

## DIPLÔME NATIONAL DE DOCTORAT

(Arrêté du 25 mai 2016)

Date de la soutenance : **18 décembre 2024**

Nom de famille et prénom de l'auteur. e : **Madame Camille GIAMPICCOLO**

Titre de la thèse : Exposition chronique et combinée à de multiples polluants de l'air et risque de cancer du sein dans la cohorte E3N : application d'approches statistiques multi-expositions

### Résumé



La pollution atmosphérique est un mélange complexe de plusieurs polluants fortement corrélés, dont la composition varie à la fois dans l'espace et dans le temps. Classée cancérogène par le CIRC, cette pollution est suspectée d'être associée à un risque positif de cancer du sein, la première cause de décès par cancer chez les femmes. Cependant, les méthodes épidémiologiques classiques ne permettent pas d'isoler les effets spécifiques de chaque polluant en raison des fortes corrélations entre les différentes variables d'exposition, ni d'évaluer l'impact combiné de l'exposition à plusieurs polluants sur la santé. L'objectif de cette thèse est d'explorer des méthodes statistiques permettant d'étudier l'effet de l'exposition combinée à de plusieurs polluants atmosphériques, fortement corrélés sur le risque de cancer du sein, tout en identifiant les contributions individuelles de ces polluants et les interactions potentielles entre eux. Plusieurs approches ont été retenues. La première consiste à développer un indicateur de multi-expositions pondéré sur des données toxicologiques relatives aux mécanismes biologiques, puis à évaluer l'association entre cet indicateur et le risque de cancer du sein à l'aide d'une régression logistique conditionnelle. Par ailleurs, deux méthodes, Weighted Quantile Sum et Quantile G-Computation, permettant d'évaluer les effets d'expositions à un mélange de polluants ainsi que les contributions individuelles de ces polluants sur la santé, ont été retenues. Ces deux méthodes se distinguent par leur formulation et les hypothèses sur lesquelles elles reposent. Une autre méthode sélectionnée est Bayesian Kernel Machine Regression, un modèle semi-paramétrique qui permet non seulement de prédire de manière flexible la relation dose-réponse, mais aussi de caractériser les interactions statistiques bivariées entre les polluants. Enfin, la dernière méthode sélectionnée est la Bayesian Profile Regression, qui regroupe les individus en fonction de leur niveau d'exposition et de leur niveau de risque, puis associe un risque à chacun de ces groupes. Ces méthodes ont été appliquées sur une étude cas-témoins

appariée nichée dans la cohorte française E3N, comprenant 5222 cas incidents de cancer du sein et 5222 témoins appariés. Pour chaque femme, la moyenne de l'exposition annuelle à huit polluants, incluant le benzo(a)pyrène, le cadmium, les dioxines, les polychlorobiphényles 153, le dioxyde d'azote, l'ozone, les matières particulaires et les particules fines, a été estimée depuis l'inclusion dans la cohorte en 1990 jusqu'à la date d'index. Globalement, ce travail de thèse a révélé un effet significatif positif de l'exposition combinée à plusieurs polluants sur le risque de cancer du sein chez toutes les femmes. Bien que nous ayons anticipé des odds ratios plus élevés associés au mélange de plusieurs polluants, les odds ratios observés pour l'effet global du mélange, évalués au sein de toute la population de notre étude, sont comparables à ceux rapportés dans les études sur les effets des polluants pris individuellement. Cependant, nous avons observé des odds ratios plus élevés chez les femmes fortement exposées au dioxyde d'azote, aux particules et aux polychlorobiphényles 153, vivant dans des zones urbaines telles que Paris et Lyon. Ce travail de doctorat a mis en évidence l'importance d'une compréhension approfondie des propriétés et des limitations des modèles, afin de prendre des décisions appropriées en fonction des caractéristiques du mélange étudié et du design de l'étude.

**Mots-clés :** polluants de l'air, cancer du sein, statistique, mélange,