



**RÉPUBLIQUE
FRANÇAISE**

*Liberté
Égalité
Fraternité*



Comité Scientifique et Éthique Plateforme de Données en Cancérologie

AVIS D'ÉVALUATION

Session du 22/09/2023

Titre	LUCIEN Etude non-interventionnelle, multisource sur les délais et les coûts des parcours diagnostiques du cancer du poumon primitif – analyse de processus (process mining)
Numéro	CSE PDC 23005-01
Transmission	01/09/2023

Avis	<input type="checkbox"/> Avis favorable <input checked="" type="checkbox"/> Avis favorable sous conditions : les réponses aux remarques permettront de délivrer un avis favorable <input checked="" type="checkbox"/> Réponses validées – Avis favorable le 30/10/2023 <input type="checkbox"/> Avis réservé : le dossier sera revu en commission après réponses aux réserves <input type="checkbox"/> Avis défavorable
------	---

Remarques associés à l'avis rendu

52, avenue André Morizet
92100 Boulogne-Billancourt - France
Tél.: +33 (0) 1 41 10 50 00 - e-cancer.fr
N° SIRET: 187 512 777 000 33
Code APE: 8412 Z

Sur le plan éthique

Le projet ne pose pas de problématique éthique spécifique concernant les patients ou les professionnels impliqués. Les données ne seront pas accessibles à l'industriel. Seuls l'INCa et le prestataire responsable des analyses pourront y accéder.

Le comité attire cependant l'attention des demandeurs sur l'interprétation des résultats issus de ce projet, qui seront pour certains très exploratoires. Leur communication doit traduire ces limites et toutes les précautions associées quant à leur fiabilité. Le vocabulaire utilisé dans le projet doit suivre ces précautions : le terme « expliquer » est impropre dans ce contexte, sous-entendant un lien de causalité qui ne peut être établi à partir de ce type d'étude. Il faudrait lui préférer la recherche d'association par exemple.

Sur les aspects portant sur l'information des sujets

La fiche projet devra être mise en conformité avec les remarques scientifiques et méthodologiques qui suivent.

Sur le plan scientifique et méthodologique

Le comité perçoit 3 finalités stratégiques distinctes dans le projet au travers des objectifs présentés et des données utilisées :

- 1) Apporter des éléments de réponses sur les parcours de la population d'intérêt : descriptif, facteurs associés...
- 2) Mettre en place et valider l'appariement des données de la plateforme de données en cancérologie (PDC) avec les comptes rendus de RCP et de génétiques moléculaires. Il s'agit du premier projet portant sur ces données, qui devraient être prochainement intégrées à la PDC. Il joue donc un rôle de validation de cette intégration.
- 3) Explorer les possibilités offertes par ces documents nouvellement intégrés à la PDC. Il s'agit donc d'un travail qui ne peut être qu'exploratoire, en raison de l'hétérogénéité de disponibilité et de qualité de ces documents selon les régions.

Le dossier soumis au CSE, se concentre sur la première finalité, présentant les 2 suivantes comme des moyens d'y parvenir. Il serait souhaitable au contraire, de faire apparaître les éléments de faisabilité, et les développements méthodologiques comme des objectifs assumés du protocole. S'agissant d'un premier projet exploitant ces données, il servira pour l'élaboration de protocoles ultérieurs. La faisabilité de l'appariement des données, la validation, la quantification de données manquantes, la description des données disponibles, leur hétérogénéité... sont autant d'informations qui serviront à adapter les futures études exploitant ces données.

Le comité souligne l'ancienneté des données exploitées, limitant la portée des résultats obtenus lorsque viendra le temps de la publication : l'étude s'intéresse aux patients traités en

2018 et 2019, et le calendrier de réalisation de l'ensemble de ces analyses d'une part, mais surtout de l'intégration des données de RCP et de génétique moléculaire est long, et incertain. De plus, la structuration et la généralisation des compte rendus de RCP et de génétiques moléculaires, est relativement récente, et les documents existants en 2018 n'avaient pas la qualité et l'exhaustivité attendues :

Concernant plus spécifiquement l'accès aux Fiches RCP, le comité souligne des éléments pouvant entraîner une certaine difficulté à la réalisation du projet :

La disponibilité des fiches RCP dans les différentes régions : Les fiches RCP sont actuellement disponibles de manière hétérogènes entre les régions. Certaines régions disposent d'un Dossier Communicant en Cancérologie (DCC) porté par le Dispositif Spécifique Régional du cancer (DSRP) comportant l'exhaustivité de fiches RCP, certaines régions ont un DCC porté par le DSRC et comportant une partie des fiches RCP, les autres étant disponibles dans les Dossiers Patients Informatisés des établissements et enfin dans certaines régions il n'existe pas encore de solution DCC régionale portée par le DSRC.

L'interface entre les DCC et la PDC : Actuellement aucun DCC n'est interfacé avec la PDC. Des travaux sont en cours pour ce faire, et nécessiteront des phases de développement informatique et l'intégration de l'INS. Ceci entraîne donc que cette étude ne pourra se réaliser que sur des données provenant de régions « pilotes » qui seront à déterminer en fonction de l'état d'avancement de leurs solutions DCC. Le comité souligne que ces régions « pilotes » devront être sollicitées en amont et retenues sur la base du volontariat étant donnée la charge de travail qui pourrait ainsi incomber aux DSRC

Qualité de remplissage des Fiches RCP : Le comité rappelle que les fiches RCP disponibles dans le DCC sont des documents synthétiques ayant pour objet de permettre aux médecins de la RCP de disposer des informations utiles et pertinentes à la proposition thérapeutique à l'instant donné. Les fiches ne sont pas à considérer comme des fiches d'enquêtes épidémiologiques et peuvent comporter des éléments partiels et non exhaustifs. Ceci est d'autant plus vrai pour les fiches les plus anciennes, la qualité de remplissage de fiches s'améliorant d'année en année.

Ces limites du projet ne sont pas abordées par les demandeurs. Elles restreignent la qualité des informations produites sur les parcours des patients, mais ne remettent pas en question les développements méthodologiques, et l'élaboration des appariements des données.

Par ailleurs, le comité souhaite aborder les points suivants qui nécessiteraient d'être pris en compte :

- Il est fait mention d'un échantillon de données de RCP qui seraient utilisées pour le projet, sans que cet échantillonnage ne soit détaillé par ailleurs. Ce point est à préciser. S'il s'agit d'une terminologie pour indiquer l'absence d'exhaustivité des comptes rendus de RCP dans la plateforme, ce terme est impropre. S'il s'agit d'un réel travail

d'échantillonnage, ce point doit être développé (pourquoi, sur quel critère, quel représentativité...).

- Le projet ne permet pas d'inclure dans l'analyse des parcours les patients pour lesquels un 1^{er} traitement du cancer ne peut être identifié. Ce sous-groupe mériterait tout de même d'être quantifié et décrit : il s'agit probablement de patients très spécifiques, il pourrait être utile de caractériser les patients qui ne peuvent être pris en compte dans l'analyse.
- Les demandeurs indiquent que les méthodes de traitement du langage naturel peuvent être appliqués aux comptes rendus (peu structurés) prochainement intégrés à la plateforme. Il n'est pas précisé si ces méthodes seront effectivement mises en œuvre dans le cadre de ce projet de recherche, ni dans quel objectif ou l'identification de quelles informations il pourrait y recourir.

