Approches analytiques innovantes pour la compréhension des parcours de soins pré-traitement dans le cancer du poumon

Exemple au travers de l'étude LUCIEN

03/11/2026

C2 - Restreint

Heva : spécialiste de l'analyse de donnés de santé depuis plus de 20 ans

Heva
Une filiale de DOCAPOSTE



Notre entrepôt SNDS FREiA :
Accélérer l'accès aux données du SNDS

IA & Data Science :
Machine learning, process mining

Datavisualisation et modélisation :
Nos outils de simulation et d'aide à la décision

Méthodologies innovantes :
Essais émulés, comparaisons externes, bras synthétiques, ...

Plateforme data / entrepôt de données de santé : Solution d'hébergement / administration et infogérance

Une association unique de savoir-faire au service de la donnée

Études

- Études et investigations cliniques / études interventionnelles
- Accès dérogatoires
- Études observationnelles

Objectifs & Finalités

- Épidémiologie
- Efficacité et sécurité
- Parcours de soins
- Bon usage et adhérence aux recommandations
- Médico-économie
- Offre de soins

Sources de données

Données primaires
- Recueil ad hoc
- Registres et cohortes cliniques

Bases médico-administratives
- SNDS, ESND, PMSI / Entrepôt FREiA

Autres données
- Données cliniques / EDS hospitaliers
- Données de biologie

LA POSTE
SANTÉ & AUTONOMIE

C2 - Restreint

Contexte et objectifs de l’étude LUCIEN

Parcours de soins et délais avant initiation d’un traitement dans le cancer du poumon : une étude d'analyse en conditions réelles à l'échelle nationale en France

C2 - Restreint

Contexte du projet LUCIEN

Parcours pré-diagnostiques du cancer du poumon en France

Contexte

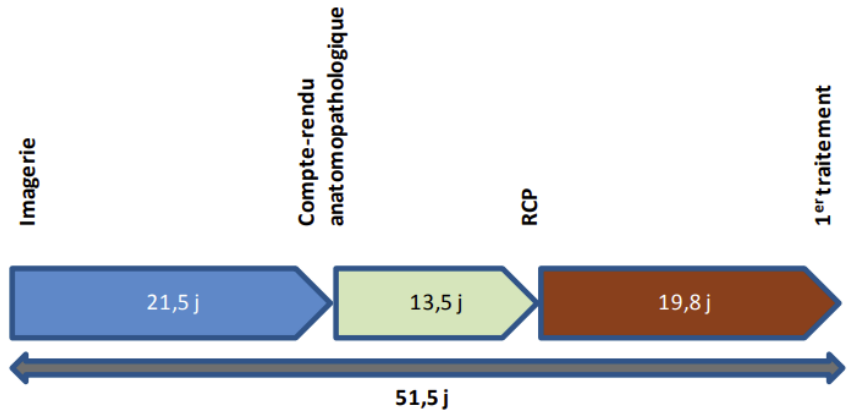
Dernière évaluation des parcours de soins du cancer du poumon en France, menée dans plusieurs régions et publiée en 2012 a fait état d'un délai moyen de 51 jours (médiane : 45 jours) entre le premier résultat d'imagerie anormal et le début du traitement, avec des variations en fonction de l'âge, du stade et de la région.

Réduction des délais entre diagnostic jusqu'au 1er traitement = enjeu clé, notamment depuis que les parcours diagnostiques se sont complexifiés avec la médecine stratifiée (recherche des mutations).

L'émergence des plateformes nationales de données multimodales et d'approches par machine learning créent de nouvelles possibilités de génération de preuves.

Données INCa 2012

Figure 9 : Distribution du délai global imagerie – premier traitement : délais moyens pour chaque étape du parcours de soins – poumon



Etude Sur Les Délais de Prise En Charge Des Cancers Du Sein et Du Poumon Dans Plusieurs Régions de France En 2011. Institut National du Cancer; 2012. https://www.e-cancer.fr/Expertises-et-publications/Catalogue-des-publications/Etude-sur-les-delais-de-prise-en-charge-des-cancers-du-sein-et-du-poumon

Objectifs de l'étude

- Identifier et décrire les différents groupes de parcours diagnostiques des patients atteints de cancer du poumon
- Explorer les potentiels facteurs associés à ces parcours à l'aide de méthodes de machine learning.

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

PRC LUCIEN

Acteurs et jalons clés sur le projet

Sponsor: MSD
Coordinateur: Filière IA & Cancers
Accès aux données: Institut National du Cancer
Analyse des données: HEVA

Les membres du Comité Scientifique

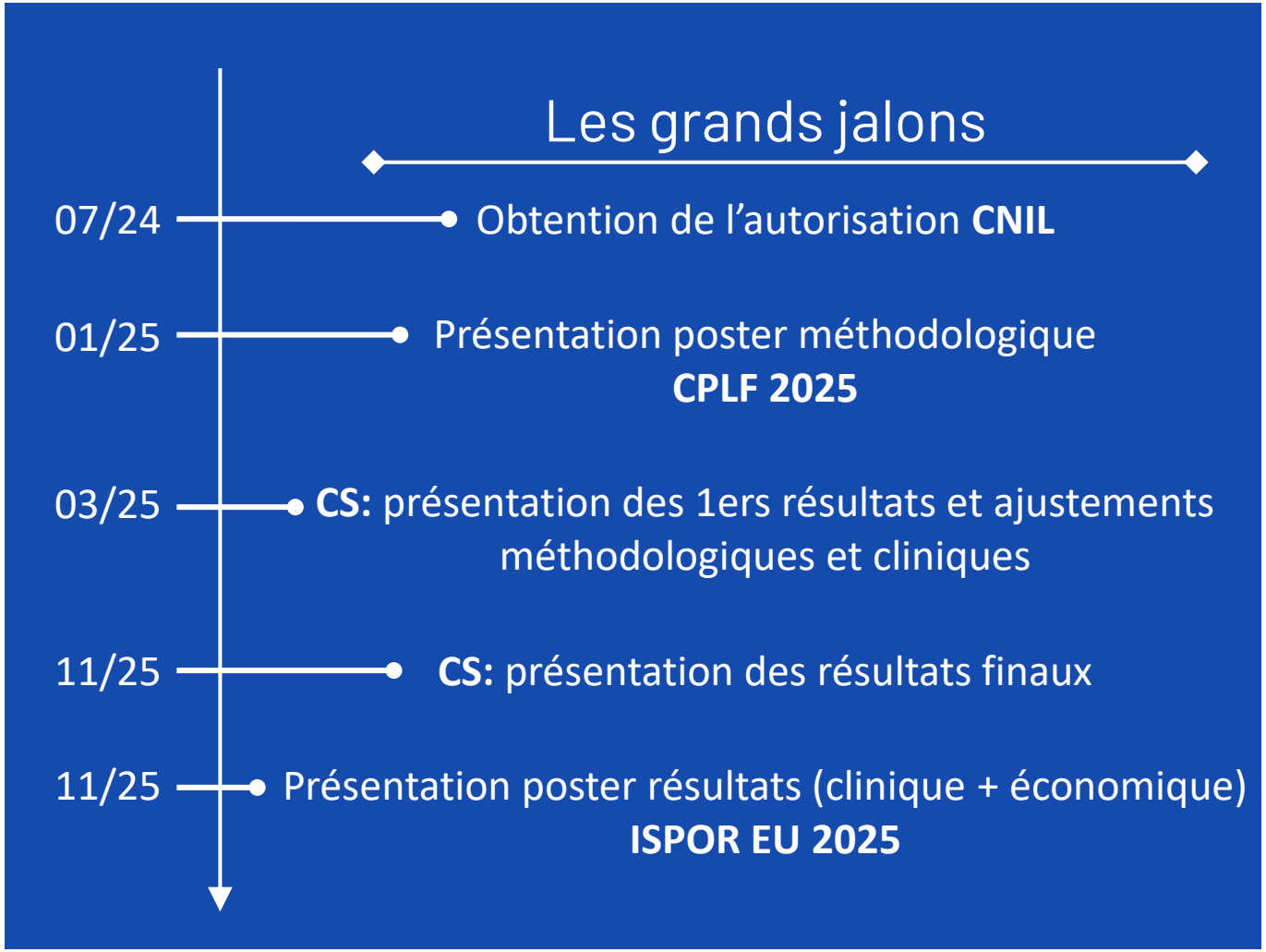
Pr Christos CHOUAID (GHI Créteil)

Pr Isabelle DURAND-ZALESKI (URC Eco, APHP Henri Mondor, CRESS UMR1153, UPEC)

Pr Laurent GREILLIER (APHM, Hôpital Nord)

Pr Pascal THOMAS (APHM, Hôpital Nord)

Pr Marie WISLEZ (APHP, Hôpital Cochin)



Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Méthodologie

C2 - Restreint

Méthode

Source de données

Registre national des cancers de l'INCa (anciennement PDC)

1ère analyse : focus sur données issues du SNDS

Population d'étude

Cohorte de **18 569 patients** traités pour un cancer pulmonaire primo-diagnostiqué entre le 1er juillet 2018 et le 30 juin 2019

Répartie en 2 groupes pour les analyses :



6 964
Patients localisés
Localement avancé
Et/ou traités par chirurgie thoracique



11 605
Patients en stade avancés/métastasiques

Les patients ont été classés comme présentant un cancer **avancé/métastatique** s'ils présentaient:

- un code de tumeur maligne secondaire (codes C78–C79 de la CIM-10) dans les trois mois suivant le début du premier traitement spécifique, OU
- un traitement systémique comme premier traitement spécifique.

Tous les autres patients, y compris tous ceux ayant bénéficié d'une intervention chirurgicale, ont été classés comme présentant un cancer localisé/traité par chirurgie.

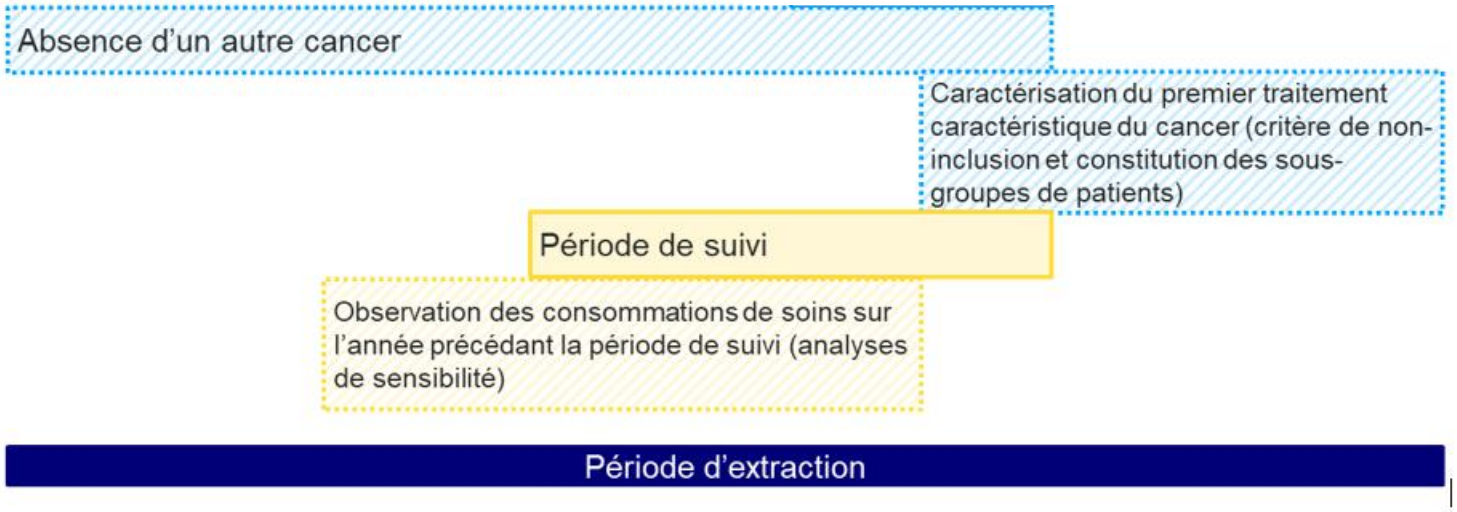
Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

PRC LUCIEN

Design et période d'étude





Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Apport des approches analytiques innovantes

Catégoriser et représenter les parcours diagnostiques



Évènements
- Épreuves fonctionnelles respiratoires
- Biopsies
- Passages aux urgences
- Consultations spécialistes
- Imageries thoraciques
- Recherche de métastases

1. Catégorisation automatique des parcours

Modèle de mélange fini

Méthode de clustering prenant en compte les trajectoires de consommations de soins dans l'année précédant la mise sous traitement



Catégorie de parcours A



Catégorie de parcours B



Catégorie de parcours C

2. Représentation automatique des parcours



PROCESS MINING







3. Description des délais entre chaque étape des catégories de parcours et étude des facteurs prédictifs

Délai entre évènement clés du parcours

- Délai court
- Délai médian
- Délai long



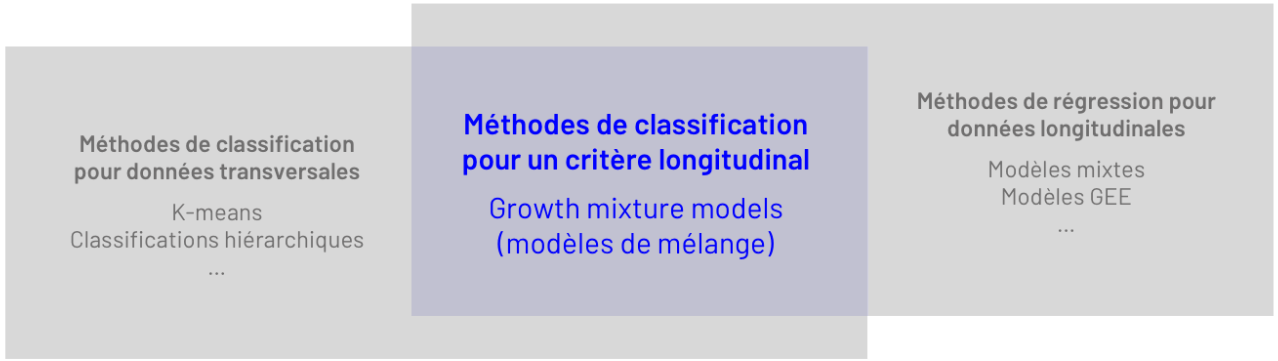
PROCESS MINING

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Catégoriser en prenant en compte les parcours de soins pré-traitement

Modèles de mélange - « mixture models »



- Nombre de groupes choisis à l'aide du critère d'information bayésien (BIC)
- Chaque groupe devait inclure au moins 5 % des patients de son groupe de stade
- Utilisation du package Flexmix (R)

Variables prises en compte dans le modèle

- Biopsies
- Epreuves d'effort
- Consultations spécialistes (pneumologue, oncologue et chirurgie thoracique)
- Imagerie médicale (angioscanner thoracique, scanographie 3 territoires, scanographie thoracique, IRM thoracique, fibroscopie, radiographie thoracique),
- Recherche de métastases (scanner cérébral, pet-scan, scintigraphie osseuse, IRM cérébrale)

Dans l'année avant initiation du traitement :

- Tous les passages aux urgences

Dans le mois qui précède l'initiation du traitement

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Analyse des parcours de soins à partir de données de vie réelle

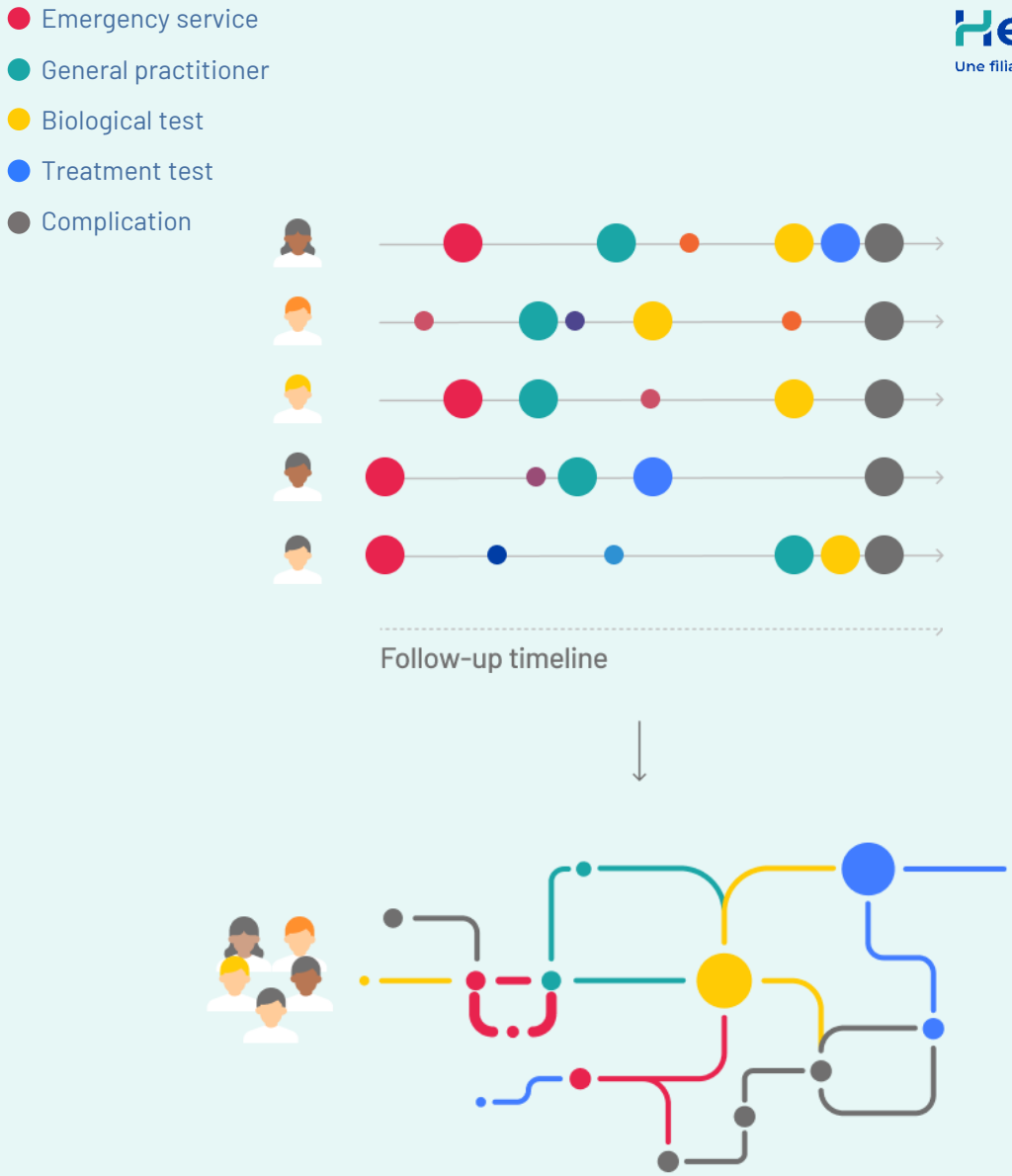
Enjeux et solutions proposées

Quels sont les enjeux de l'analyse des parcours de soins en vie réelle ?

- Chaque patient est différent et les données réelles sont imparfaites
- Des parcours de soins complexes qui peuvent être hétérogènes en fonction des lieux de prises en charge et des profils de patients
- Difficile d'obtenir une vision globale des parcours de soins et des types de patients en fonction des parcours

Pourquoi le *process mining* ?

- Agrégation de toutes les séquences de soins en un graphique unique pour en extraire des informations à l'échelle d'une cohorte.
- Modélisation des séquences temporelles en gardant un graphe lisible, représentatif et informatif.



Quelles sont les étapes clés du *process mining* ?

Approche méthodologique

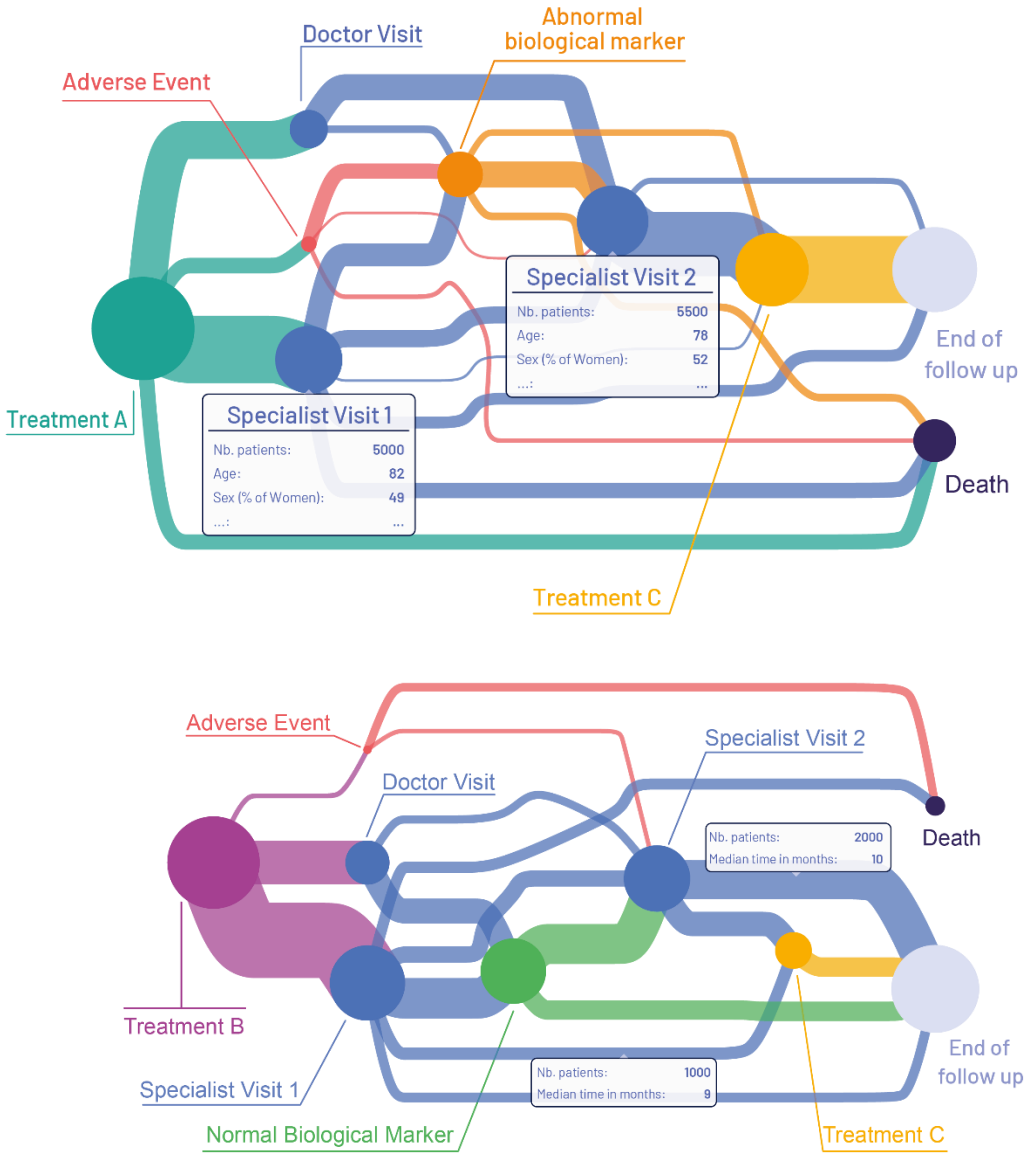
1) Identification des variables clés à intégrer dans la modélisation et des cohortes d'intérêt

- Pas trop de variables pour garder un graphe lisible
- Eviter les variables trop récurrentes
- Anticiper les regroupements d'intérêt
- Clustering envisageable en amont

2) Génération des graphes interactifs optimaux

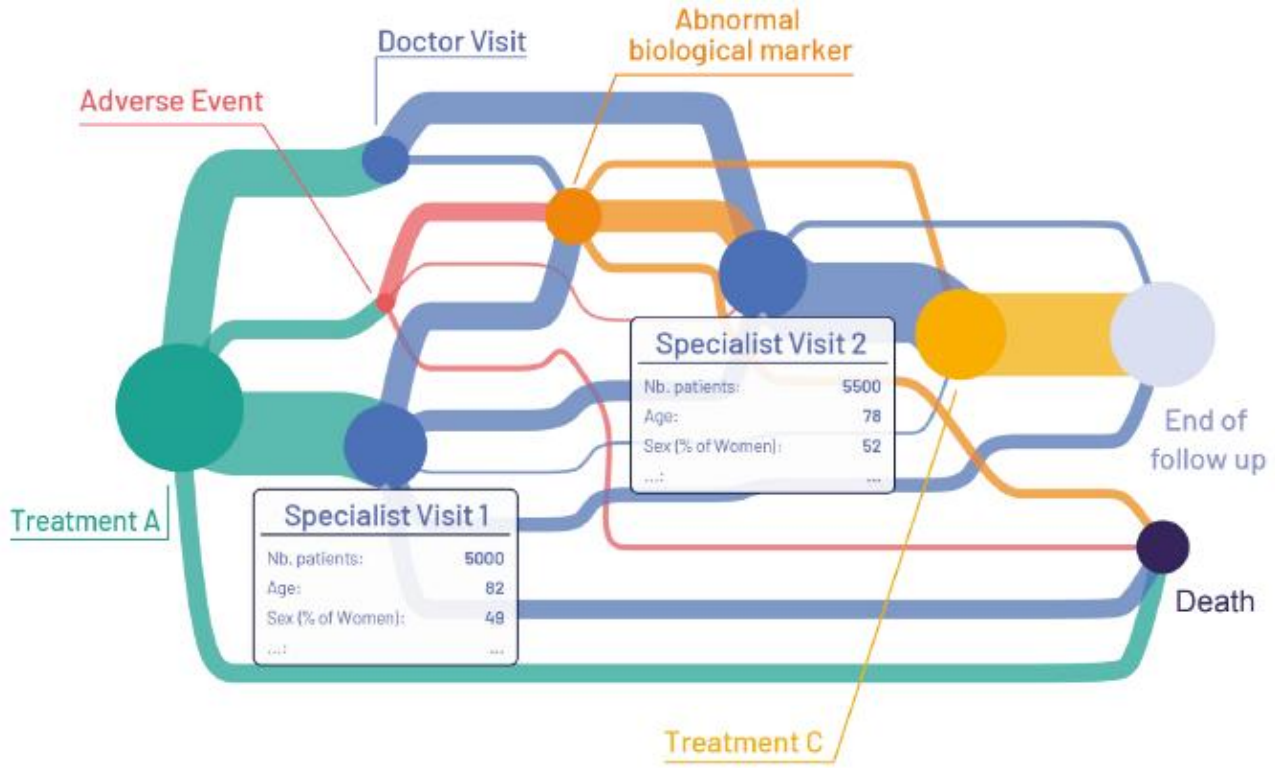
- Maximalisation de la rejouabilité (i.e. pourcentage permettant de quantifier la proportion du parcours d'un patient que l'on peut retrouver dans ce dernier)
- Descriptif des patients à chaque étape clé

3) Intégration sur une plateforme dynamique



Lecture des graphes de process mining

Caractéristiques des patients / délais disponibles à chaque étape clé du parcours

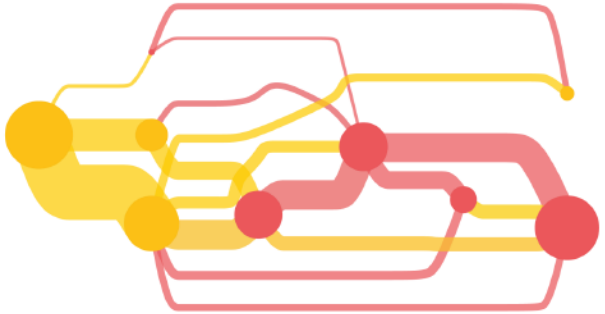


Déclinaison de présentation des résultats si on veut mettre l'accent sur les délais

Délai entre évènement clés du parcours

- Délai court
- Délai médian
- Délai long





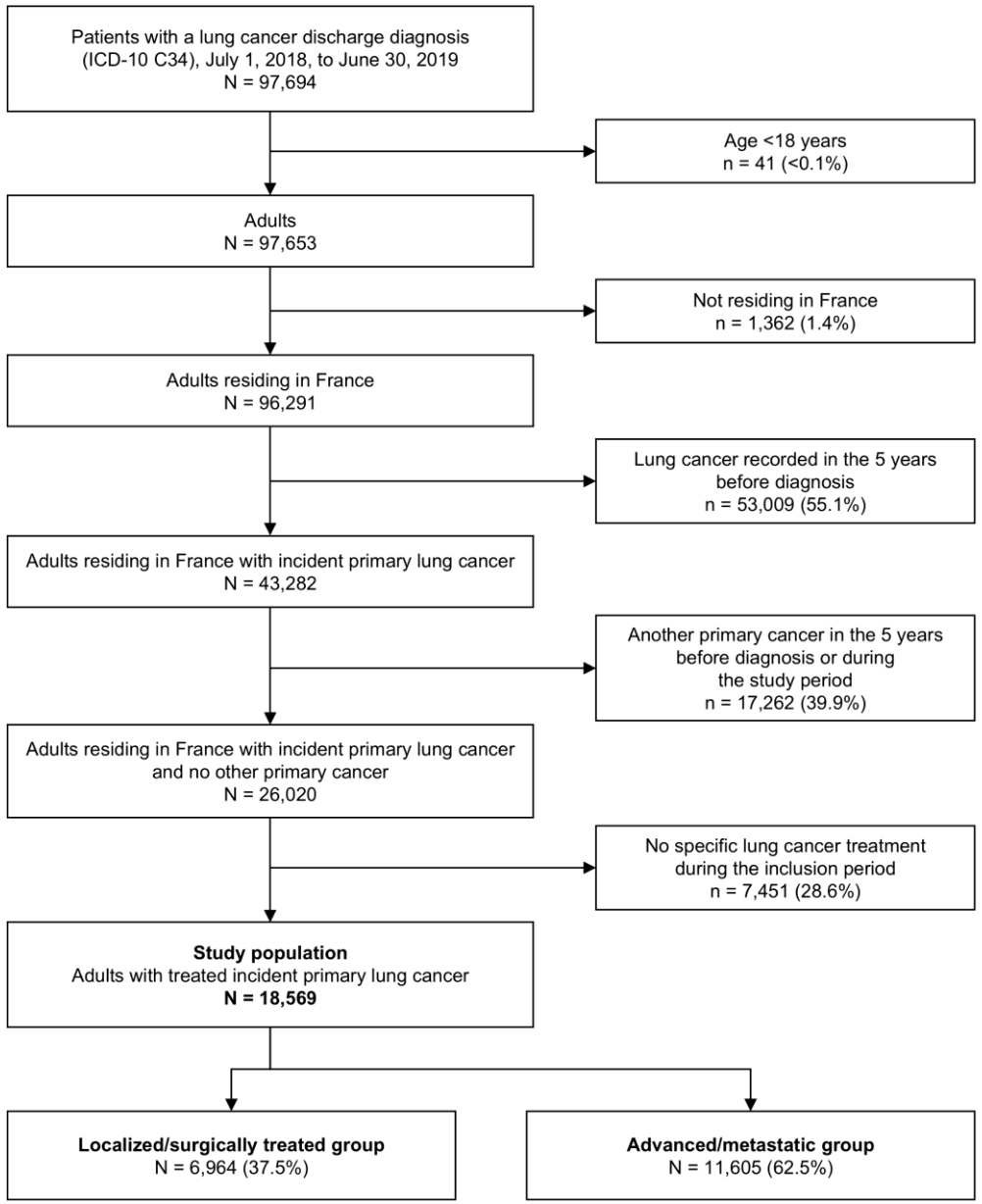
Heva
A subsidiary of DOCAPOSTE

Résultats clés de l’étude

C2 - Restreint

LA POSTE
SANTÉ & AUTONOMIE

Flowchart – sélection de la population



➔ **Modélisations effectuées sur 18 569 adultes résidant en France nouvellement traités pour un cancer du poumon primitif (sans présence d'un autre cancer dans les 5 ans précédant)**

➔ **Focus de cette présentation sur les 11 605 adultes présentant un stade avancé/métastatique**

C2 - Restreint

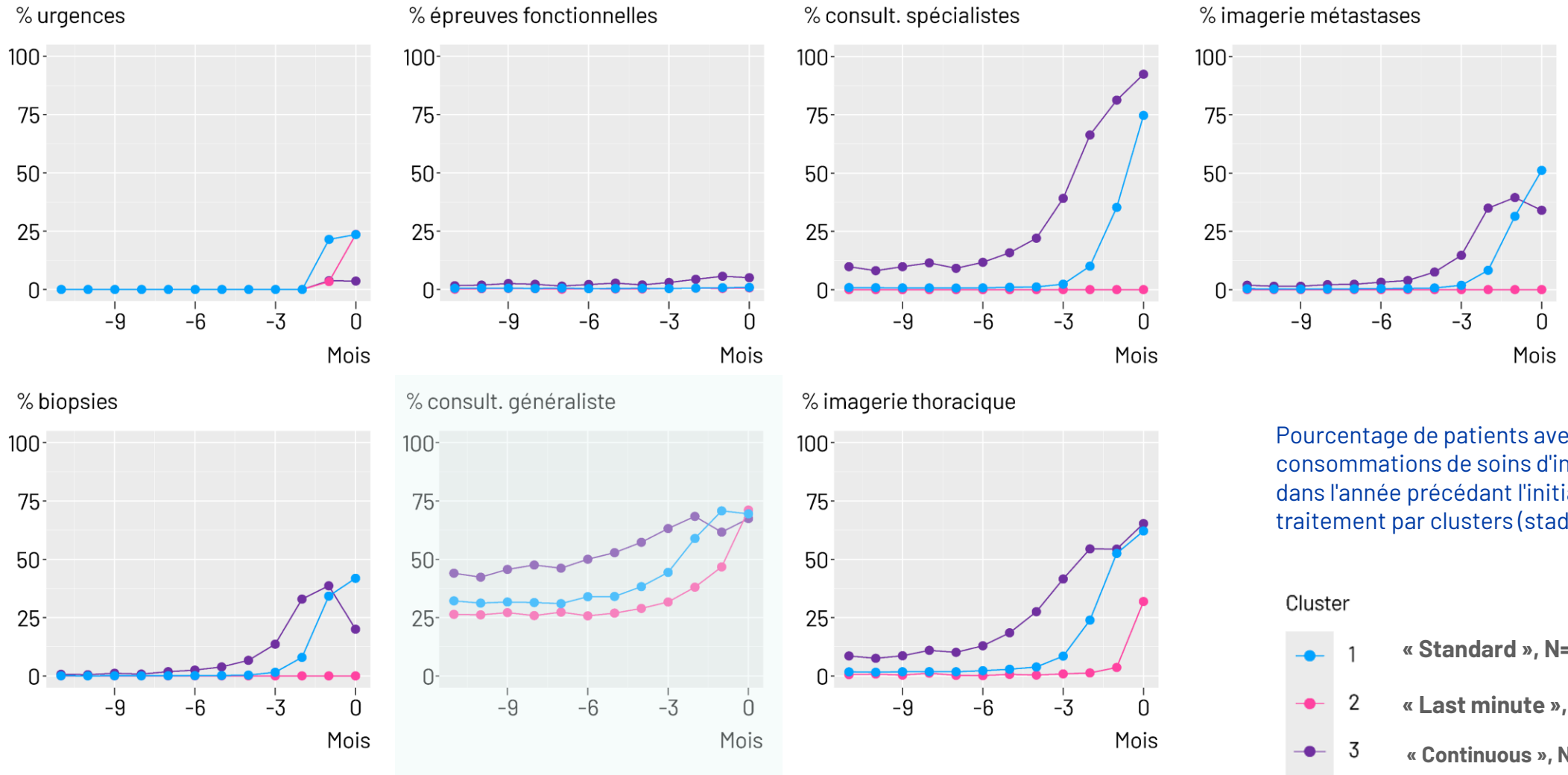
Descriptif global de la cohorte de patients – stades avancés / métastatiques

Variable	Indicator	Advanced/metastatic (N = 11,605)
Age (years)	Min–max	19–97
	Median (IQR)	65 (59–72)
Sex	Male	7,818 (67.4%)
	Female	3,787 (32.6%)
Charlson Comorbidity Index	0–2	141 (1.2%)
	3–4	1,472 (12.7%)
	5+	9,992 (86.1%)

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Stades avancés : 3 clusters identifiés qui diffèrent par leurs consommations de soins dans l'année précédant l'initiation du 1er traitement



Pourcentage de patients avec au moins une des consommations de soins d'intérêt chaque mois dans l'année précédant l'initiation du 1er traitement par clusters (stades avancés)

Cluster

1 « Standard », N=9 491 (81,8 %)

2 « Last minute », N=944 (8,1 %)

3 « Continuous », N=1 170 (10,1 %)

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Résultats clés : illustration sur les stades avancés/métastatiques

11 605 patients en stades avancés/métastatiques

→ 1 cohorte principale « standard »

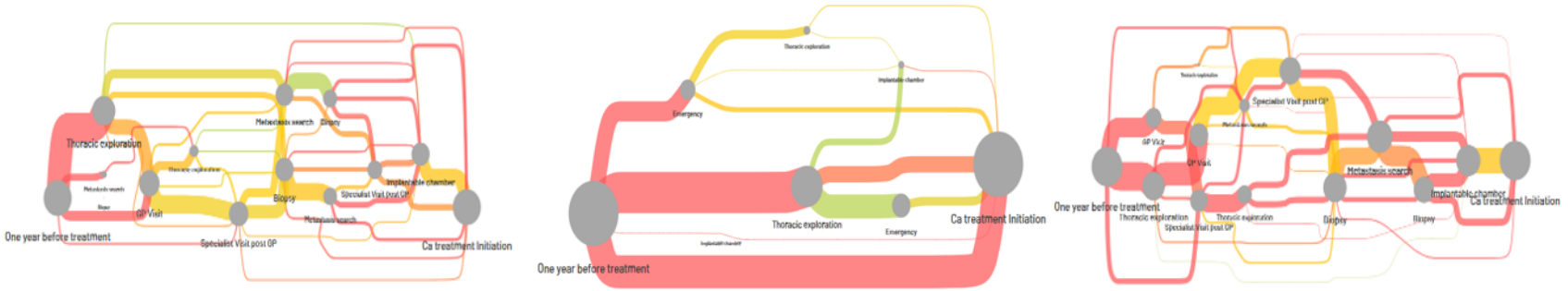
→ 1 cohorte avec diagnostic « last minute »

→ 1 cohorte avec un parcours diagnostic un peu plus long

	Cluster 1	Cluster 2	Cluster 3
N (% de la cohorte)	9 491 (**81,8** %)	944 (8,1 %)	1 170 (10,1 %)
Âge moyen (± sd)	64,9 ±10,1	64,5 ±10,5	**68,1 ± 9,3**
% d'hommes	67 %	65 %	**73 %**
Charlson moyen (± sd)	10,7 ±4,6	12,2 ±4,0	9,9 ±4,7
Délai médian (Q1; Q3) de mise sous traitement	35 jours [20-53]	**0 jours [0-4]**	58 jours [33-87]

Délai entre évènement clés du parcours *

- Délai court
- Délai médian
- Délai long



*définis en fonction de la médiane observée sur le groupe d'étude et pas par rapport à des recommandations

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Délais clés du parcours pré-traitement en fonction des clusters

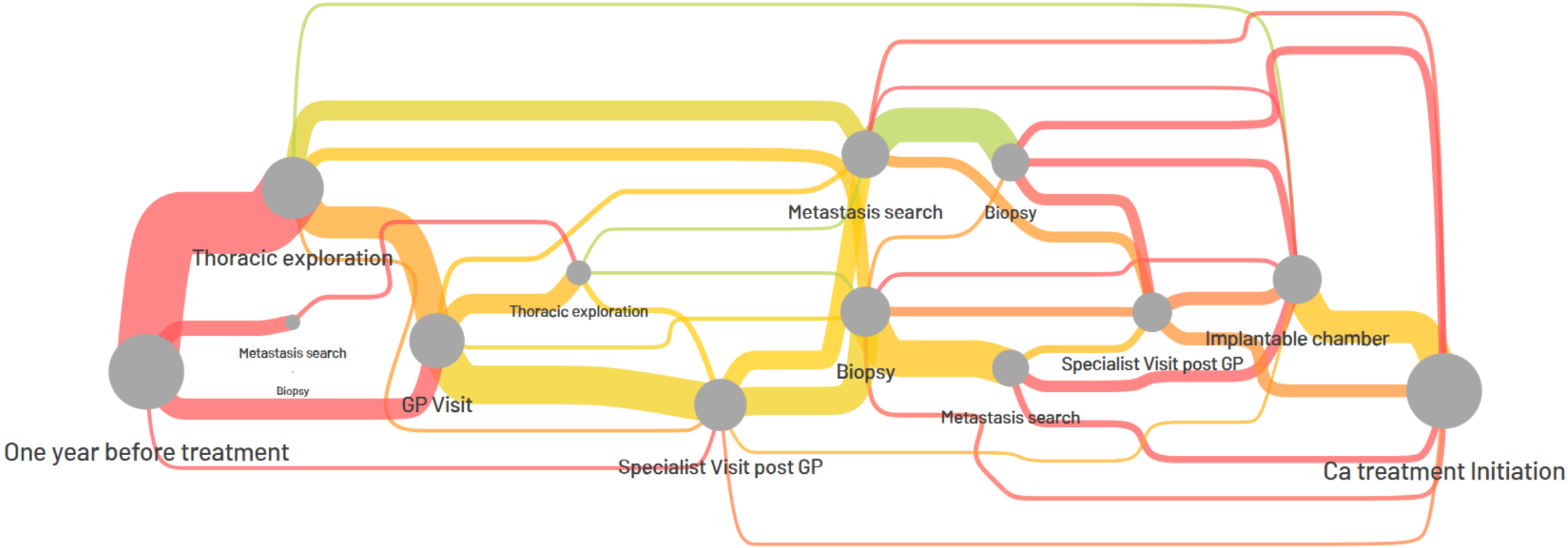
	Cluster « standard »	Cluster « last minute »	Cluster « continuous »
N (% de la cohorte)	9 491 (81,8 %)	944 (8,1 %)	1 170 (10,1 %)
Délai médian (Q1; Q3) **1ère biopsie -> Traitement**	31 (20 ; 46)	0 (0; 0)	70 (49 ; 101)
Délai médian (Q1; Q3) **1er scanner thoracique -> TT**	40 (25; 61)	0 (0; 0)	105 (72; 173)
Délai médian (Q1; Q3) **1er IRM cérébral -> TT**	27 (5; 57)	0 (0; 0)	100 (53; 193)
Délai médian (Q1; Q3) **Dernière consult. Généraliste -> Première IRM cérébrale**	9 (2; 24)	5 (0,5; 12)	18 (9; 6)

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Focus cluster 1 – cohorte « standard »

N = 9 491, rejouabilité = 86.6%

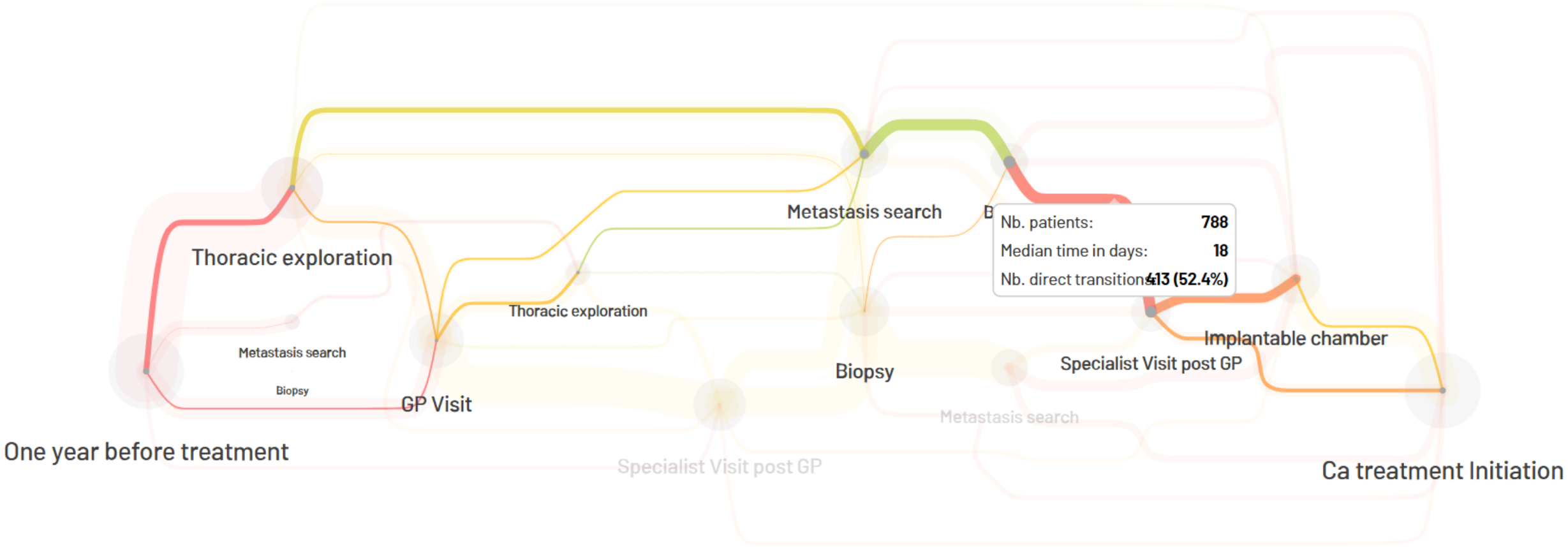


Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Focus cluster 1 – cohorte « standard »

N = 9 491, rejouabilité = 86.6%

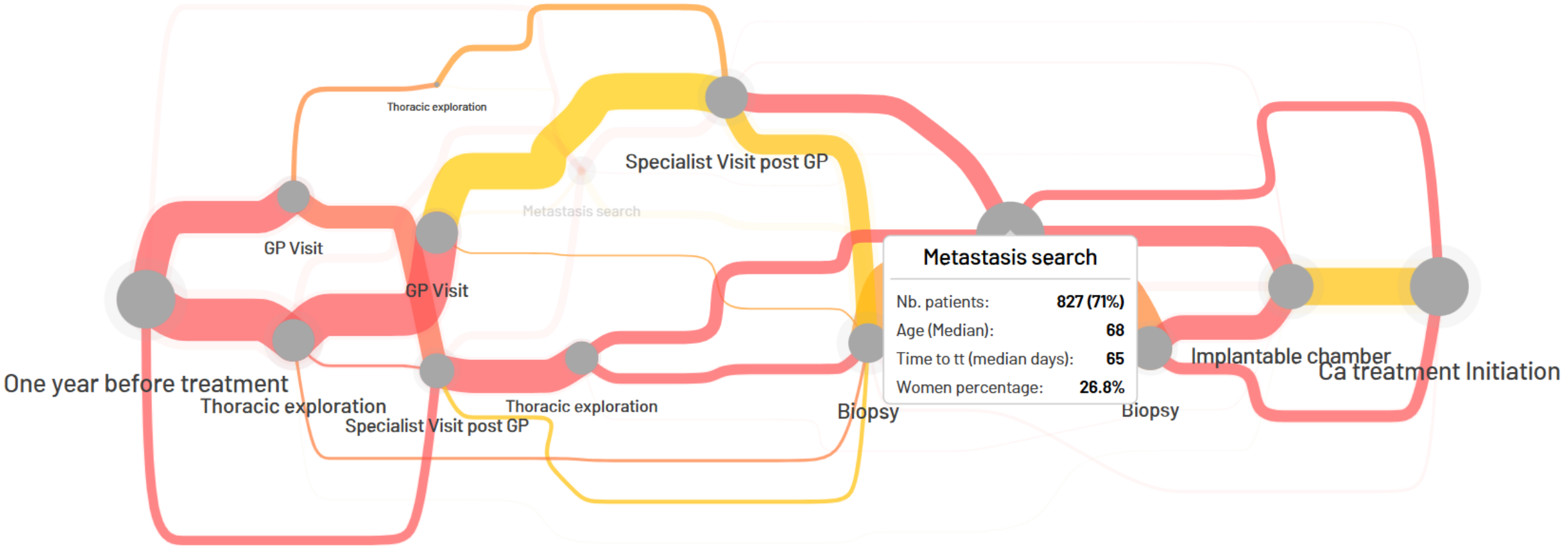


Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Focus cluster 3 – Continuous (late treatment initiation)

N=1 170, rejouabilité de 92,2%

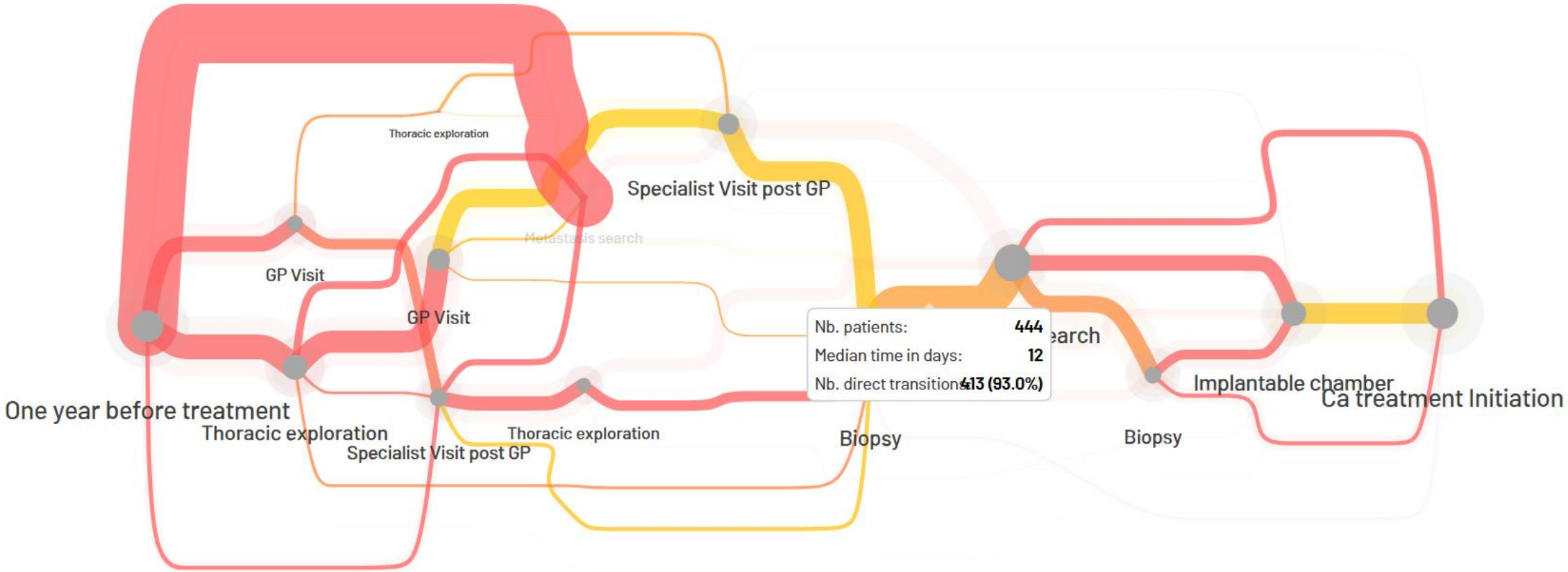


Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Focus cluster 3 – Continuous (late treatment initiation)

N=1 170, rejouabilité de 92,2%

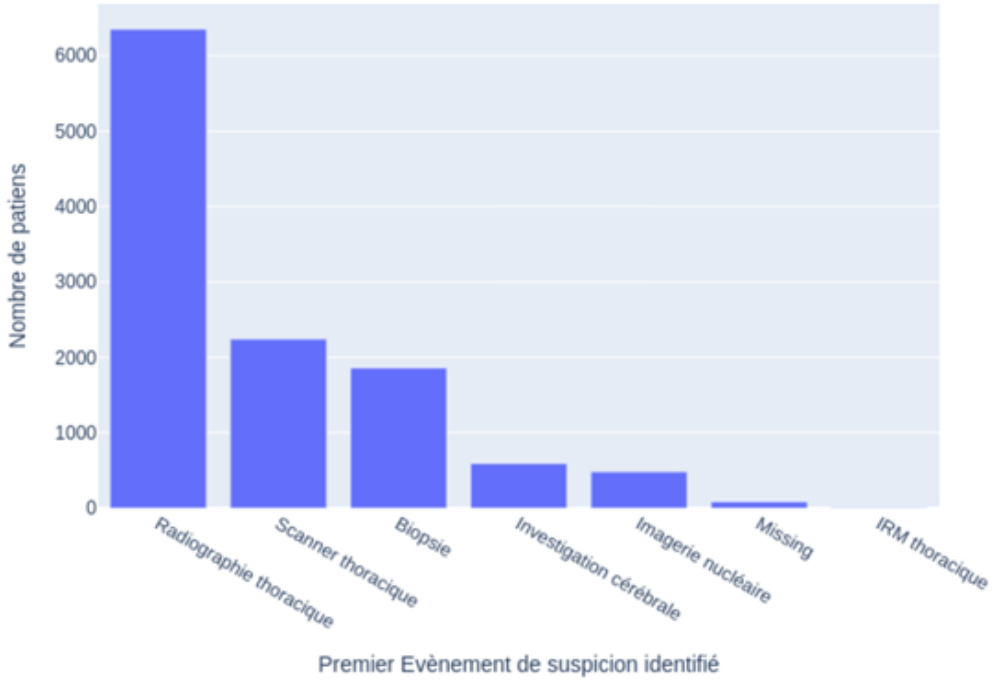


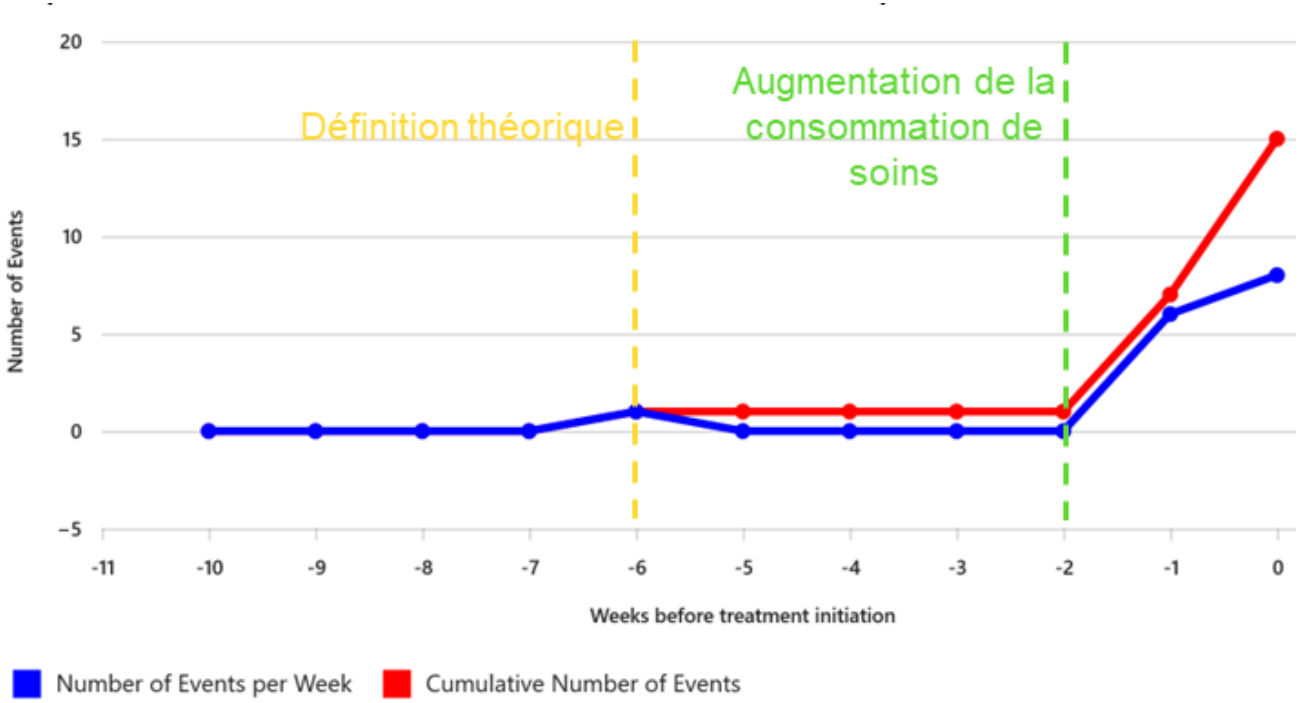
Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Exploration sur l'estimation du délai total avant initiation du traitement

1er évènement d'intérêt ou « rupture » de la consommation de soins





Premier examen pouvant aboutir à une découverte fortuite d'un cancer du poumon.

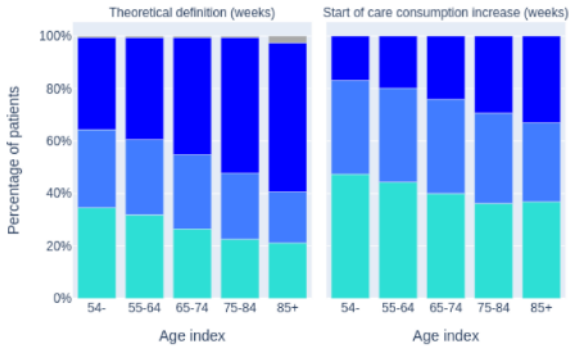
Changement de rythme de la consommation des soins en lien avec le cancer du poumon

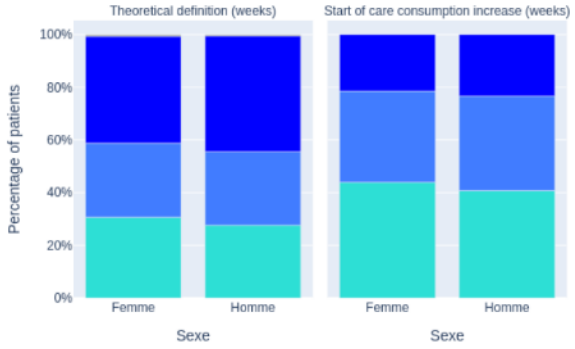
Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

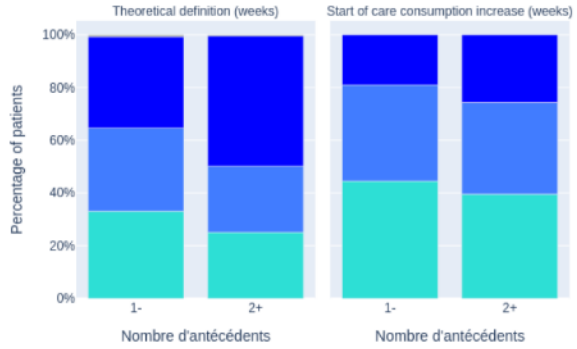
C2 - Restreint

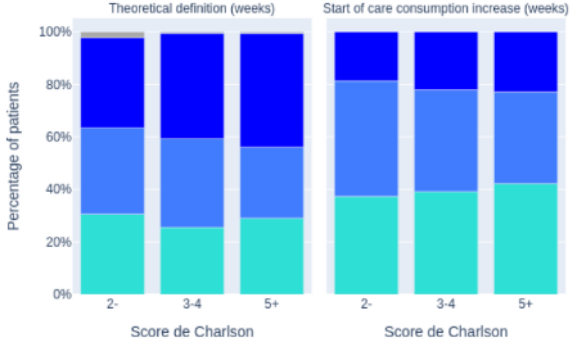
Délai de mise sous traitement

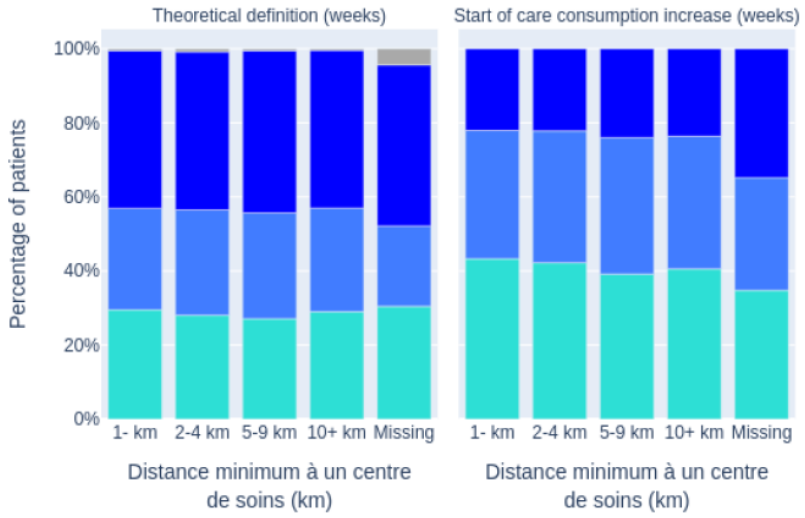
En fonction de différents paramètres d'intérêt

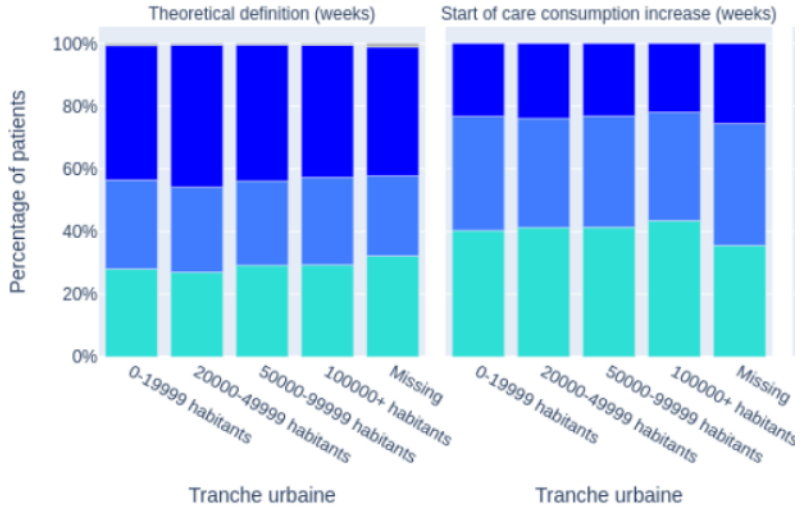


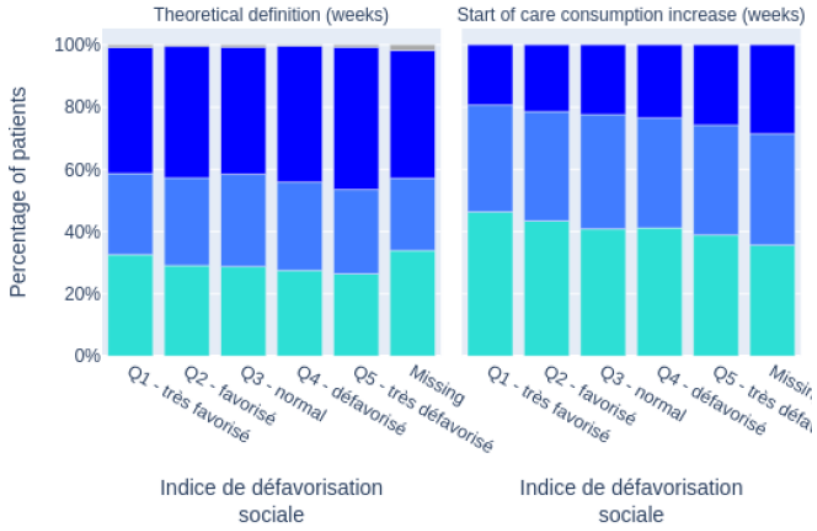


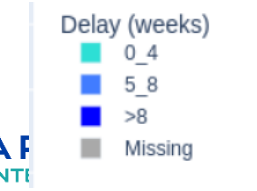












Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Consommations de soins lors du parcours de soins pré-traitement

Nombre moyen de remboursements de soins de santé pour les patients atteints d'un cancer local/localement avancé et d'un cancer avancé/métastatique au cours de l'année précédant leur inclusion dans l'étude (date de référence exclue)

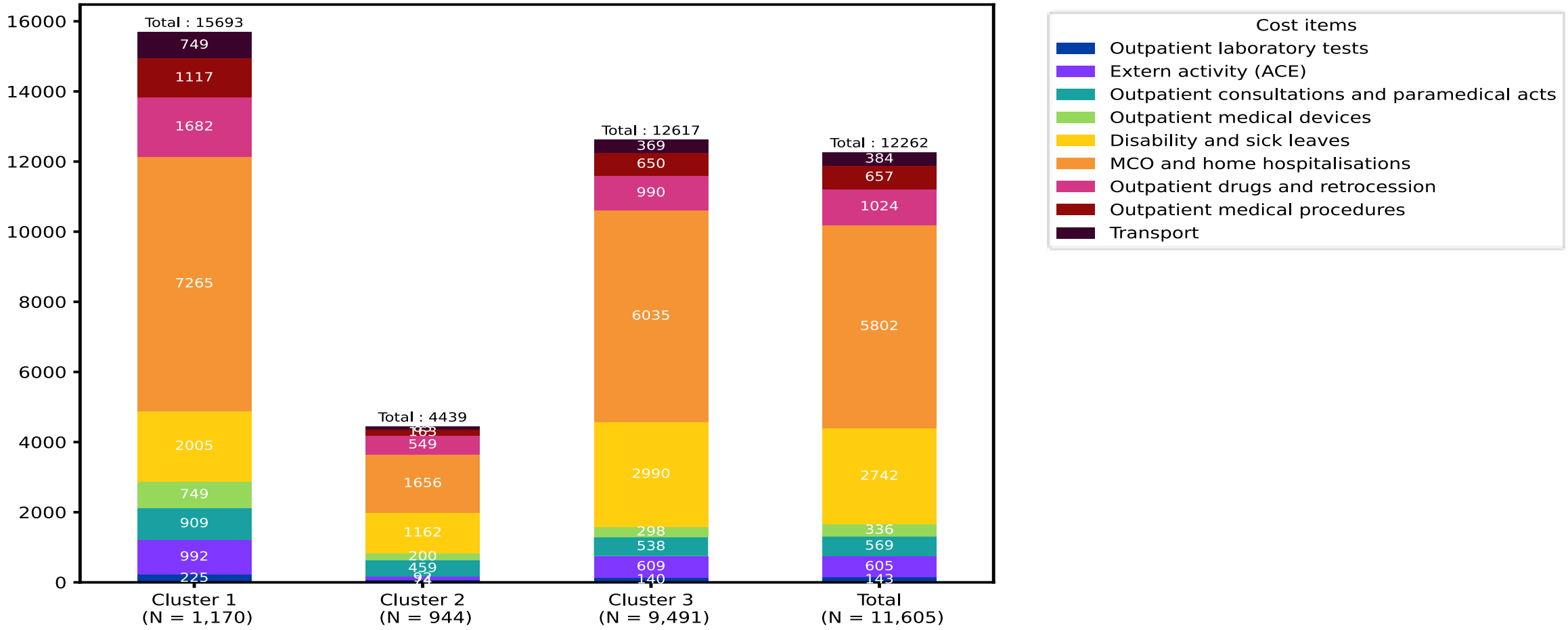
		Advanced / Metastatic (AMC) (N=11,605)		
		Cluster continuous	Cluster last minute	Cluster standard
Consultations outpatient general medicine	N	1,141	822	8,928
	Mean (± SD)	10.38 (6.39)	7.02 (4.89)	8.45 (5.43)
Consultations with an outpatient pulmonologist	N	1,143	29	6,676
	Mean (± SD)	5.32 (4.21)	1.00 (0.00)	2.77 (2.79)
Biopsies (all types combined)	N	1,029	633	7,752
	Mean (± SD)	1.86 (1.15)	1.08 (0.29)	1.39 (0.76)
Chest X-ray	N	1,129	779	8,303
	Mean (± SD)	3.41 (2.57)	1.35 (0.57)	2.26 (1.55)
CT scan and/or chest CT angiogram	N	990	495	5,862
	Mean (± SD)	1.78 (1.03)	1.12 (0.33)	1.30 (0.62)
PET scan	N	1,074	183	6,096
	Mean (± SD)	1.22 (0.48)	1.00 (0.00)	1.06 (0.24)

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Coûts des parcours pré-traitement

Coûts totaux moyens (en euros, 2023) des patients atteints d'un cancer avancé ou métastatique, en fonction des postes de coûts dans l'année précédant l'initiation du traitement

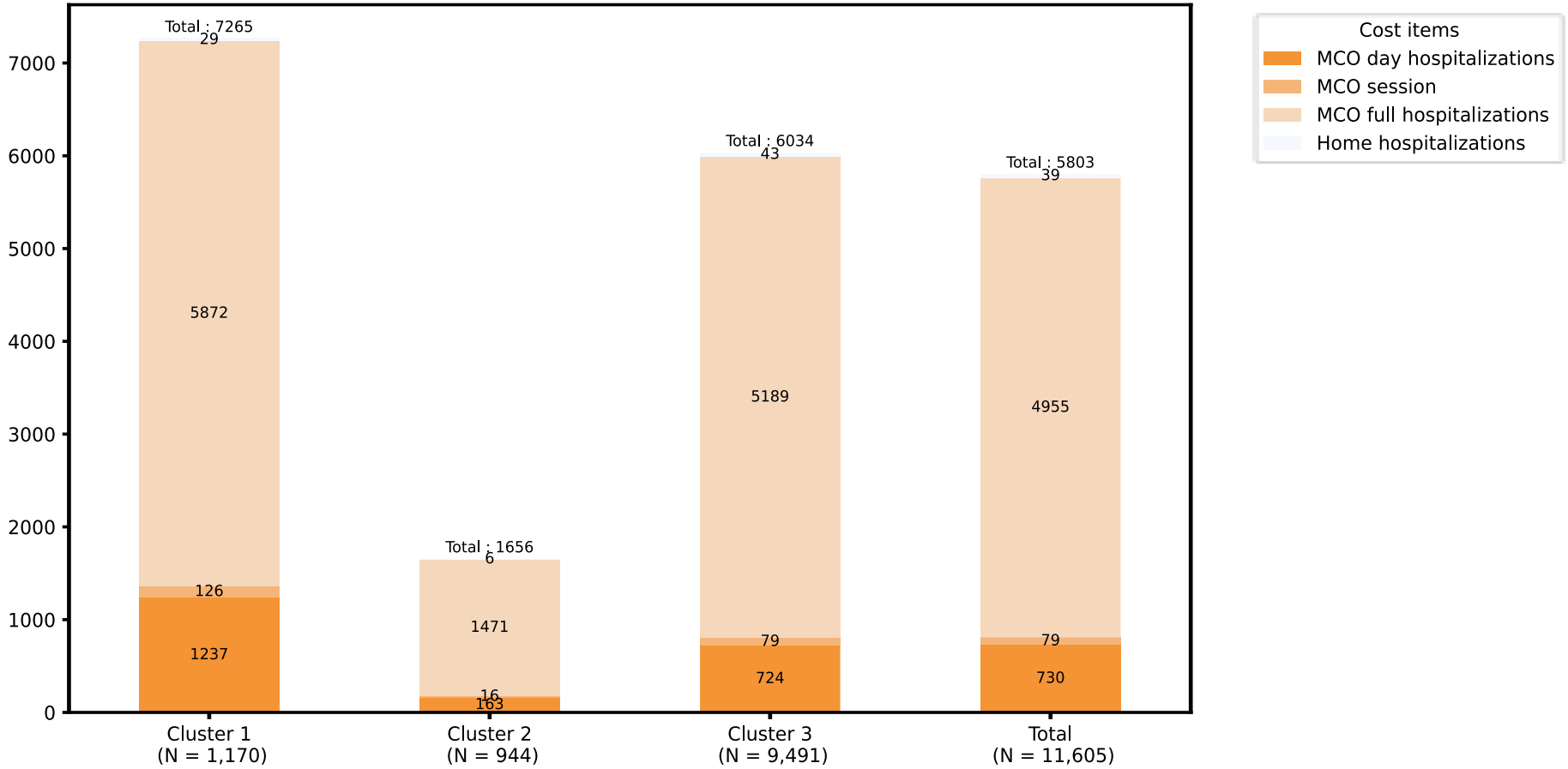


Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Coûts des parcours pré-traitement

Coûts totaux moyens (en euros, 2023) des patients atteints d'un cancer avancé ou métastatique, pour les coûts d'hospitalisation en MCO, en fonction des typologies d'hospitalisations dans l'année précédant l'initiation du traitement



Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Heva
A subsidiary of DOCAPOSTE

Conclusion

C2 - Restreint

LA POSTE
SANTÉ & AUTONOMIE

Conclusions

Les méthodes analytiques avancées et de l'IA ont permis de :

- **Visualiser de façon synthétique et interactive des parcours de soins pré-mise sous traitement**
- **Cartographier les parcours de santé au cours des 12 mois précédant le premier traitement**
- **Identifier des clusters des consommations de soins pré-traitement**

Un suivi spécifique à chaque trajectoire pourrait compléter les objectifs uniformes en matière de délai de prise en charge dans le cadre du traitement du cancer du poumon.

Ce document est la propriété exclusive de LPSA et ne peut être ni reproduit, ni communiqué à un tiers sans l'autorisation préalable de LPSA

C2 - Restreint

Merci pour votre attention

heva-data.com

linkedin.com/company/heva

Helene.denis@heva-data.com

C2 - Restreint