

Entreprises et laboratoires d'accueil

- **Institut de Radioprotection et de Sûreté Nucléaire (IRSN)**
Laboratoire d'épidémiologie des rayonnements ionisants (LEPID)
BP 17
F-92262 Fontenay-aux-Roses Cedex
- **Université Paris Descartes,**
Equipe BioSTM « Biostatistique, traitement et modélisation des données biologiques »
4, avenue de l'Observatoire
75006 Paris

Contacts :

Sophie Ancelet sophie.ancelet@irsn.fr
Tel : 01.58.35.79.89
Chantal Guihenneuc chantal.guihenneuc@parisdescartes.fr

Intitulé du sujet

Approche hiérarchique bayésienne pour la prise en compte d'erreurs de mesure partagées dans les études de cohortes professionnelles. Application en épidémiologie des rayonnements ionisants.

Contexte

Les erreurs de mesure d'exposition constituent l'une des sources d'incertitude les plus importantes en épidémiologie. Lorsqu'elles ne sont pas ou mal prises en compte, elles peuvent mener à un biais dans les estimations de risque ainsi qu'à une déformation des relations exposition-risque d'intérêt [1]. Plus généralement, cela peut contribuer à diminuer la puissance statistique des études épidémiologiques et ainsi, limiter l'interprétation et la portée des résultats de ces études. Ainsi, la question de la prise en compte, dans les analyses épidémiologiques, des incertitudes associées aux données d'exposition radiologique connaît une véritable dynamique actuelle au sein du laboratoire d'épidémiologie (LEPID) de l'IRSN et, plus largement, au niveau international.

L'une des principales raisons pour lesquelles les erreurs de mesure d'exposition sont rarement prises en compte dans les études épidémiologiques est que les méthodes classiques de correction d'erreurs de mesure (e.g., régression-calibration, méthode SIMEX...) manquent souvent de flexibilité lorsqu'il s'agit de prendre en compte des erreurs de mesure complexes telles que celles rencontrées, par exemple, dans les études de cohortes professionnelles. Dans ces études, le type et la magnitude des erreurs de mesure peuvent changer au cours du temps en fonction des techniques utilisées pour évaluer l'exposition des travailleurs. Par ailleurs, des évaluations collectives de l'exposition à travers des capteurs d'ambiance situés dans un lieu spécifique (et donc, non portés individuellement par un travailleur) ou encore des matrices emplois-expositions peuvent donner lieu à des erreurs de type Berkson combinées à des erreurs de type classique et potentiellement partagées entre plusieurs travailleurs ou au cours de la période de suivi d'un même travailleur [2]. Cette notion d'erreurs de mesure partagées implique l'existence de dépendances entre les mesures dont il faut tenir compte.

L'utilisation d'une approche hiérarchique bayésienne, connue pour sa souplesse et sa pertinence pour la prise en compte d'erreurs de mesure complexes, a fait l'objet de différents travaux de recherche au LEPID ces dernières années [2, 3, 4]. Cette approche est fondée sur l'inférence bayésienne d'un modèle probabiliste composé de plusieurs sous-modèles conditionnellement indépendants: (a) un modèle de maladie (e.g., modèle de survie) décrivant la relation entre une « vraie » exposition d'intérêt inconnue (e.g., radiologique) et un risque sanitaire d'intérêt ; (b) un modèle d'erreurs de mesure décrivant la relation entre la « vraie » exposition et l'exposition « observée » ; (c) selon la nature de l'erreur, un modèle d'exposition décrivant la distribution de la « vraie » exposition inconnue.

Notre cas d'étude concerne l'analyse de l'association entre une exposition chronique et à faibles doses au radon (et ses descendants à vie courte) et le risque de décès par pathologies radio-induites dans la cohorte française des mineurs d'uranium. Le radon, gaz radioactif d'origine naturelle, constitue la principale source

PROPOSITION DE SUJET DE STAGE

d'exposition naturelle aux rayonnements ionisants de l'Homme. Sachant qu'il fut reconnu comme cancérigène pulmonaire en 1988 par le CIRC et que sa concentration est plus élevée dans des endroits confinés comme les habitations ou les mines, il est important de connaître les risques sanitaires associés afin de permettre in fine des propositions d'amélioration des normes de radioprotection. Dans ce contexte, les cohortes de mineurs de fond qui, sont, dans le cadre de leur activité professionnelle, eux-mêmes exposés au radon constituent des populations d'étude privilégiées mais pour lesquelles les erreurs de mesure d'exposition ont souvent une structure complexe (e.g., hétéroscédastiques, Berkson/classique, partagées/non-partagées...). L'objectif est d'affiner, si possible, l'estimation actuelle du risque de décès par cancers (e.g., poumon, rein, cerveau, leucémie...) potentiellement associés à une exposition chronique et à faibles doses au radon à partir des données de la cohorte française des mineurs d'uranium en tenant compte de la structure complexe des erreurs de mesure sur l'exposition au radon dans cette cohorte.

Jusqu'à présent, un modèle hiérarchique bayésien supposant l'existence d'erreurs de mesure non-partagées de type Berkson ou classique (selon la période calendaire) a été proposé et inféré à l'aide d'algorithmes Monte Carlo par Chaînes de Markov (MCMC) afin d'estimer le risque de décès par cancer du poumon radio-induit dans la cohorte française des mineurs d'uranium. L'estimation de risque obtenue n'a changé que très marginalement par rapport à l'estimation obtenue sans prendre en compte les erreurs de mesure d'exposition au radon [3]. Néanmoins, des études par simulation, fondées sur les données d'exposition de la cohorte française des mineurs d'uranium et l'utilisation de modèles de survie (e.g. modèle de Cox et en excès de risque instantané), ont ensuite montré que des erreurs de mesure d'exposition partagées sur plusieurs années de suivi d'un même individu conduisent à des biais plus élevés dans les estimations de risque ainsi qu'à une déformation plus sévère de la relation exposition-risque que dans le cas d'erreurs de mesure non partagées ou partagées par plusieurs individus [5].

Mission

Tout en s'inspirant des derniers travaux de thèse de Sabine Hoffmann, les objectifs de ce stage sont :

- 1- de proposer un modèle hiérarchique bayésien supposant l'existence d'erreurs de mesure partagées entre plusieurs travailleurs et/ou sur plusieurs années de suivi successives d'un même travailleur dans la cohorte française des mineurs d'uranium. Il s'agira notamment de proposer un modèle de mesure pertinent permettant de décrire les erreurs de mesure partagées rencontrées après une reconstruction rétrospective d'expositions par dires d'experts. Cette technique d'évaluation de l'exposition au radon a notamment été utilisée dans les premières années de suivi de la cohorte française des mineurs d'uranium (1946-1956). Il s'agira également d'inclure de possibles facteurs susceptibles de modifier la relation exposition-risque : période d'exposition, temps depuis l'exposition...
- 2- d'implémenter un algorithme d'inférence bayésienne efficace afin d'estimer les paramètres inconnus du (des) modèle(s) hiérarchique(s) proposé(s) et de l'appliquer aux données de survie relatives à différents types de cancer (e.g., poumon, rein, cerveau, leucémie...) de la cohorte française des mineurs d'uranium
- 3- de mener une étude par simulations permettant a) de valider l'algorithme d'inférence bayésienne proposé ; b) d'étudier la robustesse du modèle proposé par rapport à une mauvaise spécification de modèles (modèle de mesure, modèle de maladie, modèle d'exposition).

Résultats attendus

Le stagiaire devra fournir :

- une **estimation corrigée** de l'excès de risque de décès associé à différents types de cancers d'intérêt ainsi qu'une comparaison avec une estimation de risque ignorant les incertitudes d'exposition ;
- une **analyse critique de la robustesse** de l'approche hiérarchique bayésienne proposée pour la prise en compte d'erreurs de mesure partagées dans les études de cohortes professionnelles en se fondant sur les résultats de l'étude par simulations.

Intérêts du stage

Ce travail de stage présente trois intérêts principaux :

- il s'insère parfaitement dans l'axe programme de recherche IRSN « Effets des faibles doses des rayonnements ionisants en situations d'exposition chronique » (D3P8). Par ailleurs, il s'inscrit dans la priorité de recherche en radioprotection suivante : « Forme de la relation dose-réponse pour les cancers » déclinée dans les agendas stratégiques de recherche des plateformes européennes MELODI, EURADOS et EURAMED;
- un intérêt en santé publique et en radioprotection. En effet, la production d'une estimation de risque de mortalité par pathologies radio-induites, corrigée des erreurs de mesure d'exposition radiologique et assortie de l'incertitude qui l'entache, devrait constituer une nouvelle source d'information intéressante pour la Commission Internationale de Protection Radiologique (CIPR), à l'origine du système de radioprotection actuelle;
- un intérêt méthodologique en contribuant à promouvoir l'utilisation de l'approche hiérarchique bayésienne qui est encore très rarement utilisée en épidémiologie des rayonnements ionisants malgré sa grande souplesse et sa pertinence pour la prise en compte d'erreurs de mesure d'exposition complexes et pour l'estimation du risque de pathologies radio-induites en situation peu informative.

Profil recherché

Ce stage de 6 mois (à partir de février, mars ou avril 2019) est destiné à un(e) étudiant(e) de M2 ou équivalent avec spécialité statistique appliquée aux sciences du vivant ayant des bases solides en modélisation probabiliste et en statistique bayésienne ainsi qu'un intérêt pour la programmation informatique et les applications en épidémiologie/santé publique. La maîtrise du langage R est indispensable. Une bonne maîtrise de l'anglais est exigée et une connaissance du langage de programmation Python serait un plus.

Informations pratiques

- Le stage sera entièrement financé par l'IRSN. Néanmoins, le stagiaire partagera les 6 mois de son stage entre l'IRSN (pendant 3 mois) sous la supervision directe de Sophie Ancelet et l'Université Paris-Descartes (pendant 3 mois) sous la supervision directe de Chantal Guihenneuc.
- Le stage sera rémunéré à hauteur de 1200 euros par mois et sera accompagné d'un accès au self-service de l'IRSN à prix réduit et d'une prise en charge partielle des titres de transport.
- Le stagiaire aura à sa disposition les ressources nécessaires pour mener à bien sa mission au sein de l'IRSN et de l'université Paris-Descartes: bureau, ordinateur de bureau personnel (système d'exploitation : Windows), logiciels statistiques.

Bibliographie :

- [1] Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. (2006) Measurement Error in Nonlinear Models : A Modern Perspective. Chapman & Hall, CRC Press : Boca Raton, FL.
- [2] Hoffmann S. (2017) Approche hiérarchique bayésienne pour la prise en compte d'erreurs de mesure d'exposition chronique et à faibles doses aux rayonnements ionisants dans l'estimation du risque de cancers radio-induits. Application à une cohorte de mineurs d'uranium. Manuscrit de thèse disponible online. Université Paris-Saclay.
- [3] Hoffmann S., Rage E., Laurier D., Laroche P., Guihenneuc C., Ancelet S. (2017) Accounting for Berkson and classical measurement error in radon exposure assessment using a Bayesian structural approach in the analysis of lung cancer mortality in the French cohort of uranium miners. *Radiation Research*.187(2): 196-209
- [4] Belloni M. (2017) Modélisation hiérarchique bayésienne pour la prise en compte d'erreurs de mesure d'exposition radiologique dans l'estimation du risque de décès par cancer du poumon avec application à la cohorte française des mineurs d'uranium – Cas des expositions aux rayonnements gamma et des expositions manquantes. Rapport de stage M2. IRSN.
- [5] Hoffmann S., Laurier D., Rage E., Guihenneuc C., Ancelet S. (2018) Shared and unshared measurement error in occupational cohort studies and their effects on statistical inference in proportional hazards models. *Plos One*. 13(2):e0190792