# Apport de l'estimation non paramétrique de l'importance des variables dans la recherche de biomarqueurs pronostiques en grande dimension pour le cancer du col de l'utérus.

S. Cyrille, X. Paoletti et **A. Latouche\***

**(\*Institut Curie, Inserm U1331 et Conservatoire national des arts et métiers)**

**Contexte**

Dans la recherche de biomarqueurs pronostiques, l'importance des variables (VIMP) permet d'identifier les marqueurs moléculaires les plus pertinents à partir de données multi-omiques, telles que la génomique et la transcriptomique. Toutefois, les mesures traditionnelles de l'importance variable dépendent du modèle d’analyse, ce qui entraîne des problèmes de comparabilité. Pour remédier à cette limitation, Williamson et .al ont proposé une méthode non-paramétrique pour l'estimation de la VIMP dans le cas où la variable à prédire est binaire. S'appuyant sur cette méthode et en employant une méthode de staking pour les données de survie, Wolock et al. ont étendu cette VIMP non-paramétrique à des données de survie (censurées à droite).

Nous avons analysé les données de l'étude BioRAIDs en utilisant ce cadre de VIMP pour les données de survie afin d'identifier des biomarqueurs pronostiques pour la survie sans progression chez les patientes atteintes d'un cancer du col de l'utérus.

Nous détaillerons les points clés pour mettre en oeuvre la méthode de Wolock qui nécessite l’estimation des espérances conditionnelles de la réponse (censurée) conditionnellement aux caractérisques et des paramètres de nuisance comme la distribution de la censure. Afin de se prémunir d’un problème de mauvaise spécification, nous utilisons des méthodes d’apprentissage de type Super Learner (aggégation d’algorithmes).

La valeur incrémentale de deux voies génétiques pour la prédiction d’un critère de jugement censuré illustre l'intérêt de la VIMP non-paramétrique dans un contexte multi-omique et améliore l’interprétabilité des potentiels candidats.

[Williamson 2020] Williamson, Brian D., Gilbert, Peter B. Carone Marco , and Simon Noah.

2020. Nonparametric Variable Importance Assessment Using Machine Learning

Techniques. *Biometrics* 77 (1): 9–22. https://doi.org/10.1111/biom.13392.

[Wolock 2025] Wolock, Charles J. and Gilbert, Peter B. and Simon, Noah and Carone, Marco.

2025. Assessing variable importance in survival analysis using machine learning

Biometrika, Volume 112, Issue 2, 2025, asae061, <https://doi.org/10.1093/biomet/asae061>