

DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **03 juillet 2024**

Nom de famille et prénom de l'auteur : **Madame ANTOINE Alison**

Titre de la thèse : « *Utilisation des données de vie réelle pour estimer l'effet traitement en cancérologie. Application à l'étude de la survie des patientes prises en charge pour un cancer du sein métastatique* »



Résumé

L'essai contrôlé randomisé (ECR) est le schéma expérimental de référence pour évaluer l'effet d'un nouveau traitement. En oncologie, les modes de prise en charge des patients évoluent rapidement avec l'avancée des connaissances sur de nouveaux mécanismes d'action des molécules et le développement de thérapies ciblant des sous-groupes rares de patients. Ainsi, il peut parfois être difficile, voire impossible de réaliser dans des délais raisonnables un ECR dans des sous-populations de cancers définies par la présence d'une mutation, altération ou délétion rare. De fait, le recours à des essais monobras, non contrôlés s'impose souvent lors du développement des molécules concernées. Cependant, en l'absence de bras contrôle, ce schéma d'étude ne permet qu'une démonstration partielle et peu robuste de l'effet traitement. L'émulation d'essai cible est un cadre méthodologique consistant à appliquer les principes de conception des ECR à des données observationnelles lorsqu'un ECR ne peut pas être réalisé pour des raisons éthiques ou logistiques. L'analyse des données de vie réelle *via* l'émulation d'essai cible offre de nouvelles perspectives pour générer des preuves de haut niveau à partir de larges bases de données. Une démonstration rétrospective de la capacité de l'émulation d'essai cible à reproduire l'effet traitement des ECR en oncologie s'avérerait nécessaire. A titre d'application, cette thèse visait à émuler jusqu'à 14 ECR existants pré-identifiés, à savoir des études pivots évaluant des innovations thérapeutiques de rupture dans le cancer du sein métastatique (CSM) depuis les années 2000. Des techniques d'ajustement avancées telles que la pondération par l'inverse de la probabilité d'être traité ou la g-computation ont été utilisées pour équilibrer les bras de traitement *a posteriori* et prendre en compte le biais de confusion. Le biais résiduel a été exploré à travers des analyses quantitatives du

biais par simulation d'un facteur de confusion hypothétique non mesuré et en calculant la E-value qui est la force d'association minimale qu'un facteur de confusion non mesuré devrait avoir de manière simultanée avec le traitement et le résultat pour invalider les résultats. Il a été possible d'émuler 9 des 14 ECR pré-identifiés, les 5 ECR restants n'ont pu être réalisées en raison d'une taille d'échantillon insuffisante. L'effet traitement estimé dans chacun des 9 ECR servait de référence pour juger de la fiabilité globale de la méthodologie employée en mesurant la concordance des estimations entre l'ECR et l'émulation. Dans l'ensemble, le niveau de concordance obtenu était satisfaisant, bien que des divergences liées à des facteurs de confusion résiduels et à des changements significatifs dans les pratiques de prescription après l'approbation du médicament ont parfois pu être observées. Des analyses de contrôles négatifs étaient également prévues mais n'ont pas pu être réalisées, faute d'avoir pu obtenir les autorisations réglementaires dans les délais requis. En effet, un chaînage des données ESMÉ CSM aux données du Système National des Données de Santé était nécessaire afin d'accéder à des données hors cancer pour réaliser les analyses de contrôles négatifs. Ces analyses seront néanmoins réalisées à l'issue de cette thèse. En conclusion, ce travail a établi que l'émulation des ECR dans le CSM à l'aide de la base de données ESMÉ, combinée à un ajustement approprié, pouvait être une approche prometteuse pour étudier l'impact à long terme des thérapies innovantes dans la cohorte ESMÉ-CSM tout en minimisant les biais. Elle offre ainsi des opportunités pour évaluer l'efficacité comparative des thérapies dans des situations où un ECR ne peut être réalisé, notamment en utilisant des bras de contrôle externes issus de données de vie réelle.