

RÉUTILISATION DES DONNÉES DE SANTÉ DES PATIENTS : RETOUR D'EXPÉRIENCE DE L'ICO

JOURNÉE RÉGIONALE ONCOPL
05/12/2024

Pr François Bocquet, PU-PH pharmacien
Responsable du service Data Factory & Analytics
(Centre des données numériques)



Contexte : vers une approche intégrée du cancer par le recours aux données massives de santé

LE CANCER : UNE PATHOLOGIE COMPLEXE

Pas une seule maladie mais un ensemble de maladies **rares** et **hétérogènes**.

Des **connaissances** sur la biologie des cancers **en progrès permanent**

Évolution des cancers vers des **pathologies chroniques**



La nécessité d'une approche intégrée (ou systémique) du cancer



MÉDICALE
BIOLOGIQUE
ÉCONOMIQUE

L'intérêt des données massives de santé (Big Data) :
appréhender le cancer dans toutes ses dimensions

SOCIALE
PSYCHOLOGIQUE
ENVIRONNEMENTALE



ENTREPÔTS DE DONNÉES DE SANTÉ HOSPITALIERS

Associer et agréger des données de formats et de sources hétérogènes à des fins d'optimisation des soins

Données multi-sources, multi-domaines et multi-échelles





Contexte : entrepôts de données de santé, un enjeu clé des centres de lutte contre le cancer



Avancées médicales majeures en oncologie

CONTINUUM SOINS – RECHERCHE : RECHERCHE CLINIQUE ↔ VRAIE VIE

La problématique : données d'essais cliniques ≠ données de vie réelle

		ESSAIS CLINIQUES	ÉTUDES DE DONNÉES DE « VIE RÉELLE »
Données cliniques	Origine	Cadre expérimental (standardisé, normé et contrôlé)	Hors cadre expérimental (générées par les soins de routine)
		Données colligées durant les essais	Comptes rendus d'hospitalisation, de consultation, données de biologie, prescription / administration, des médicaments, anatomopathologie.
	Structuration	Structurées 	Pas ou peu structurées 
	Qualité	Élevée	Faible
	Complétude		
Collecte	Manuelle 	Automatisé 	
Avantages		Qualité, homogénéité et complétude des données	Prise en charge « réelle », effectif et horizon temporel non restreints
Inconvénients		Effectif et horizon temporel restreints, souvent éloignés des soins de routine	Données non-structurées, hétérogènes et de qualité parfois faible.



Data Factory & Analytics (Centre des données numériques) : les objectifs

- **Entrepôt de données de santé**
 - **Continuum soins-recherche** : équipe dédiée, pluridisciplinaire et transverse.
 - **Faire converger les données administratives et médicales** (CR de consultation, d'hospitalisation, d'imagerie, résultats de biologie, d'anatomopathologie, etc...) issues des différents logiciels / applications utilisés par les professionnels de santé dans une seule et même base.
 - **Standardiser, mettre en qualité, organiser et rendre accessibles ces données** dans le cadre de projets de recherche.
 - **Améliorer l'accès à la donnée, sa qualité, son exhaustivité** dans le but de répondre aux besoins des utilisateurs.
 - **Gouvernance des données** → évolution du DPI.
 - **Guichet unique** professionnels de santé et chercheurs.
- **Études de données de vie réelle**

Une suite de solutions et projets

- Intégration des **sources de données structurées** (applis du SIH, ESME, RedCap, EPC, autres bases de pathologies adhoc ...) dans un EDS local
 - 280.000 patients à mai 2024 (depuis janvier 2014)
 - EDS « au fil de l'eau » (actualisation hebdo)
- **Expérimentation de structuration de données** par IA (NLP, LLM) – partenaires : Quinten, PoC Circular Lab, IMT, ARC
- **Déploiement outils de fouille de données / data mining** (CONSORE, Medexprim)
- Alimentation des entrepôts UNICANCER (ODH, ONCODS)
- Conduite du changement pour la structuration de la donnée par les producteurs (/ filière)

Projet ISIS : fertilité et cancer

- Incidence du cancer du sein en augmentation y compris chez les femmes jeunes : 5,000 à 6,000 femmes < 45 ans par an et mortalité en diminution

- Recul de l'âge de la première grossesse

PLAN CANCER 2021-2030

AXE 2 du plan cancer: limiter les séquelles et améliorer la qualité de vie

OBJECTIF :

Mieux connaître les séquelles liées à la maladie et les effets secondaires des traitements afin de les prévenir lorsque c'est possible et de les traiter le plus efficacement le cas échéant.

Mettre en place des ZU21 un appel à projets multithématiques au service des orientations de la stratégie (désescalade, médecine de précision, reconstruction, séquelles et qualité de vie, soins de support, préservation de la fertilité, études, emploi, accompagnement) et un programme de recherche clinique (FIRC-K) dédié à

PRÉSERVATION DE LA FERTILITÉ ET CANCER: RECOMMANDATIONS INCA 2021

R1. Il est recommandé d'informer tout enfant, adolescente (et leurs représentants légaux) et femme en âge de procréer du niveau de risque de gonadotoxicité et/ou d'insuffisance ovarienne prématurée en fonction du protocole de chimiothérapie envisagé dès la consultation d'annonce du diagnostic et/ou de la proposition thérapeutique (Grade A).

R2. Tout enfant, adolescente et femme en âge de procréer devant recevoir un protocole de chimiothérapie peut bénéficier d'une information spécialisée sur l'impact de ces traitements sur sa réserve ovarienne (Grade A).

R3. Il est recommandé de proposer de façon systématique une consultation de préservation de la fertilité pour toute patiente de moins de 38 ans devant recevoir un protocole de soins comprenant des agents alkylants bifonctionnels à haute toxicité ovarienne (Grade A). Entre 38 et 40 ans, il est recommandé de discuter au cas par cas l'indication de préservation de la fertilité (AE).

Description synthétique du projet

Objectif principal :

Evaluer le risque d'infertilité après cancer du sein de stade 1 à 3 chez des femmes âgées de 18 à 41 ans selon:

01

L'âge

02

Les traitements reçus

Type de chimiothérapie, thérapie ciblée et hormonothérapie

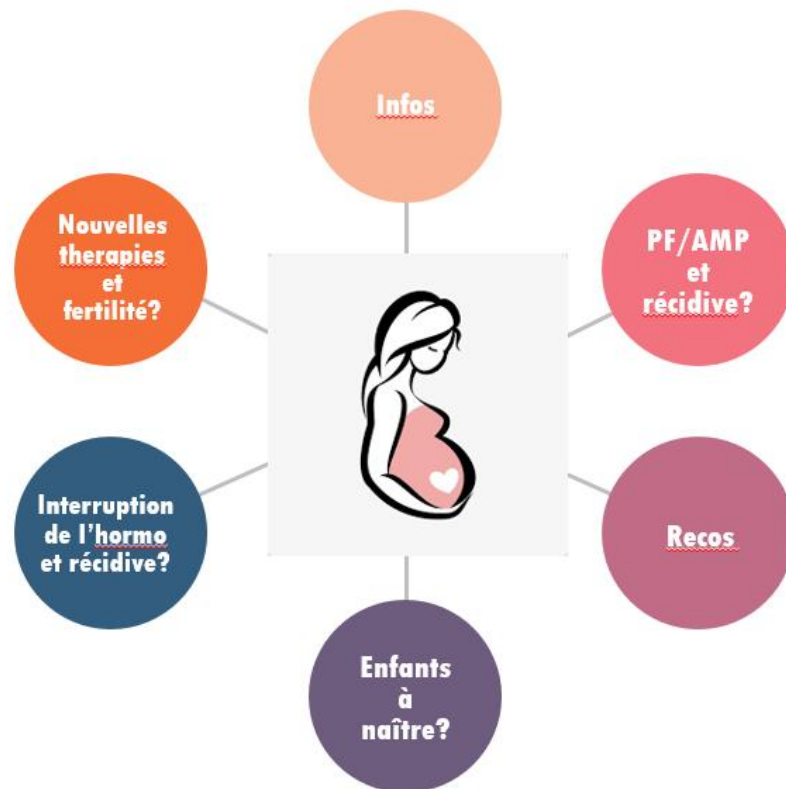
03

La prescription d'agonistes de la LH-RH

Pendant la chimiothérapie



Bénéfices pour les patients et la recherche



Caractère innovant du projet

Première **base de données dédiée à la fertilité** après cancer du sein.

Peu de données en vie réelle existent sur l'incidence des grossesses après traitement pour un cancer du sein en France.

Cette question nécessite un suivi sur le long terme rendu possible par ce projet avec possibilité de continuer le suivi des patientes tout juste traitées ou diagnostiquées.

Les données recueillies pourraient servir à étudier plusieurs questions sur la fertilité après cancer du sein :

- impact des nouvelles thérapies sur la fertilité
- impact des techniques de préservation de la fertilité sur la récurrence
- impact de la PMA après cancer du sein sur la récurrence
- impact des traitements sur les grossesses
- impact de la durée de la fenêtre thérapeutique du tamoxifène sur la récurrence, etc...

Méthodologie du projet

4 centres participants

Porteur : Institut de Cancérologie de l'Ouest (ICO)

Période 2000-2022

TOTAL



CENTRE
DE
CANCÉROLOGIE
LEON
BERARD



Centre
Oscar Lambret

• Patientes âgées de 18-41 ans	11 846	2 000	1 900	6 388	1 558
• Diagnostic de cancer du sein localisée I-III					



Évaluation du risque d'infertilité après cancer du sein de stade 1 à 3 chez des femmes âgées de 18 à 41 ans



Impact de l'âge, des **traitements anticancéreux** (chimiothérapie, thérapie ciblée, hormonothérapie) et **analogues LH-RH pendant la chimiothérapie** sur la fertilité

Méthodologie du projet

4 centres participants

Porteur : Institut de Cancérologie de l'Ouest (ICO)

Période 2000-2022

TOTAL



Centre Oscar Lambret

• Patientes âgées de 18-41 ans	11 846	2 000	1 900	6 388	1 558
• Diagnostic de cancer du sein localisée I-III					

- Collecte de données avec contrôle de cohérence et de fiabilité
- Pseudonymisation des données
- Intégration et appariement des données sur le HDH
- Analyse statistique des données (plan d'analyse statistique robuste validé)

Données utilisées dans le cadre du projet

Données hospitalières multi-sources (ConSoRe et entrepôts de données) sur la période 2000-2022

Données démographiques des patients

Données PMSI : diagnostic, actes médicaux, dates de séjours hospitaliers

Données d'anatomopathologie : RH+/HER2- ; triple négatif ; RH+/HER2+ ; RH-/HER2+ ; TNM

Dates de diagnostic (date CR anatomopathologie tumeur primitive)

Prescriptions et administrations médicamenteuses (chimiothérapie, thérapie ciblée, hormonothérapie, analogues LH-RH)



Données SNDS sur la même période (extraction du jeu de données par la CNAM)



Consommations en anticancéreux et analogues LH-RH ville et hôpital : dates début et de fin (base SNIIRAM et PMSI)

Grossesse (O/N), dates de début et fin de grossesse, données préservation fertilité, césarienne (O/N) (base PMSI)

Appariement données hospitalières et SNDS

Si appariement déterministe direct : NIR, dates de naissance, sexe

Si appariement déterministe indirect : GHS, date d'entrée et sortie, code postal

