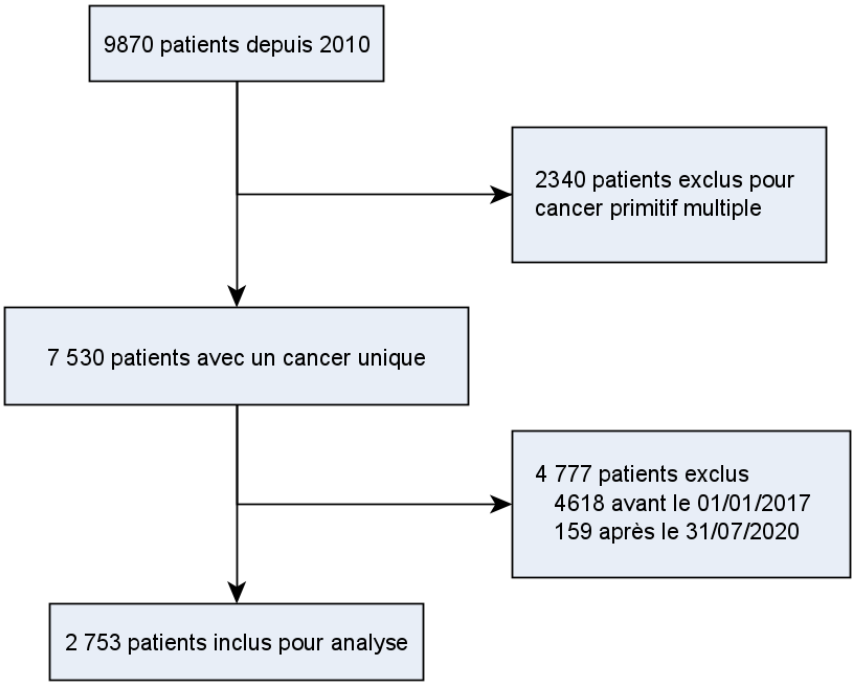


Figure 6 : Diagramme de flux des patients identifiés *via* l'entrepôt de données de santé du CHU de Bordeaux

Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

Parmi ces 2 753 patients, 1 093 (38,8 %) étaient identifiés par chacune des trois sources de données (Figure 7). Ces patients hospitalisés au CHU dans le cadre de leur cancer avaient des données du PMSI comportant en DP ou DR un des diagnostics recherchés, avaient eu un prélèvement tissulaire au CHU (CRAP dans les données d'anatomo-pathologie) et leur dossier avait été discuté en RCP au CHU. Par ailleurs, 632 (23 %) patients n'étaient repérés que par des données de RCP, 298 (10,8 %) par des données de CRAP et 133 (4,8 %) par les seules données diagnostics du PMSI. Les autres sources de données combinées sont détaillées dans la figure suivante.

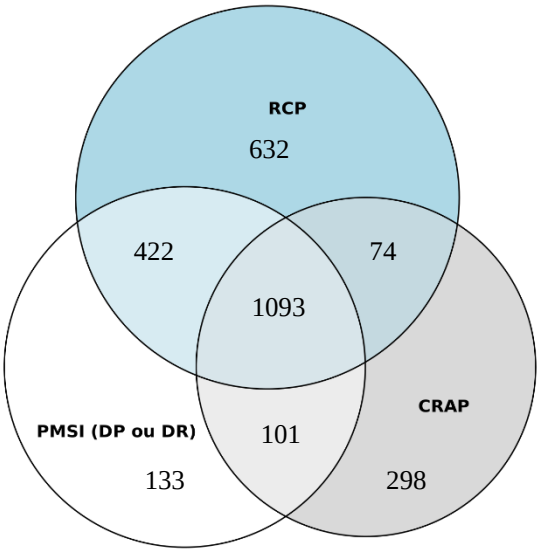


Figure 7 : Diagramme de Venn des patients inclus selon les critères et sources de données de l'entrepôt de données de santé du CHU de Bordeaux

PMSI (DP : diagnostic principal, DR : diagnostic relié), CRAP : compte rendu anatomopathologique, RCP : réunion de concertation pluridisciplinaire

Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

Le nombre de patients inclus par mois et année de début de prise en charge montrait une répartition qui semblait homogène entre 2017 et 2019 (Figure 8a). L'analyse des données cumulées met en évidence une diminution globale des cas nouvellement pris en charge en 2020 (422 patients entre janvier et juillet 2020versus 500 sur la même période en 2019), accentuée en mars 2020 suivie d'un rebond en juillet 2020 (Figure 8b).

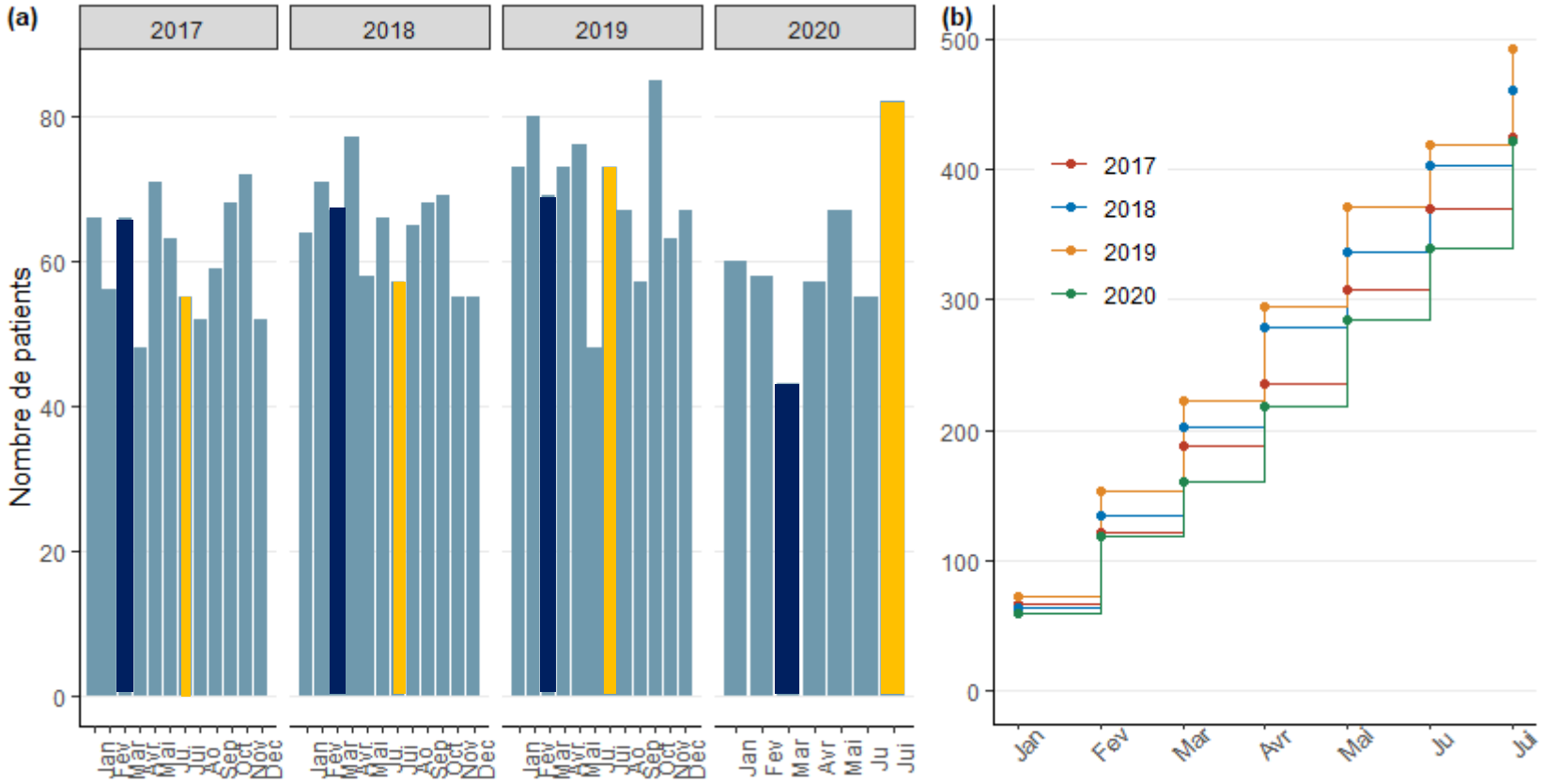


Figure 8 : Nombre de patients atteints de cancer infiltrant du poumon identifiés via l'entrepôt de données de santé du CHU de Bordeaux
(8a) Nombre de patients par mois et par année de 1ère prise en charge au CHU de Bordeaux (**mois de mars en bleu foncé, mois de juillet en jaune**).
(8b) Nombre cumulé de patients par année de 1ère prise en charge au CHU de Bordeaux, pour les mois de janvier à juillet
Projet : indicateurs des parcours patient avec cancers du poumon 2017-2020

L'âge moyen lors de la première prise en charge était similaire quelle que soit l'année d'observation (entre 65 et 66 ans en moyenne) alors que la proportion de patients de sexe féminin augmentait de 36,0 % en 2017 à 42,4 % en 2020 (Tableau 5).

Au sein des données de l'entrepôt, la source de la première information relative à un cancer, qui nous a permis d'attribuer une date de prise en charge au CHU pour chaque patient, était stable entre les années et correspondait le plus souvent à une RCP (42,2 % au total) ou une donnée du PMSI (36,2 % au total). De plus en 2020, 87,9 % des patients ont bénéficié d'au moins une RCP au CHU et 61,6 % avaient au moins une RCP qui correspondait, à l'échelle de l'entrepôt, à la définition d'une RCP dite « *pré-thérapeutique* » soit davantage qu'en 2019, avec respectivement 81,9 et 55,0 % des patients.

Enfin, la même proportion de patients a reçu un traitement anti-cancéreux au CHU (identifié par les données codées décrites dans le paragraphe 3.2.4) entre 2017 et 2020 (56,5 à 58,2 %). Concernant le type du premier traitement reçu au CHU, on observait en 2020 moins de chirurgie (25,6 % des patients) et plus de radiothérapie (17,8 %) qu'en 2019 (32,5 % de chirurgie initiale et 13,5 % de radiothérapie).

Tableau 5: Caractéristiques des patients identifiés, par année de première prise en charge au CHU de Bordeaux

Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

	Année de 1ère prise en charge au CHU								**Total**	
	2017		**2018**		**2019**		**2020*** **			
	n	%	n	%	n	%	n	%	*n*	*%*
Nombre de patients	728		772		831		422		*2753*	
Sexe										
Hommes	466	*(64,0)*	503	*(65,2)*	504	*(60,7)*	243	*(57,6)*	*1716*	*(62,3)*
Femmes	**262**	***(36,0)***	269	*(34,8)*	327	*(39,3)*	**179**	***(42,4)***	*1037*	*(37,7)*
Age moyen à la date de 1ère prise en charge au CHU (en années)	65,4		66,2		65,7		65,9		*65,8*	
Source du 1er code cancer*										
RCP	**312**	***(42,9)***	**325**	***(42,1)***	**347**	***(41,8)***	**177**	***(41,9)***	*1161*	*(42,2)*
PMSI (DP, DR, DAS ou DAD)	231	*(31,7)*	287	*(37,2)*	319	*(38,4)*	159	*(37,7)*	*996*	*(36,2)*
CRAP**	180	*(24,7)*	160	*(20,7)*	163	*(19,6)*	85	*(20,1)*	*588*	*(21,4)*
Autre formulaire	5	*(0,7)*	0	*(0,0)*	2	*(0,2)*	1	*(0,2)*	*8*	*(0,3)*
RCP au CHU										
>= 1 RCP	**626**	***(85,9)***	630	*(81,6)*	681	*(81,9)*	**371**	***(87,9)***	*2308*	*(83,8)*
>= 1 RCP « *pré-thérapeutique* »	408	*(56,0)*	409	*(53,0)*	457	*(55,0)*	**260**	***(61,6)***	*1534*	*(55,7)*
Traitement initial au CHU	411	*(56,5)*	447	*(57,9)*	484	*(58,2)*	245	*(58,1)*	*1587*	*(57,6)*
Chirurgie	191	*(26,2)*	222	*(28,8)*	270	*(32,5)*	**108**	***(25,6)***	*791*	*(28,7)*
Radiothérapie	101	*(13,9)*	122	*(15,8)*	112	*(13,5)*	**75**	***(17,8)***	*410*	*(14,9)*
Chimiothérapie	119	*(16,3)*	103	*(13,3)*	102	*(12,3)*	62	*(14,7)*	*386*	*(14,0)*

*code permettant d'attribuer une date de début de prise en charge au CHU de Bordeaux

**Compte Rendu Anatomo-Pathologique

***janvier à juillet 2020

4.2 Caractéristiques des patients par cohorte

Parmi les 2 753 patients identifiés par notre algorithme, 2 308 (83,8%) ont bénéficié d'au moins une RCP au CHU, dont 1 534 une RCP dite « *pré-thérapeutique* ». L'imbrication de ces différentes cohortes est schématisée dans la Figure 9, qui montre que certains patients avaient bénéficié d'un traitement sans trace de RCP dans leur dossier informatisé au CHU (39 patients, soit 2,5% des patients traités). En proportion, les patients de la cohorte avec « *RCP pré-thérapeutique* » avaient plus souvent bénéficié d'un premier traitement au CHU que les patients de la cohorte avec « *RCP* » (72,9 contre 67,1 %, Tableau 6).

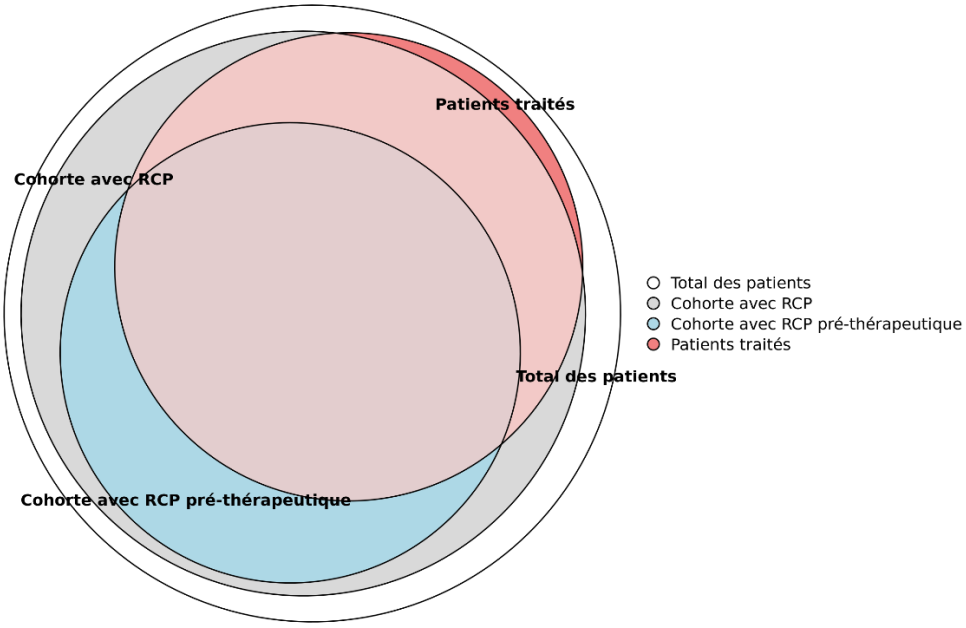


Figure 9 : Diagramme d'Euler des cohortes identifiées parmi l'ensemble des 2753 patients inclus, au CHU de Bordeaux
Cohorte avec RCP « pré-thérapeutique » incluse dans la cohorte avec « RCP », comprenant des patients traités et non traités ;
Cohorte de patients traités, comprenant des patients avec et sans RCP (les couleurs de chevauchement entre cohortes apparaissent différentes de la légende, celle-ci indiquant les couleurs propres aux cohortes)
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

Tableau 6 : Caractéristiques des patients pris en charge au CHU de Bordeaux, par cohorte d'analyse. Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

	Cohorte avec RCP		**Cohorte avec RCP *pré-thérapeutique***	
	n	*%*	n	*%*
Nombre de patients avec :	2308		1534	
Sexe				
Hommes	1419	*61,5*	953	*62,1*
Femmes	889	*38,5*	581	*37,9*
Age moyen à la date de 1ère prise en charge (en années)	65,5		65,7	
Source du 1er code cancer				
RCP	1161	*50,3*	817	*53,3*
PMSI (DP, DR, DAS ou DAD)	826	*35,8*	503	*32,8*
CRAP	314	*13,6*	207	*13,5*
Autre formulaire	7	*0,3*	7	*0,5*
Traitement initial au CHU	**1548**	***67,1***	**1119**	***72,9***

La RCP d'intérêt, qui nous a servi à construire les indicateurs de stade et de délai, était différente selon la cohorte analysée (Tableau 7) : par définition, pour la cohorte « *avec RCP* », cette RCP correspondait à la première RCP du parcours patient au CHU, alors que pour la cohorte « avec *RCP pré-thérapeutique* », le rang moyen était de 1,3 pour cette RCP (avec un rang maximum égal à 6). Les motifs de réalisation de la RCP d'intérêt étaient également différents entre les deux cohortes : moins de mise en traitement et plus d'avis diagnostique ou de suivi de traitement pour la cohorte « avec RCP ». Enfin, la proposition de prise en charge au décours de cette RCP aboutissait dans 40,9% des cas à une proposition qui ne relevait pas d'une prise en charge thérapeutique (proposition « autres ») pour les patients de la cohorte « avec RCP » alors que, par notre définition, les patients de la cohorte avec « *RCP pré-thérapeutique* » avaient tous une proposition de traitement au décours de cette RCP d'intérêt. Parmi les patients pour lesquels nous avions identifié un traitement anti-cancéreux au CHU, 19,4 % des patients de la cohorte « avec RCP » avaient bénéficié d'un traitement avant leur première RCP au CHU, contre 13,0 % des patients de la cohorte « avec *RCP pré-thérapeutique* ». Pour ces traitements intervenant tôt dans le parcours patient, la majorité étaient des chirurgies curatives (78,7 % pour la cohorte « avec RCP » et 72,1 % pour la cohorte « avec *RCP pré-thérapeutique* »).

Tableau 7 : Caractéristiques de la RCP d'intérêt et du premier traitement réalisés au CHU de Bordeaux, par cohorte d'analyse, au 27/10/2020.
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

	Cohorte avec RCP		**Cohorte avec *RCP pré-thérapeutique***	
	n	%	n	%
Nombre de patients avec :	2308		1534	
Rang de la RCP d'intérêt				
Rang moyen	1		**1,3**	
Rang médian (min - max)	1	(1 - 1)	1	(**1 - 6**)
Motif de la RCP d'intérêt				
Mise en traitement	**1084**	**(47,0)**	1153	(75,2)
Avis diagnostique	**641**	**(27,8)**	271	(17,7)
Suivi de traitement	**551**	**(23,9)**	104	(6,8)
Surveillance	25	(1,1)	5	(0,3)
Autre	5	(0,2)	1	(0,1)
Proposition de la RCP d'intérêt				
Mise en traitement	1092	(47,3)	1332	(86,8)
Surveillance	207	(9,0)	135	(8,8)
Abstention thérapeutique	64	(2,8)	67	(4,4)
Autres	**945**	**(40,9)**	0	(0,0)
Patients avec traitement au CHU	1548		1119	
Réalisé avant la RCP	**300**	**(19,4)**	**147**	**(13,0)**
Type du 1er traitement				
Chirurgie	765	(49,4)	530	(47,4)
Radiothérapie	400	(25,8)	271	(24,2)
Chimiothérapie	383	(24,7)	318	(28,4)
Traitement réalisé avant la RCP				
Chirurgie	**236**	**(78,7)**	**106**	**(72,1)**
Radiothérapie	**41**	**(13,7)**	**26**	**(17,7)**
Chimiothérapie	**23**	**(7,7)**	**15**	**(10,2)**

*1ère RCP réalisée au CHU pour la cohorte *« avec RCP »*,

RCP « *pré-thérapeutique* » pour la cohorte « *avec RCP pré-thérapeutique* »

4.3 Indicateurs

Nous décrivons ici les résultats des trois indicateurs choisis pour l'analyse (stade TNM au diagnostic, délai d'accès à la proposition thérapeutique et délai d'accès au premier traitement), estimés pour chacune des cohortes décrites au paragraphe précédent.

4.3.1 Stade TNM au diagnostic

Pour la cohorte « avec RCP », nous avons dénombré en 2020 plus de stades inconnus (36,4 %), moins de stades localisés (14,8 %) et une proportion similaire de stades métastatiques (40,4 %) comparativement aux autres années (Figure 10a). A un niveau mensuel, 50% des patients pris en charge en avril 2020, en pleine période de confinement, avaient un stade TNM au diagnostic métastatique, et le pourcentage pour les stades localisés et localement avancés était plus faible (Figure 10b). Les autres mois étaient difficilement comparables du fait la variation importante de la proportion de stades inconnus entre les années.

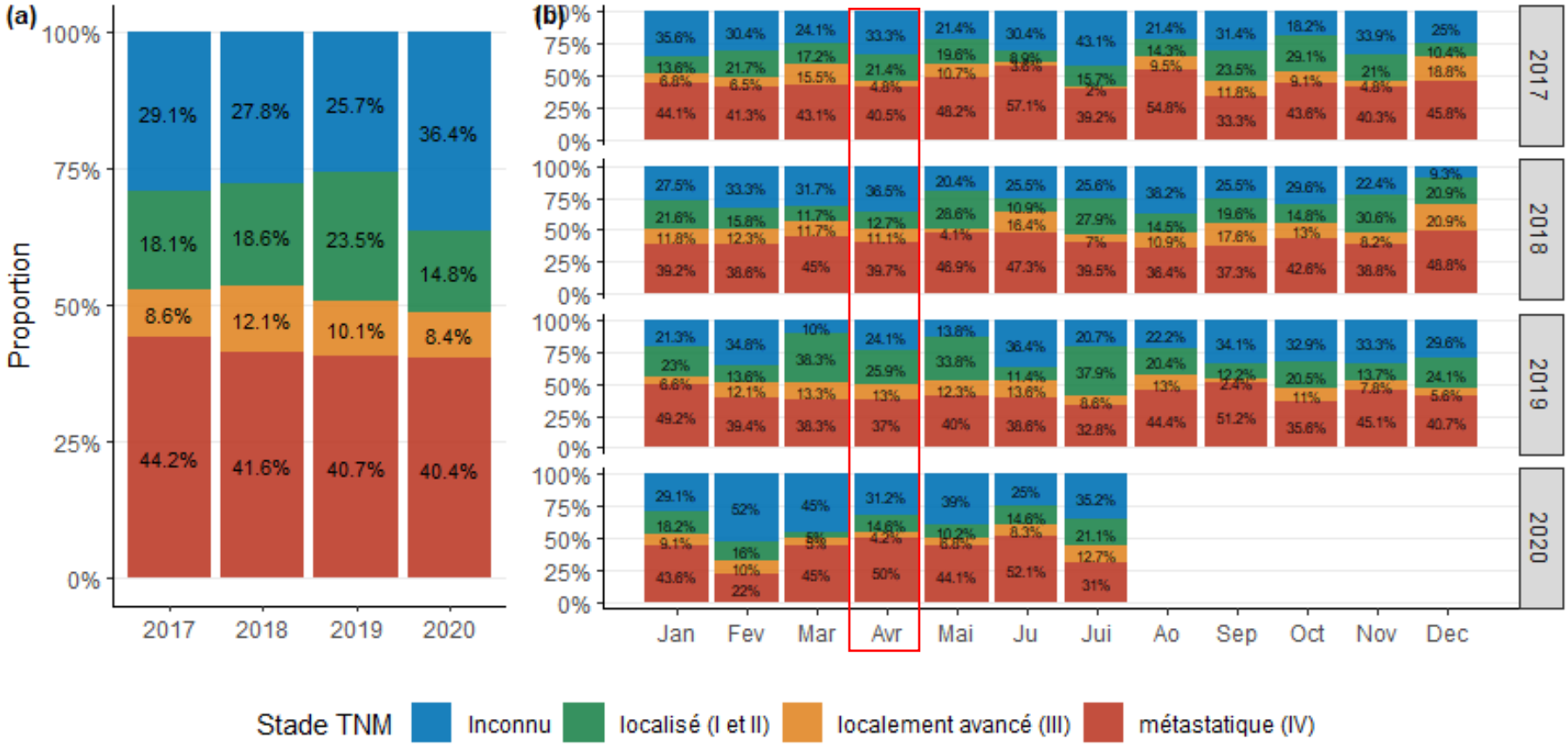


Figure 10 : Evolution du stade TNM au diagnostic, des cas de cancer du poumon identifiés via l'EDS pour la cohorte de patients « avec RCP »
(10a) Stades par année de 1ère prise en charge. (10b) Stades par année et par mois de 1ère prise en charge
encadré rouge : proportion de stades métastatiques élevée en avril 2020
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020, CHU de Bordeaux

Pour les résultats de la cohorte de patients « avec *RCP pré-thérapeutique* » (Figure 11), les résultats par année (Figure 11a) changeaient : en modifiant la donnée de l'entrepôt nous permettant d'estimer un stade TNM au diagnostic, nous trouvions une proportion plus grande de stades métastatiques en 2020 (48,1 % contre 42,5 % en 2019 par exemple), moins de stades localisés (16,9 % contre 25,6 % en 2019), avec une proportion de stades inconnus similaire entre les années (entre 21,5 % et 25,0 %). Cette proportion de stades inconnus était également plus faible par rapport à la cohorte des patients « avec RCP ».

De plus, le calcul des indicateurs à un niveau mensuel montre que cette tendance d'un diagnostic des patients à un stade plus avancé en 2020 semblait être accentuée pour les mois d'avril, mai et juin 2020, avec respectivement 57,6 %, 58,1 % et 64,5 % de patients diagnostiqués à un stade métastatique (contre 38,9 %, 34,1 % et 45,5 % en 2019), moins de stades localisés et localement avancés, et une proportion comparable entre années de stades inconnus.

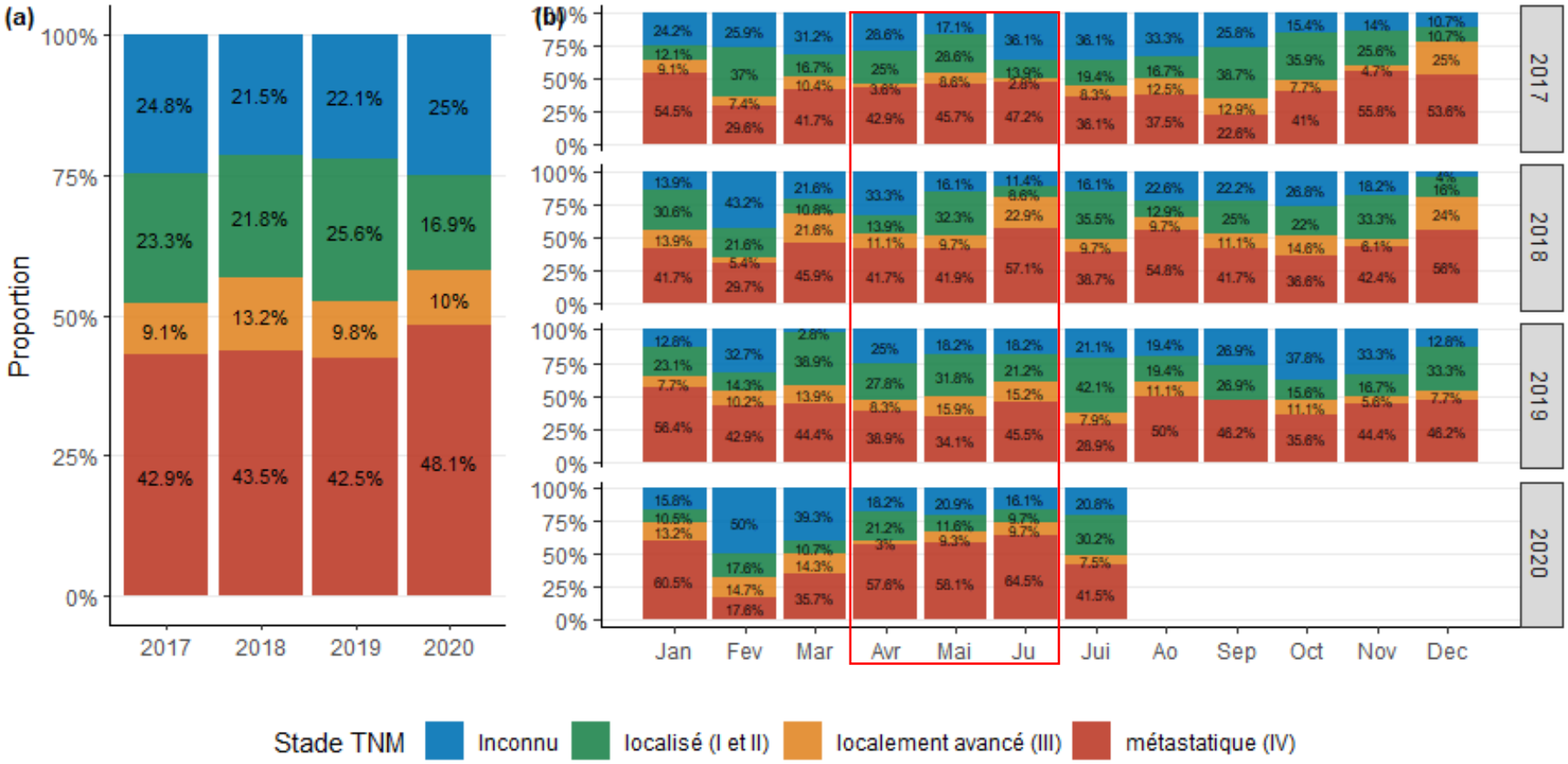


Figure 11 : Evolution du stade TNM au diagnostic, pour la cohorte de patients « avec RCP pré-thérapeutique ».
(11a) Stades par année de 1ère prise en charge. (11b) Stades par année et par mois de 1ère prise en charge
encadré rouge : proportion de stades métastatiques élevée en avril, mai et juin 2020
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020, CHU de Bordeaux

4.3.2 Délai d'accès à la proposition thérapeutique en RCP

Pour chaque délai, un graphique en boite à moustache (« boxplot ») composé d'un rectangle duquel sortait deux droites (les « moustaches ») a été produit (Figure 12, Figure 13, Figure 14, Figure 15). La valeur centrale du graphique était le délai médian en jours, les bords du rectangle les premiers et troisièmes quartiles, et les extrémités des moustaches étaient calculées en utilisant 1,5 fois l'espace interquartile (la distance entre le premier et le troisième quartile).

Le délai médian d'accès à la proposition thérapeutique en RCP était de zéro jours pour les deux cohortes, avec un délai moyen de 19,4 jours (écart-type 79,5 jours) pour la cohorte « avec RCP » et de 20,3 jours (écart-type 49 jours) pour la cohorte « avec *RCP pré-thérapeutique* » (Tableau 8). Pour les patients de ces deux cohortes, nous ne retrouvions pas de délais négatifs, la proportion de délais nuls (date de première prise en charge au CHU et date de la RCP identiques) étant de 50,7 % pour la cohorte « avec RCP » et 37,7 % pour la cohorte « avec *RCP pré-thérapeutique* ».

Tableau 8 : Délai en jours d'accès à la proposition thérapeutique en RCP pour les patients pris en charge au CHU de Bordeaux, par cohorte d'analyse
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

	Cohorte avec RCP		**Cohorte avec *RCP pré-thérapeutique***	
Nombre de patients avec :	N = 2308		N = 1534	
Délai négatif, n	0		0	
Délai nul, n (%)	1171	(50,7)	579	(37,7)
Délai >= 1 jour, n (%)	1137	(49,3)	955	(62,3)
Délai en jours				
Délai moyen (écart-type)	19,4	(79,5)	20,3	(49)
Délai médian	0		9	
Délai min, max	0	, 1174	0	, 800

Pour la cohorte « *avec RCP* », le délai médian était stable entre années, et montrait de légères variations pour les mois de janvier à juillet 2020 (Figure 12). Pour la cohorte « avec *RCP pré-thérapeutique* », nous notions une diminution du délai médian en 2019, pour revenir au niveau de 2017 et 2018 en 2020. Nous ne notions pas de variation nette entre les mois de janvier à juillet 2020 (Figure 13).

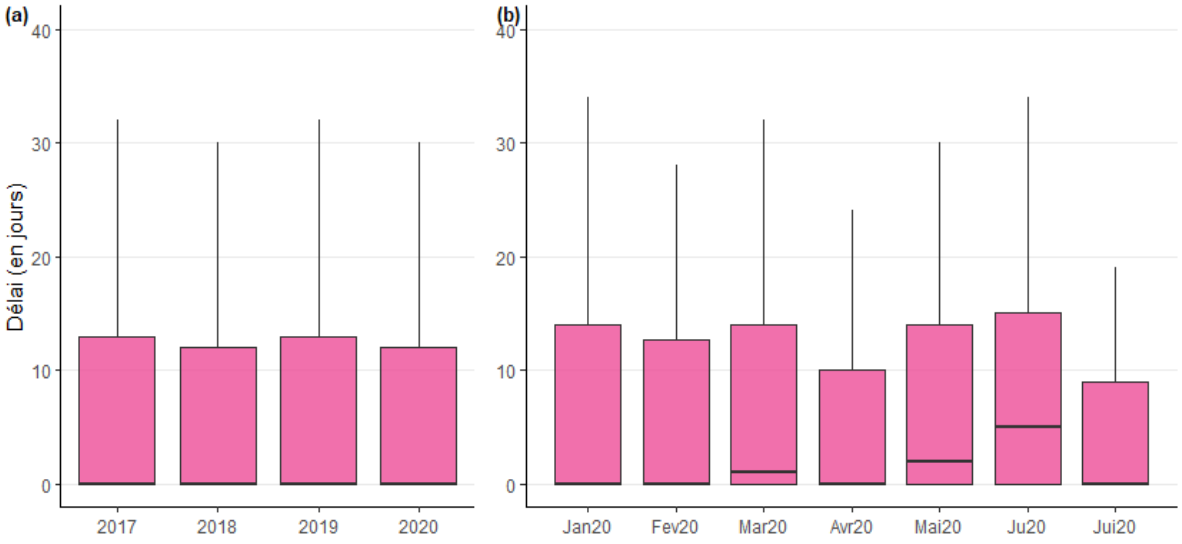


Figure 12 : Délai d'accès à la proposition thérapeutique en RCP, pour la cohorte de patients « avec RCP ».
(12a) Délais par année de 1ère prise en charge au CHU de Bordeaux.
(12b) Délais pour l'année 2020, pour les patients pris en charge entre janvier et juillet
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

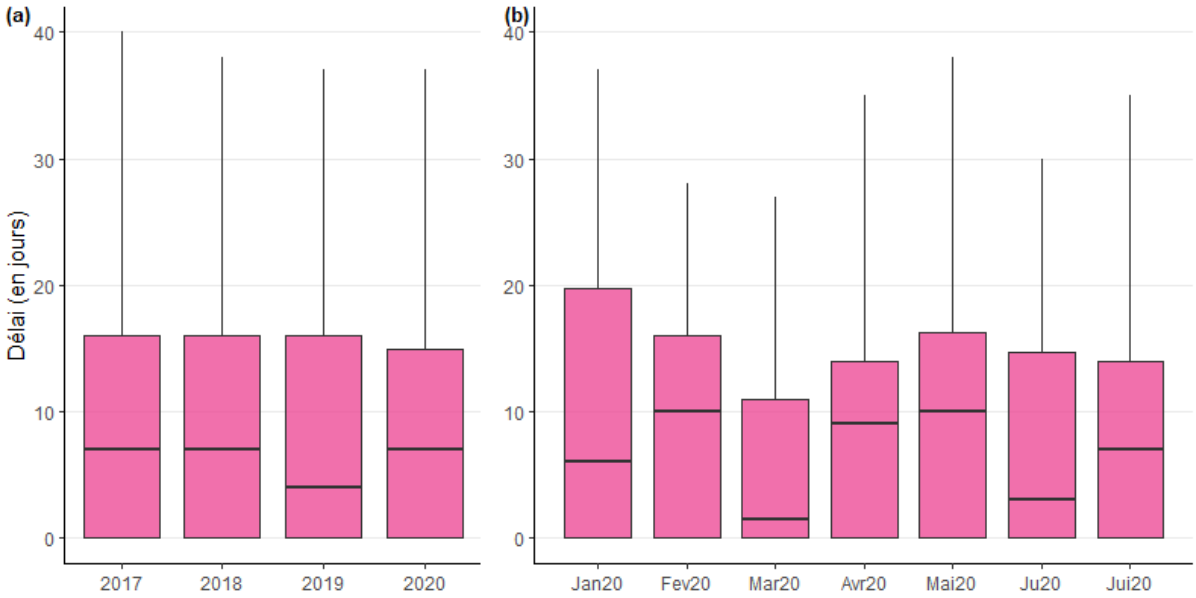


Figure 13 : Délai d'accès à la proposition thérapeutique en RCP, pour la cohorte de patients « avec RCP pré-thérapeutique ».
(13a) Délais par année de 1ère prise en charge au CHU de Bordeaux.
(13b) Délais pour l'année 2020, pour les patients pris en charge entre janvier et juillet
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

4.3.3 Délai d'accès au premier traitement

Le délai médian d'accès au premier traitement était similaire pour les deux cohortes (34 jours pour la cohorte « avec RCP » contre 35 jours pour la cohorte « avec *RCP pré-thérapeutique* »), avec un délai moyen de 48,4 jours (± 134 jours) pour la cohorte « avec RCP » et de 51,5 jours (± 133 jours) pour la cohorte « avec *RCP pré-thérapeutique* » (Tableau 9). Pour les patients traités de ces deux cohortes, nous retrouvions moins de délais négatifs dans la cohorte « avec *RCP pré-thérapeutique* » (0,8 % contre 1,7 %), la proportion de délais nuls (date de première prise en charge au CHU et date du 1er traitement identiques) étant de 75,6 % pour la cohorte « avec RCP » et 51,7 % pour la cohorte « avec *RCP pré-thérapeutique* ».

Tableau 9 : Délai en jours d'accès au premier traitement pour les patients pris en charge au CHU de Bordeaux, par cohorte d'analyse
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

	Cohorte avec RCP		**Cohorte avec *RCP pré-thérapeutique***	
Nombre de patients traités	1548		1119	
Délai négatif (n, %)	27	(1,7)	9	(0,8)
Délai nul, n (%)	1171	(75,6)	579	(51,7)
Délai >= 1 jour, n (%)	350	(26,1)	531	(47,5)
Délai en jours				
Délai moyen (écart-type)	48,4	(134)	51,5	(133)
Délai médian	34		35	
Délai min, max	-3529	, 892	-3529	, 714

Pour les deux cohortes, le délai médian était stable entre années (Figure 14a et Figure 15a). Pour les mois de janvier à juillet 2020, nous n'avons pas remarqué de variation nette pour la cohorte « avec RCP » (Figure 14b) alors qu'il semblait être majoré entre janvier et mars 2020 pour la cohorte « avec *RCP pré-thérapeutique* » (Figure 15b).

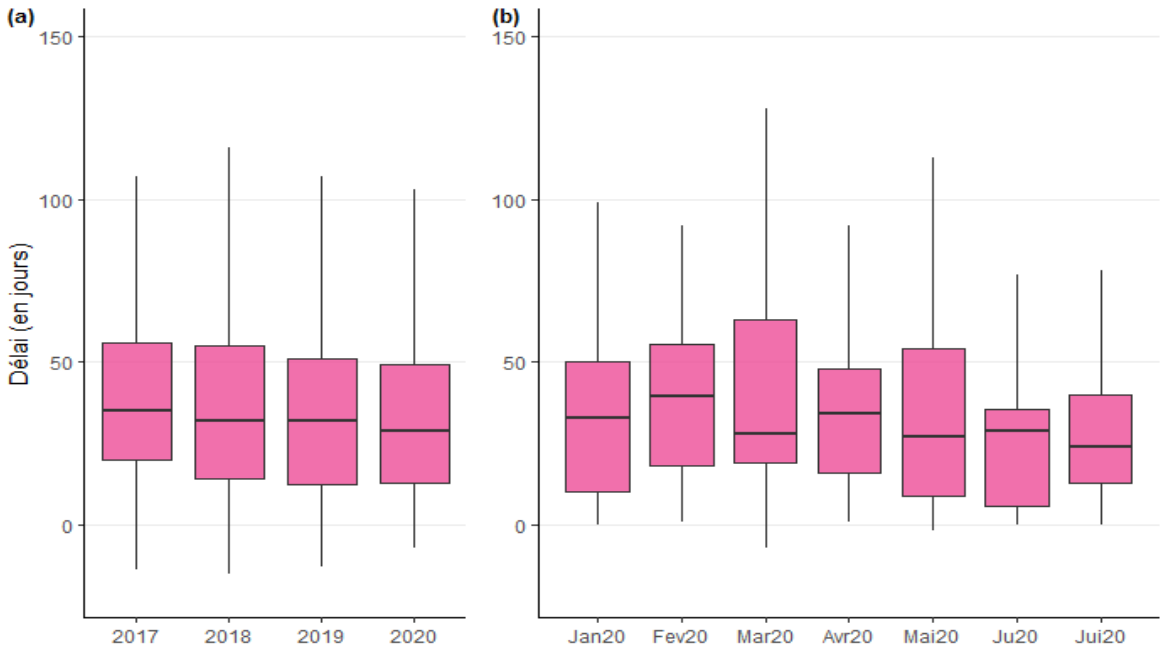


Figure 14 : Délai d'accès au premier traitement, pour la cohorte de patients « avec RCP »
(14a) Délais par année de 1ère prise en charge au CHU de Bordeaux.
(14b) Délais pour l'année 2020, pour les patients pris en charge entre janvier et juillet
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

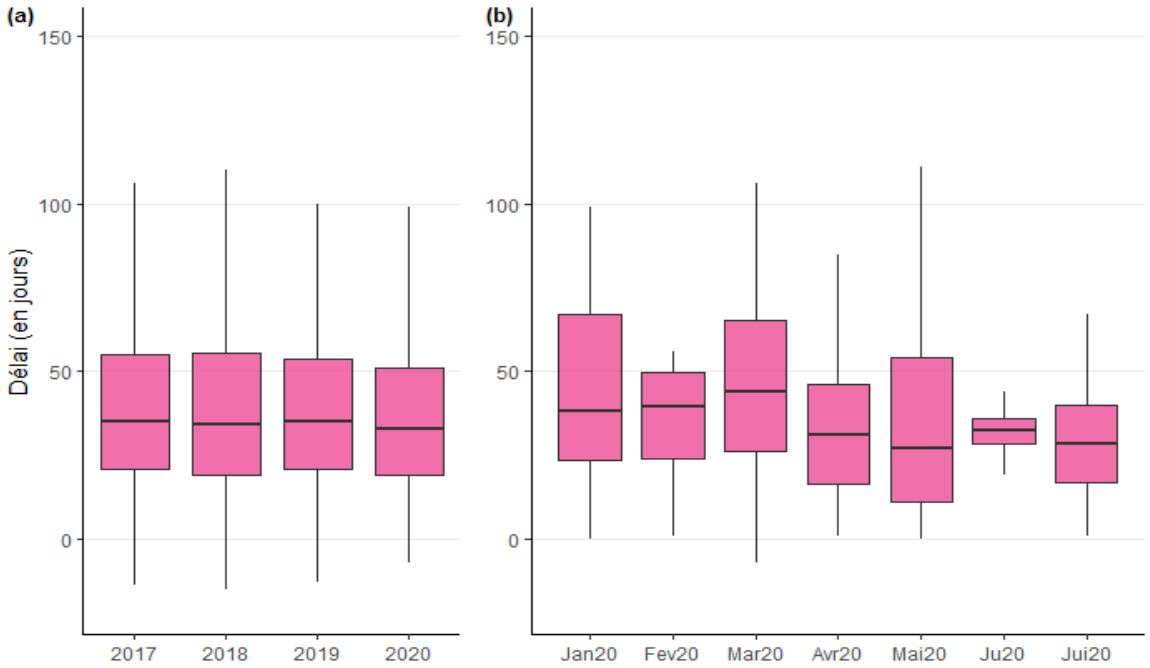


Figure 15 : Délai d'accès au premier traitement, pour la cohorte de patients « avec RCP pré-thérapeutique ».
(15a) Délais par année de 1ère prise en charge au CHU de Bordeaux.
(15b) Délais pour l'année 2020, pour les patients pris en charge entre janvier et juillet
Projet : indicateurs des parcours patients avec cancers du poumon 2017-2020

5 Discussion

5.1 Résultats principaux

Nous présentons ici les principaux résultats des trois indicateurs construits dans le but de décrire les parcours de soins des patients atteints de cancer du poumon pendant la période de crise sanitaire de 2020.

Le suivi dans le temps de ces indicateurs, depuis 2017, montre une tendance à un diagnostic de cancer du poumon à un stade plus avancé pour les patients dont la prise en charge en cancérologie au CHU débute entre les mois d'avril et juin 2020. Cette tendance est observable pour les deux cohortes étudiées mais de façon plus nette pour les patients de la cohorte avec « *RCP pré-thérapeutique* » pour lesquels la proportion de stade métastatique au diagnostic est estimée à respectivement 57,6 %, 58,1 % et 64,5 % pour les patients entrant au CHU en avril, mai et juin 2020 alors qu'elle était de 38,9 %, 34,1 % et 45,5 % à ces mêmes mois en 2019. Park *et al.* confirment ce résultat en Corée, avec une augmentation de la proportion de patients atteints de cancer du poumon non à petites cellules et diagnostiqués à un stade III ou IV entre février et juin 2020, comparativement à la même période en 2017, 2018 et 2019 (41).

A l'inverse, nous n'observons pas de retard de prise en charge lorsque le patient entre dans la filière de soins hospitalière, avec des délais d'accès à la proposition thérapeutique en RCP et d'accès au premier traitement anti-cancéreux similaires en 2017, 2018, 2019 et 2020. La réorganisation des soins en cancérologie mise en place au CHU de Bordeaux durant la crise sanitaire semble donc avoir permis de maintenir la qualité de la prise en charge hospitalière des patients atteints de cancer du poumon. Le diagnostic des patients atteints de cancer du poumon à un stade plus avancé pourrait alors s'expliquer par un retard au diagnostic pendant la pandémie de CoviD-19, avec un impact sur les parcours patients pré hospitaliers, comme le suggère la littérature avec une augmentation du délai d'orientation du patient du médecin généraliste vers un médecin spécialiste (42,43) ou encore une diminution de l'activité de dépistage et de diagnostic (14,19,44–47).

5.2 Fiabilité de nos indicateurs et comparaison avec la littérature

Notre méthode de construction des indicateurs repose sur la réutilisation de données hospitalières mono centriques, à travers l'EDS du CHU de Bordeaux. Il est ainsi primordial d'évaluer la fiabilité de cette méthode par rapport à une production d'indicateurs *via* le Système National des Données de Santé (SNDS) ou *via* les registres de cancers en population, méthodes préconisées pour « produire de façon automatique et sans biais des indicateurs qu'autrement seules des enquêtes *ad hoc* pourraient permettre » (48).

Selon l'INCa, les patients atteints de cancer du poumon sont diagnostiqués à un stade localisé pour 15 à 30 % d'entre eux, pour 20 % à un stade localement avancé et pour 40 à 55 % à un stade métastatique (24). En comparaison, les résultats de nos indicateurs en 2017, 2018 et 2019, hors période de crise sanitaire, sont cohérents avec ces intervalles pour nos deux cohortes d'analyse sauf pour les diagnostics à un stade localement avancé pour lesquels nous retrouvons une proportion plus faible (entre 9 et 13 % pour la cohorte « avec *RCP pré-thérapeutique* » et entre 8 et 12 % pour la cohorte « avec RCP ») et compte tenu d'une proportion non négligeable de stades inconnus (autour de 22 % pour la cohorte avec « *RCP pré-thérapeutique* » et 27 % pour la cohorte « avec RCP »).

En ce qui concerne les délais de prise en charge, leur interprétation dépend de leur définition mais aussi des données utilisées pour leur calcul : une étude des délais de prise en charge des cancers du poumon au Luxembourg en 2013 a montré par exemple, à partir des données du registre national des cancers, un délai médian de 21 jours entre la date de confirmation histologique et le traitement initial pour cancer et un délai médian de 29 jours entre la date de confirmation histologique et la date de la première RCP (avec une proportion importante de premières RCP réalisées après le 1er traitement, 51 %) (34). En France, une étude similaire réalisée en 2011 conclut à un délai médian de 11 jours d'accès à la proposition thérapeutique (défini comme le délai entre le diagnostic histologique et la première RCP) et un délai médian de 15 jours d'accès au premier traitement (délai entre la première RCP et le traitement initial) (33). Rivera *et al.* calculent quant à eux, par revue manuelle des dossiers patients, un délai médian entre le diagnostic histologique et la première RCP de 14 jours et un délai médian entre le diagnostic histologique et le 1er traitement de 31 jours (27).

La date à t0 la plus souvent utilisée pour le calcul des délais dans la littérature correspondait à la date du diagnostic histologique alors que dans notre travail, cette date t0 était définie par la date de première prise en charge au CHU pour cancer et correspondait le plus souvent à la première RCP du patient (50,3 % pour la cohorte « *avec RCP* », 53,3 % pour la cohorte avec « *RCP pré-thérapeutique* ») et seulement pour moins de 15% d'entre eux au CRAP signant le diagnostic histologique du cancer. Notre méthode de construction des délais de prise en charge à partir de l'EDS, bien que permettant un calcul pour un nombre plus important de patients, rend ainsi difficile la comparaison de ces résultats avec la littérature. Pour autant, notre méthode de construction étant identique pour toutes les années et nos délais reflétant la durée complète du parcours intra-hospitalier du patient, nous estimons qu'il est possible d'utiliser ces résultats pour suivre l'évolution dans le temps des délais de prise en charge à l'échelle de l'établissement.

Finalement, la cohérence de nos résultats avec la littérature concernant le stade TNM au diagnostic suggère la possibilité d'utiliser l'EDS pour produire de façon automatisée des indicateurs de prise en charge des patients atteints de cancer du poumon et tirer parti des avantages de cette méthode, à la fois règlementaires (l'EDS n'ayant pas les contraintes d'utilisation du SNDS) et liés à la disponibilité des données de santé (contrairement à l'EDS, plusieurs années sont nécessaires à l'intégration des données hospitalières au sein des registres de cancers en France).

5.3 Limites du travail

Néanmoins, plusieurs éléments, en lien avec les choix méthodologiques réalisés pour identifier notre population cible et construire nos indicateurs, limitent l'interprétation de nos résultats et nous incitent à être prudent sur leur signification en vie réelle.

5.3.1 Identification des patients atteints de cancer du poumon

Dès 2002, Pearson *et al.* soulignaient l'importance d'identifier de façon spécifique et exhaustive les patients pour lesquels nous voulons évaluer la qualité des soins reçus, dans le but de minimiser les biais de mesure (49).

Notre algorithme d'identification des patients repose sur des règles de décision utilisant uniquement des données codées du SIH, et construites à l'aide de la littérature. L'identification des patients atteints de cancer du poumon à partir des données codées des bases de données médico-administratives ou hospitalières (50–52) est réalisée avec une sensibilité variant de 77,7 à 92,2 %, une spécificité (probabilité de ne pas identifier de patient en excès) de 98,8 à 99,7 %, et une valeur prédictive positive (probabilité que le patient soit réellement malade lorsque l'algorithme l' identifie) de 45,2 à 80,9 % . Bousquet *et al.* soulignent de plus que « *la prise en compte des seuls DP du PMSI sous-estime le nombre de cas* » (53), réflexion confirmée par nos résultats avec 36,5 % (1004/2753) des patients identifiés uniquement avec des données de RCP ou de CRAP. Ainsi, même si nous pouvons être confiants sur la capacité de notre algorithme à détecter spécifiquement le cancer du poumon, une des principales limites de notre méthode d'identification des patients est l'absence de « *Gold standard* » nécessaire à son évaluation.

5.3.2 Méthode de construction des indicateurs

Une autre limite majeure à notre travail provient de l'incertitude et de la temporalité des données utilisées pour la construction de nos indicateurs : (a) à quelle étape du parcours correspond la date de première prise en charge au CHU de Bordeaux, (b) à quoi correspond la RCP utilisée pour le calcul du stade au diagnostic et du délai d'accès à la proposition thérapeutique, et (c) le traitement que l'on qualifie de « *traitement initial* » correspond-il réellement au premier traitement anti-cancéreux reçu par le patient ?

Nos résultats suggèrent l'importance de repérer avec précision la RCP amenant une première proposition thérapeutique (qualifiée de « *RCP pré-thérapeutique* » dans ce travail). En effet, la première RCP réalisée au CHU avait pour indication une « mise en traitement » pour seulement 47 % des patients identifiés, alors que cette indication concernait 75,2 % des patients au moment de la RCP dite *« pré-thérapeutique »*. De plus, la séquence <RCP, traitement initial> était plus logique pour la cohorte de patients avec *« RCP pré-thérapeutique »* (13 % de traitement anti-cancéreux réalisé avant cette RCP contre 19% pour la cohorte « *avec RCP* »). Enfin, même si nos résultats sont cohérents avec la littérature pour les deux cohortes de patients entre le stade réel au diagnostic et le stade construit dans ce travail, la proportion plus faible de stades inconnus pour la cohorte avec *« RCP pré-thérapeutique »* (autour de 23 %, contre 30% pour la cohorte « *avec RCP* ») confirme la pertinence de notre définition de la RCP « *pré-thérapeutique* » pour construire des indicateurs fiables et reproductibles.

En ce qui concerne la date de première prise en charge, t0 de nos indicateurs de délai, celle-ci correspondait dans ce travail à la date de la première donnée codée « *cancer* » présente dans le dossier patient. La source de ce *« code cancer »* était variable en fonction des patients (RCP, PMSI ou CRAP) et rend ainsi la comparaison des résultats avec la littérature difficile. Néanmoins, cette source restait homogène entre les années et entre les cohortes, permettant une comparaison et un suivi dans le temps des délais de prise en charge à l'échelle du CHU. Enfin, il est probable que le traitement initial identifié dans ce travail ne soit pas toujours le premier traitement reçu par le patient au cours de son parcours médical, avec des délais d'accès au premier traitement anormalement longs (jusqu'à 892 jours) pour certains patients identifiés dans notre étude.

5.4 Perspectives

Les limites identifiées amènent à envisager plusieurs perspectives d'amélioration, à la fois de notre méthode d'identification des patients et de notre méthode de construction des indicateurs.

5.4.1 Evaluer et améliorer l'identification des patients

Pour construire notre algorithme d'identification des patients atteints de cancer du poumon, nous avons fait deux choix méthodologiques : (a) l'utilisation d'une approche basée sur des règles de décision, (b) l'utilisation des données structurées codées du SIH. En l'absence de *Gold standard* nécessaire à l'évaluation de notre algorithme, ces deux choix nous ont permis d'identifier rapidement des patients avec un certain niveau de confiance, validé par la littérature. Pour autant, de nombreux auteurs suggèrent l'utilisation des données médicales textuelles (comptes rendus d'hospitalisation, d'imagerie…) pour améliorer les performances des algorithmes d'identification, notamment dans le domaine de la cancérologie (54–58). L'apport de ces données textuelles est indéniable mais complexifie beaucoup la construction des algorithmes, rendant obsolète une approche basée sur des règles construites manuellement (54). Ainsi, les algorithmes intégrant des données textuelles utilisent de préférence une approche par apprentissage automatique (*machine learning*) par laquelle les règles d'identification sont construites automatiquement par le programme. On distingue classiquement deux types d'algorithme de *machine learning* : les algorithmes supervisés et les algorithmes non supervisés. L'avantage des méthodes non supervisées est qu'elles ne nécessitent pas de connaitre à l'avance le résultat du programme (dans notre cas, « *le patient a-t-il un cancer du poumon ?* ») pour construire leurs règles d'identification (celles-ci sont alors construites uniquement à partir des données médicales fournies au programme, grâce à des mesures de similarité par exemple), nous permettant ainsi de se passer de *Gold standard* pour identifier nos patients. Cette approche est de plus en plus utilisée dans le domaine médical (59–63), notamment dans le domaine spécifique de la cancérologie (64–66).

Dans tous les cas, la construction d'une BDD « *Gold standard* », avec des patients dont on sait à l'avance qu'ils sont atteints de cancer du poumon et d'autres dont on sait qu'ils sont indemnes de cette maladie, est indispensable pour évaluer notre algorithme d'identification en termes de spécificité et de sensibilité, quelle que soit l'approche de construction utilisée.

5.4.2 Reconstruire le parcours patient à l'échelle d'un établissement

Enfin, il nous semble important de développer une méthode pour se situer dans le parcours patient afin d'identifier avec précision notre population cible, les patients atteints d'une tumeur maligne primitive infiltrante du poumon et pris en charge au CHU de Bordeaux dans le cadre d'un premier cancer, et ainsi d'améliorer la fiabilité de nos indicateurs, qui ont pour but de refléter la prise en charge initiale d'un patient atteint de cancer du poumon et uniquement cette prise en charge initiale. Dans ce travail, nous avons fait le choix d'identifier la RCP amenant à une première proposition thérapeutique pour repérer cette prise en charge initiale au sein du CHU, aboutissant à la construction de deux cohortes d'analyse aux caractéristiques différentes et soulignant le fait que plusieurs trajectoires de soins existent pour une même maladie (67,68) et que tout patient atteint de cancer du poumon pris en charge au CHU ne l'est pas nécessairement dans le cadre de sa prise en charge initiale et/ou peut bénéficier d'une prise en charge dans un autre établissement de santé. Ainsi, l'ensemble des évènements clés (encore appelés « évènements traceurs ») pour se repérer au sein du parcours patient n'est pas présent explicitement au sein du SIH (69). Pour résoudre cette problématique sans les contraintes liées à l'utilisation d'autres sources de données (registres des cancers, SNDS…), nous proposons deux choix méthodologiques basés sur ce qu'on appelle la fouille de données [2] (« *data mining* »), champ de recherche en plein essor dans le domaine médical (70) : (a) extraire des comptes rendus médicaux les évènements clés réalisés en dehors du CHU, à l'aide de méthodes de traitement automatique de la langue (71,72), ou (b) reproduire les séquences de soins du parcours intra-hospitalier à l'aide de méthodes de « *process mining* », « *sequence mining* » et « *temporal mining* » (sous-domaines du « *data mining* », s'intéressant aux évènements de soins et à leur relation temporelle) (69,73–75). L'objectif sous-jacent de cette reconstruction des parcours de soins est d'identifier des sous-groupes de patients répondant de façon plus précise aux caractéristiques de notre population cible et ainsi améliorer la fiabilité des indicateurs produits.

[2] Extraction de connaissance à partir de données, à l'aide de modèles statistiques

6 Conclusion

Notre étude avait pour objectif d'identifier des descripteurs des parcours de soins de patients en secteur hospitalier afin de mesurer leurs éventuelles variations durant la pandémie Covid-19.

Depuis 2017, le suivi montre une tendance à un diagnostic de cancer du poumon à un stade plus avancé pour les patients dont la prise en charge en cancérologie au CHU débute entre les mois d'avril et juin 2020. A l'inverse, nous n'observons pas de retard de prise en charge lorsque le patient entre dans la filière de soins hospitalière, avec des délais d'accès à la proposition thérapeutique en RCP et d'accès au premier traitement anti-cancéreux similaires entre 2017 et 2020. La réorganisation des soins en cancérologie mise en place au CHU de Bordeaux durant la crise sanitaire semble donc avoir permis de maintenir, sur ces critères, la qualité de la prise en charge hospitalière des patients atteints de cancer du poumon. Le diagnostic des patients atteints de cancer du poumon à un stade plus avancé pourrait alors s'expliquer par un retard au diagnostic pendant la pandémie de Covid-19, avec un impact sur les parcours patients pré hospitaliers.

Néanmoins, plusieurs éléments méthodologiques d'identification de la population cible et de construction de nos indicateurs limitent encore l'interprétation de ces résultats et incitent à être prudent sur leur signification en vie réelle. Ces limites pourraient à l'avenir être dépassées en travaillant sur deux voies d'amélioration :

- inclure dans l'algorithme d'identification des patients les données médicales textuelles (comptes rendus d'hospitalisation, d'imagerie…),

- reconstruire le parcours patient à l'échelle d'un établissement en identifiant les évènements traceurs dans les comptes rendus médicaux à l'aide d'un traitement automatique de la langue ou encore en reproduisant les séquences de soins du parcours intra-hospitalier par « *data mining* ».

Ainsi, ce travail a permis de montrer que le SIH du CHU de Bordeaux, interrogé à travers son EDS, dispose de suffisamment de données pour mettre en place et suivre des indicateurs de prise en charge en cancérologie, sans retour manuel au dossier patient. Ces indicateurs pourraient alors être disponibles en période de crise sanitaire afin d'aider à la mise en place de mesures adaptées dans l'ensemble des secteurs sanitaires.

7 Bibliographie

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

2. Defossez G, Le Guyader-Peyrou S, Uhry Z, Grosclaude P, Remontet L, Colonna M, et al. Estimations nationales de l'incidence et de la mortalité par cancer en France métropolitaine entre 1990 et 2018. Étude à partir des registres des cancers du réseau Francim. Résultats préliminaires. Synthèse. Saint-Maurice : Santé publique France, 2019. 20 p. Disponible à partir des l'URL : www.santepubliquefrance.fr ; https://geodes.santepubliquefrance.fr ; http://lesdonnees.e-cancer.fr/ ; https://www.e-cancer.fr/

3. Plan cancer 2014-2019 [Internet]. [cited 2019 May 3]. Available from: https://solidarites-sante.gouv.fr/IMG/pdf/2014-02-03_Plan_cancer-2.pdf

4. Circulaire DHOS/SDO no 2005-101 du 22 février 2005 relative à l'organisation des soins en cancérologie [Internet]. [cited 2020 Aug 7]. Available from: https://solidarites-sante.gouv.fr/fichiers/bo/2005/05-03/a0030034.htm

5. Circulaire DHOS/CNAMTS/INCa no 2007-357 du 25 septembre 2007 relative aux réseaux régionaux de cancérologie [Internet]. [cited 2020 Aug 7]. Available from: https://solidarites-sante.gouv.fr/fichiers/bo/2007/07-10/a0100154.htm

6. Kutikov A, Weinberg DS, Edelman MJ, Horwitz EM, Uzzo RG, Fisher RI. A War on Two Fronts: Cancer Care in the Time of COVID-19. Ann Intern Med. 2020 Mar 27;172(11):756–8.

7. Sud A, Jones ME, Broggio J, Loveday C, Torr B, Garrett A, et al. Collateral Damage: The Impact on Cancer Outcomes of the COVID-19 Pandemic [Internet]. Rochester, NY: Social Science Research Network; 2020 Apr [cited 2020 Aug 7]. Report No.: ID 3582775. Available from: https://papers.ssrn.com/abstract=3582775

8. van de Haar J, Hoes LR, Coles CE, Seamon K, Fröhling S, Jäger D, et al. Caring for patients with cancer in the COVID-19 era. Nat Med. 2020 May;26(5):665–71.

9. HCSP. Avis provisoire Recommandations relatives à la prévention et à la prise en charge du COVID-19 chez les patients à risque de formes sévères [Internet]. Rapport de l'HCSP. Paris: Haut Conseil de la Santé Publique; 2020 Mar [cited 2020 Oct 23]. Available from: https://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=775

10. Info Coronavirus COVID-19 - Les actions du Gouvernement [Internet]. Gouvernement.fr. [cited 2020 Oct 23]. Available from: https://www.gouvernement.fr/info-coronavirus/les-actions-du-gouvernement

11. You B, Ravaud A, Canivet A, Ganem G, Giraud P, Guimbaud R, et al. The official French guidelines to protect patients with cancer against SARS-CoV-2 infection. Lancet Oncol. 2020 May 1;21(5):619–21.

12. COVID-19 : Ré-organisation des soins [Internet]. Onco-Nouvelle-Aquitaine. [cited 2020 Oct 23]. Available from: https://onco-nouvelle-aquitaine.fr/covid-19-reorganisation-reprise-des-soins/

13. Chiaravalli S, Ferrari A, Sironi G, Gattuso G, Bergamaschi L, Puma N, et al. A collateral effect of the COVID-19 pandemic: Delayed diagnosis in pediatric solid tumors. Pediatr Blood Cancer [Internet]. 2020 Aug 6 [cited 2020 Aug 30]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7435564/

14. Vincentiis LD, Carr RA, Mariani MP, Ferrara G. Cancer diagnostic rates during the 2020 'lockdown', due to COVID-19 pandemic, compared with the 2018–2019: an audit study from cellular pathology. J Clin Pathol [Internet]. 2020 Jun 19 [cited 2020 Aug 7]; Available from: https://jcp.bmj.com/content/early/2020/06/19/jclinpath-2020-206833

15. Yang Y, Shen C, Hu C. Effect of COVID-19 Epidemic on Delay of Diagnosis and Treatment Path for Patients with Nasopharyngeal Carcinoma. Cancer Manag Res. 2020 May 25;12:3859–64.

16. Sud A, Torr B, Jones ME, Broggio J, Scott S, Loveday C, et al. Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: a modelling study. Lancet Oncol. 2020 Aug 1;21(8):1035–44.

17. London JW, Fazio-Eynullayeva E, Palchuk MB, Sankey P, McNair C. Effects of the COVID-19 Pandemic on Cancer-Related Patient Encounters. JCO Clin Cancer Inform. 2020 Jul 27;(4):657–65.

18. Jones D, Neal RD, Duffy SRG, Scott SE, Whitaker KL, Brain K. Impact of the COVID-19 pandemic on the symptomatic diagnosis of cancer: the view from primary care. Lancet Oncol. 2020 Jun;21(6):748–50.

19. Rutter MD, Brookes M, Lee TJ, Rogers P, Sharp L. Impact of the COVID-19 pandemic on UK endoscopic activity and cancer detection: a National Endoscopy Database Analysis. Gut [Internet]. 2020 Jul 19 [cited 2020 Aug 30]; Available from: https://gut.bmj.com/content/early/2020/07/19/gutjnl-2020-322179

20. Diamand R, Ploussard G, Roumiguié M, Oderda M, Benamran D, Fiard G, et al. Timing and delay of radical prostatectomy do not lead to adverse oncologic outcomes: results from a large European cohort at the times of COVID-19 pandemic. World J Urol [Internet]. 2020 Aug 10 [cited 2020 Aug 30]; Available from: https://doi.org/10.1007/s00345-020-03402-w

21. Safran C. Reuse of Clinical Data. Yearb Med Inform. 2014 Aug 15;9(1):52–4.

22. Safran C. Update on Data Reuse in Health Care. Yearb Med Inform. 2017 Aug;26(1):24–7.

23. Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc. 2010 Mar 1;17(2):124–30.

24. Prise en charge thérapeutique du cancer du poumon non à petites cellules : synthèse des recommandations - Ref : RECOSYNTPNP10 [Internet]. [cited 2020 Sep 7]. Available

from: https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Prise-en-charge-therapeutique-du-cancer-du-poumon-non-a-petites-cellules-synthese-des-recommandations

25. Brundage MD, Davies D, Mackillop WJ. Prognostic Factors in Non-small Cell Lung Cancer. Chest. 2002 Sep;122(3):1037–57.

26. Kasymjanova G, Small D, Cohen V, Jagoe RT, Batist G, Sateren W, et al. Lung cancer care trajectory at a Canadian centre: an evaluation of how wait times affect clinical outcomes. Curr Oncol. 2017 Oct;24(5):302–9.

27. Rivera C, Mathiaux J, Haaser T, Begueret H, Jougon J, Trouette R. Délais de prise en charge initiale des patients atteints de cancer bronchique traités en radiothérapie. Bull Cancer (Paris). 2012 Dec 1;99(12):1117–22.

28. Salazar MC, Rosen JE, Wang Z, Arnold BN, Thomas DC, Herbst RS, et al. Association of Delayed Adjuvant Chemotherapy With Survival After Lung Cancer Surgery. JAMA Oncol. 2017 May 1;3(5):610.

29. Nwagbara UI, Ginindza TG, Hlongwana KW. Health systems influence on the pathways of care for lung cancer in low- and middle-income countries: a scoping review. Glob Health. 2020 Mar 18;16(1):23.

30. Malalasekera A, Nahm S, Blinman PL, Kao SC, Dhillon HM, Vardy JL. How long is too long? A scoping review of health system delays in lung cancer. Eur Respir Rev [Internet]. 2018 Sep 30 [cited 2020 Nov 8];27(149). Available from: https://err.ersjournals.com/content/27/149/180045

31. Mohammed N, Kestin LL, Grills IS, Battu M, Fitch DL, Wong CO, et al. Rapid Disease Progression With Delay in Treatment of Non–Small-Cell Lung Cancer. Int J Radiat Oncol. 2011 Feb;79(2):466–72.

32. Maiga AW, Deppen SA, Pinkerman R, Callaway-Lane C, Massion PP, Dittus RS, et al. Timeliness of Care and Lung Cancer Tumor-Stage Progression: How Long Can We Wait? Ann Thorac Surg. 2017 Dec 1;104(6):1791–7.

33. LEDESERT B, Giraud J, POURCEL G, Bousquet P-J. Etude sur les délais de prise en charge des cancers du sein et du poumon dans plusieurs régions de France en 2011. INCa; 2011.

34. COUFFIGNAL S, JACOBS J, Saleh S, Untereiner M. Prise en charge des cancers du poumon au Luxembourg en 2013 – 2014 : Résultats issus du Registre National du Cancer - Rapport technique [Internet]. 2018 [cited 2020 Nov 8]. Available from: https://www.lih.lu/crp/publication/prise-en-charge-des-cancers-du-poumon-au-luxembourg-en-2013-2014-resultats-issus-du-registre-national-du-cancer-rapport-scientifique-book-11993

35. Organization WH. Classification statistique internationale des maladies et des problèmes de santé connexes [Internet]. Genève : Organisation mondiale de la Santé; 2009 [cited 2020 Nov 10]. Available from: https://apps.who.int/iris/handle/10665/44082

36. CCAM en ligne - Téléchargement [Internet]. [cited 2020 Nov 10]. Available from: https://www.ameli.fr/accueil-de-la-ccam/telechargement/index.php

37. Terminologie – ADICAP [Internet]. [cited 2020 Nov 10]. Available from: https://esante.gouv.fr/terminologie-adicap

38. Vogt A, Schmid S, Heinimann K, Frick H, Herrmann C, Cerny T, et al. Multiple primary tumours: challenges and approaches, a review. ESMO Open. 2017 May 1;2(2):e000172.

39. Soerjomataram I, Coebergh JW. Epidemiology of Multiple Primary Cancers. In: Verma M, editor. Cancer Epidemiology [Internet]. Totowa, NJ: Humana Press; 2009 [cited 2020 Sep 7]. p. 85–105. (Methods in Molecular Biology). Available from: https://doi.org/10.1007/978-1-59745-416-2_5

40. Fritz AG. Classification internationale des maladies pour l'oncologie: CIM-O. Genève: Organisation mondiale de la santé; 2008.

41. Park JY, Lee YJ, Kim T, Lee CY, Kim HI, Kim J-H, et al. Collateral effects of the coronavirus disease 2019 pandemic on lung cancer diagnosis in Korea. BMC Cancer. 2020 Oct 29;20(1):1040.

42. Statistics » Cancer Waiting Times [Internet]. [cited 2020 Nov 11]. Available from: https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/

43. Lai A, Pasea L, Banerjee A, Denaxas S, Katsoulis M, Chang WH, et al. Estimating excess mortality in people with cancer and multimorbidity in the COVID-19 emergency. 2020.

44. Song H, Bergman A, Chen AT, Ellis D, David G, Friedman AB, et al. Disruptions in preventive care: Mammograms during the COVID-19 pandemic. Health Serv Res [Internet]. [cited 2020 Nov 11];n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6773.13596

45. Dinmohamed AG, Cellamare M, Visser O, de Munck L, Elferink MAG, Westenend PJ, et al. The impact of the temporary suspension of national cancer screening programmes due to the COVID-19 epidemic on the diagnosis of breast and colorectal cancer in the Netherlands. J Hematol OncolJ Hematol Oncol. 2020 Nov 4;13(1):147.

46. Brugel M, Carlier C, Essner C, Debreuve-Theresette A, Beck M-F, Merrouche Y, et al. Dramatic changes in oncology care pathways during the COVID-19 pandemic: The French ONCOCARE-COV study. The Oncologist [Internet]. [cited 2020 Nov 11];n/a(n/a). Available from: https://theoncologist.onlinelibrary.wiley.com/doi/abs/10.1002/onco.13578

47. Kaufman HW, Chen Z, Niles J, Fesko Y. Changes in the Number of US Patients With Newly Identified Cancer Before and During the Coronavirus Disease 2019 (COVID-19) Pandemic. JAMA Netw Open [Internet]. 2020 Aug 4 [cited 2020 Nov 12];3(8). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7403918/

48. Quilleta A, Defossez G, Ingrand P. Surveillance des délais d'accès aux traitements chez des patientes atteintes d'un cancer du sein : une approche automatisée à partir des données d'un registre des cancers. Rev DÉpidémiologie Santé Publique. 2015 Mar;63:S7.

49. Pearson ML, Ganz PA, McGuigan K, Malin JR, Adams J, Kahn KL. The Case Identification Challenge in Measuring Quality of Cancer Care. J Clin Oncol. 2002 Nov 1;20(21):4353–60.

50. Baldi I, Vicari P, Di Cuonzo D, Zanetti R, Pagano E, Rosato R, et al. A high positive predictive value algorithm using hospital administrative data identified incident cancer cases. J Clin Epidemiol. 2008 Apr;61(4):373–9.

51. Setoguchi S, Solomon DH, Glynn RJ, Cook EF, Levin R, Schneeweiss S. Agreement of diagnosis and its date for hematologic malignancies and solid tumors between medicare claims and cancer registry data. Cancer Causes Control CCC. 2007 Jun;18(5):561–9.

52. Bronson MR, Kapadia NS, Austin AM, Wang Q, Feskanich D, Bynum JPW, et al. Leveraging Linkage of Cohort Studies with Administrative Claims Data to Identify Individuals with Cancer. Med Care. 2018 Dec;56(12):e83–9.

53. Bousquet P-J, Caillet P, Coeuret-Pellicer M, Goulard H, Kudjawu YC, Le Bihan C, et al. Recherche d'algorithmes d'identification des cancers dans les bases médico-administratives : premiers résultats des travaux du groupe REDSIAM Tumeurs sur les cancers du sein, du côlon-rectum et du poumon. Rev DÉpidémiologie Santé Publique. 2017 Oct 1;65:S236–42.

54. Xu H, Fu Z, Shah A, Chen Y, Peterson NB, Chen Q, et al. Extracting and Integrating Data from Entire Electronic Health Records for Detecting Colorectal Cancer Cases. AMIA Annu Symp Proc. 2011;2011:1564–72.

55. Zhou S. Extracting Phenotypes of Cancer Patients from Electronic Health Records. In: 2019 IEEE International Conference on Healthcare Informatics (ICHI). 2019. p. 1–2.

56. Kocbek S, Cavedon L, Martinez D, Bain C, Manus CM, Haffari G, et al. Text mining electronic hospital records to automatically classify admissions against disease: Measuring the impact of linking data sources. J Biomed Inform. 2016 Dec 1;64:158–67.

57. Sada Y, Hou J, Richardson P, El-Serag H, Davila J. Validation of Case Finding Algorithms for Hepatocellular Cancer From Administrative Data and Electronic Health Records Using Natural Language Processing. Med Care. 2016 Feb;54(2):e9-14.

58. Friedlin J, Overhage M, Al-Haddad MA, Waters JA, Aguilar-Saavedra JJR, Kesterson J, et al. Comparing Methods for Identifying Pancreatic Cancer Patients Using Electronic Data Sources. AMIA Annu Symp Proc. 2010;2010:237.

59. Zhang Y, Cai T, Yu S, Cho K, Hong C, Sun J, et al. High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). Nat Protoc. 2019 Dec;14(12):3426–44.

60. Colbaugh R, Glass K, Rudolf C, Tremblay Volv Global, Lausanne, Switzerland M. Learning to Identify Rare Disease Patients from Electronic Health Records. AMIA Annu Symp Proc. 2018 Dec 5;2018:340–7.

61. Khalid S, Prieto-Alhambra D. Machine Learning for Feature Selection and Cluster Analysis in Drug Utilisation Research. Curr Epidemiol Rep. 2019 Sep 1;6(3):364–72.

62. Hodapp C. Unsupervised Learning for Computational Phenotyping. ArXiv161208425 Cs Stat [Internet]. 2016 Dec 29 [cited 2020 May 31]; Available from: http://arxiv.org/abs/1612.08425

63. Guan W-J, Jiang M, Gao Y-H, Li H-M, Xu G, Zheng J-P, et al. Unsupervised learning technique identifies bronchiectasis phenotypes with distinct clinical characteristics. Int J Tuberc Lung Dis. 2016 Mar 1;20(3):402–10.

64. Miotto R, Li L, Kidd BA, Dudley JT. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. Sci Rep. 2016 May 17;6(1):26094.

65. Chen Y, Carroll RJ, Hinz ERM, Shah A, Eyler AE, Denny JC, et al. Applying active learning to high-throughput phenotyping algorithms for electronic health records data. J Am Med Inform Assoc. 2013 Dec 1;20(e2):e253–9.

66. Niţică Ş, Czibula G, Tomescu V. A comparative study on using unsupervised learning based data analysis techniques for breast cancer detection. In: 2020 IEEE 14th International Symposium on Applied Computational Intelligence and Informatics (SACI). 2020. p. 000099–104.

67. Roussot A, Nuemi G, Amat-Roze J-M, Quantin C. Analyse spatiale des trajectoires de prise en charge des patients atteints de cancer primitif du poumon en région Bourgogne. J Econ Medicale. 2012;Vol. 30(2):96–109.

68. Riou F, Jarno P. Représentation et modélisation des trajectoires de soins. ITBM-RBM. 2000 Oct 1;21(5):313–7.

69. Defossez G, Rollet A, Dameron O, Ingrand P. Temporal representation of care trajectories of cancer patients using data from a regional information system: an application in breast cancer. BMC Med Inform Decis Mak. 2014 Apr 2;14(1):24.

70. Lu J, Hales A, Rew D, Keech M, Fröhlingsdorf C, Mills-Mullett A, et al. Data Mining Techniques in Health Informatics: A Case Study from Breast Cancer Research. In: Renda ME, Bursa M, Holzinger A, Khuri S, editors. Information Technology in Bio- and Medical Informatics. Cham: Springer International Publishing; 2015. p. 56–70. (Lecture Notes in Computer Science).

71. Tate AR, Martin AG, Murray-Thomas T, Anderson SR, Cassell JA. Determining the date of diagnosis – is it a simple matter? The impact of different approaches to dating diagnosis on estimates of delayed care for ovarian cancer in UK primary care. BMC Med Res Methodol. 2009 Jun 23;9(1):42.

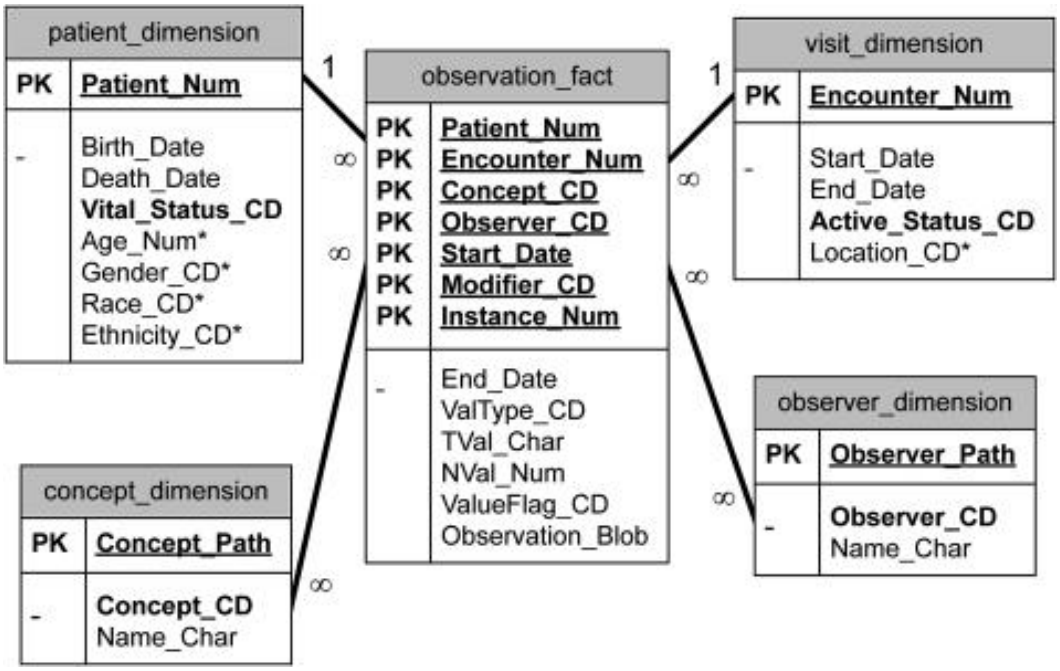
72. Solarte-Pabon O, Torrente M, Rodriguez-González A, Provencio M, Menasalvas E, Tuñas JM. Lung Cancer Diagnosis Extraction from Clinical Notes Written in Spanish. In: 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS). 2020. p. 492–7.

73. Bettencourt-Silva JH, Clark J, Cooper CS, Mills R, Rayward-Smith VJ, Iglesia B de la. Building Data-Driven Pathways From Routinely Collected Hospital Data: A Case Study on Prostate Cancer. JMIR Med Inform. 2015;3(3):e26.

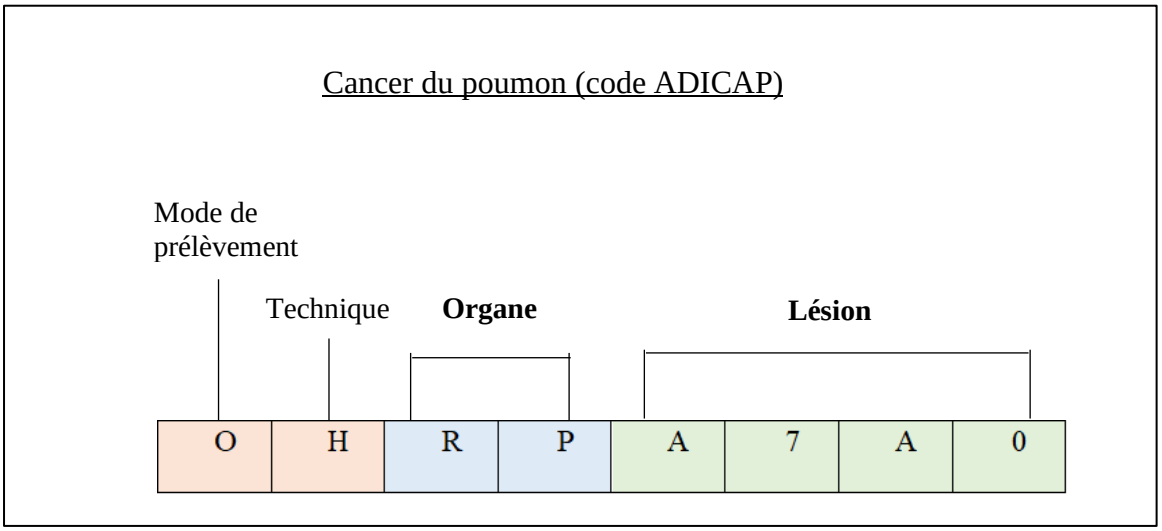
74. Baker K, Dunwoodie E, Jones RG, Newsham A, Johnson O, Price CP, et al. Process mining routinely collected electronic health records to define real-life clinical pathways during chemotherapy. Int J Med Inf. 2017 Jul 1;103:32–41.

75. Dagliati A, Sacchi L, Zambelli A, Tibollo V, Pavesi L, Holmes JH, et al. Temporal electronic phenotyping by mining careflows of breast cancer patients. J Biomed Inform. 2017 Feb 1;66:136–47.

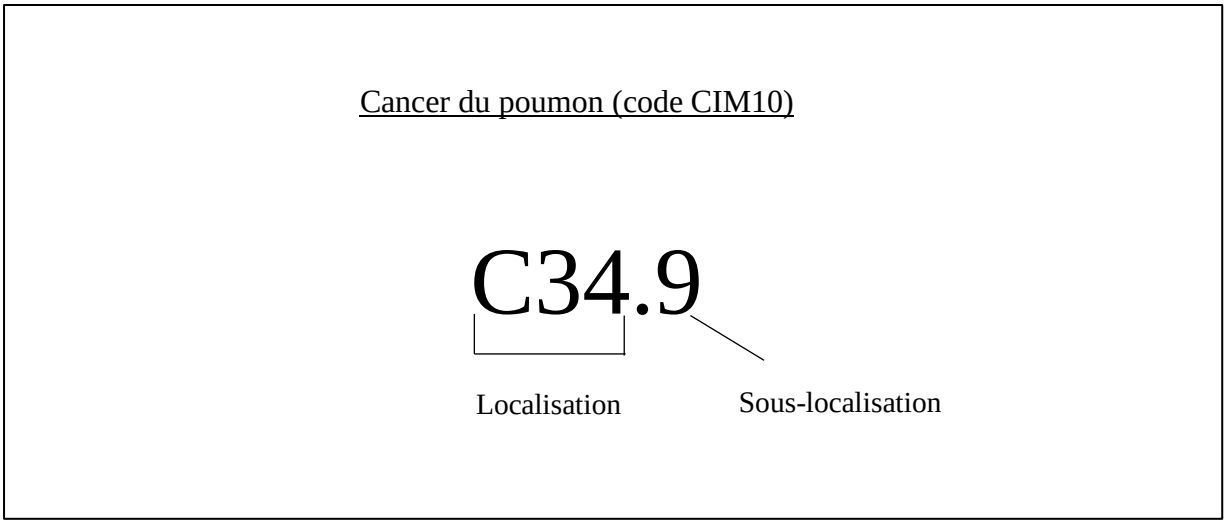
Annexes



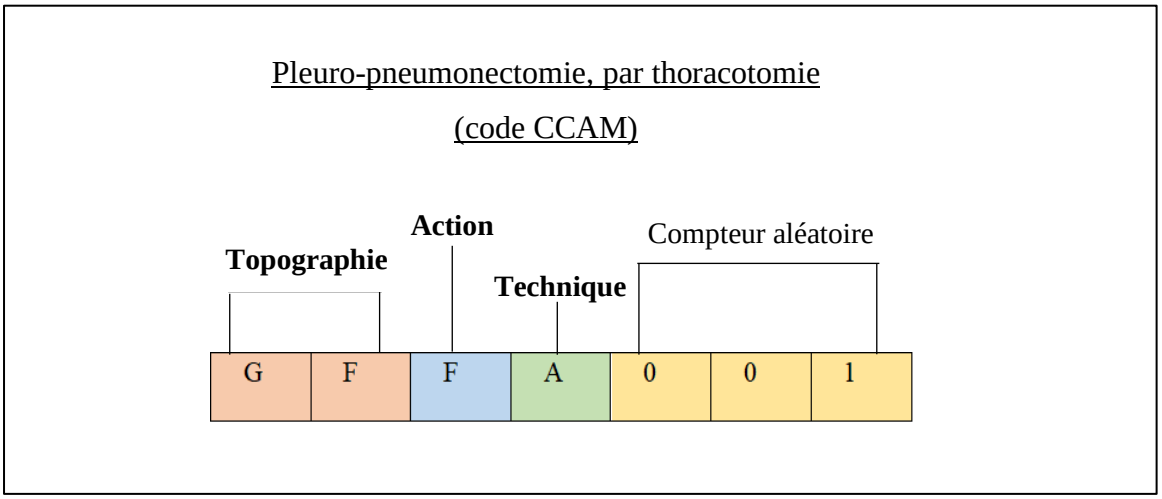
Annexe 1 : Modèle physique (dit « en étoile ») d'un entrepôt de données basé sur le modèle i2b2



Annexe 2 : Huit premiers caractères obligatoire d'un code ADICAP indiquant un adénocarcinome invasif (A7) SAI (A0) du poumon (RP), par histologie (H) d'une pièce opératoire (O)



Annexe 3 : Code CIM10 indiquant un cancer de bronche ou poumon (C34), sans précision (.9)



Annexe 4 : Code CCAM indiquant une chirurgie curative de l'appareil respiratoire, ici une exérèse (F) du poumon (GF), par abord ouvert (A)

Résumé :

En France, la cancérologie s'est structurée afin de garantir à tous les patients un parcours de soins sécurisé et de qualité. Ce parcours a été bouleversé en 2020 par la pandémie Covid-19, entrainant une possible perte de chance pour les patients atteints de cancer. Avec l'exemple du cancer du poumon, notre objectif principal a été de construire de façon automatisée des indicateurs pour décrire ce parcours pendant la période de crise sanitaire.

A partir de l'entrepôt de données de santé du CHU de Bordeaux, nous avons ainsi développé un algorithme d'identification des patients atteints de cancer du poumon et pris en charge au CHU, puis reconstruit leur stade TNM au diagnostic et calculé leur délai de prise en charge initiale.

Au total, 2 753 patients ont bénéficié d'une première prise en charge au CHU entre le 01/01/2017 et le 31/07/2020. En 2020, la proportion de patients diagnostiqués à un stade métastatique semblait plus importante (entre avril et juin 2020), alors que leur délai de prise en charge intra-hospitalière était équivalent aux autres années.

Ces résultats suggèrent un impact de la pandémie Covid-19 sur le parcours de soins des patients atteints de cancer du poumon. Néanmoins, plusieurs éléments, en lien avec les choix méthodologiques réalisés pour identifier notre population cible et construire nos indicateurs, limitent leur interprétation et nous incitent à être prudent sur leur signification en vie réelle.

Discipline

Santé publique

Mots clés :

Cancer du poumon, Covid-19, informatique médical, données de santé, phénotypage, indicateurs, parcours de soins

U.F.R.

Unité d'Informatique et d'Archivistique Médical (UIAM), CHU de Bordeaux