

BAYESIAN MODEL AVERAGING À L'AIDE D'UN ÉCHANTILLONNAGE PRÉFÉRENTIEL ADAPTATIF ET MULTIPLE POUR L'ÉSTIMATION DU RISQUE DE LEUCÉMIE INFANTILE RADIO-INDUITE.

Sophie Ancelet ¹ & Merlin Keller ²

¹ *Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PRP-HOM, SRBE, Laboratoire d'épidémiologie (LEPID), BP17, 92262 Fontenay-aux-Roses, France. Email: sophie.ancelet@irn.fr*

² *EDF R&D, Département Management des Risques Industriels, 6 quai Wattier, Chatou, France. Email: merlin.keller@edf.fr*

Résumé.

En épidémiologie des rayonnements ionisants (RI), plusieurs modèles dose-risque de cancer radio-induit peuvent s'ajuster de manière équivalente à un même jeu de données. Ces modèles diffèrent de par le choix d'une relation dose-risque, en excès de risque relatif ou absolu, et des facteurs de risque susceptibles de modifier cette relation. Cette incertitude de modèle est ignorée en pratique: un seul modèle est généralement sélectionné pour estimer les risques de cancer radio-induit. Une telle approche peut mener à des estimations d'excès de risque biaisées en forçant un unique modèle à s'ajuster à des données qui n'en vérifient pas nécessairement toutes les caractéristiques et à une sous-estimation de l'incertitude de ces estimations. Walsh et Kaiser (2011) ont récemment proposé de combiner plusieurs modèles de risque de cancer radio-induit en utilisant une approche fréquentiste, appelée inférence multi-modèles (MMI), basée sur le calcul de poids AIC. Nous proposons d'utiliser une approche alternative bayésienne, connue pour ses bonnes propriétés en termes d'estimation et de prédiction: le Bayesian Model Averaging (BMA). Pour inférer les modèles en compétition et réaliser le BMA, nous proposons une approche par échantillonnage préférentiel adaptatif et multiple basé sur une fonction d'importance originale: une copule Gaussienne avec lois marginales de Student décentrées. Nous analysons les forces et faiblesses de l'algorithme AMIS implémenté dans un cas d'étude réel portant sur l'estimation de l'excès de risque de leucémie infantile chez les survivants des bombardements d'Hiroshima-Nagasaki. Enfin, nous comparons les approches BMA et MMI à partir de ce cas d'étude et de simulations.

Mots-clés. Bayesian Model Averaging, échantillonnage préférentiel, épidémiologie, leucémie infantile, rayonnements ionisants, relation dose-risque, incertitude.

Abstract.

In radiation epidemiology, several radiation-related cancer risk models may fit a same dataset equally well. These models differ according to the choice of an excess relative or

excess absolute dose-risk structure and of potentially risk-modifying factors. Usual practice ignores such a model uncertainty: generally, a single model is selected to estimate the radiation-related cancer risks. Such an approach may lead to biased estimations of excess risks by forcing a unique model to fit to data that do not necessarily support it in all its characteristics and to an underestimation of the uncertainty of these estimations. Walsh and Kaiser (2011) have recently proposed to combine several radiation-related cancer risk models using a frequentist approach, called multi-model inference (MMI), based on AIC weights. We propose to use a Bayesian alternative approach, known for its good properties in terms of both estimation and prediction: Bayesian model averaging (BMA). To infer the competing models and carry out BMA, we propose an adaptive multiple importance sampling approach based on an original importance function: Gaussian copulas with non-central Student marginal distributions. We analyse the strengths and weaknesses of the implemented AMIS algorithm from a real-life case study dealing with the estimation of childhood leukemia excess risks among the survivors of the atomic bombings of Hiroshima and Nagasaki. Finally, we compare the BMA and MMI approaches from this case-study and simulations.

Keywords. Bayesian Model Averaging, importance sampling, epidemiology, childhood leukemia, ionising radiation, dose-risk relationship, uncertainty.

1 Introduction

En épidémiologie des rayonnements ionisants (RI), plusieurs modèles probabilistes décrivant la relation entre une dose reçue de RI et un risque de mortalité par cancer radio-induit peuvent s'ajuster de manière équivalente à un même jeu de données. Malgré des performances très proches en terme d'ajustement, ces modèles diffèrent de par le choix de la forme de la relation dose-risque (e.g., linéaire, linéaire quadratique exponentielle, linéaire par morceaux, ...), la forme du risque de base (i.e., à dose nulle), le choix d'une relation dose-risque en excès de risque relatif (ERR) ou absolu (EAR) et des facteurs de risque susceptibles de modifier cette relation.

Cette incertitude sur le modèle dose-risque est ignorée en pratique: un seul modèle est généralement sélectionné pour estimer les risques de cancers radio-induits. Une telle approche peut mener à des estimations biaisées des excès de risque d'intérêt, en forçant un unique modèle à s'ajuster à des données qui ne vérifient pas nécessairement toutes les caractéristiques du modèle, et à une sous-estimation de l'incertitude de ces estimations. À terme, cela peut donc mener à des conclusions erronées quant à l'ampleur et la significativité statistique de l'impact de l'exposition aux RI sur le risque de cancer d'intérêt.

L'impact de la combinaison de plusieurs modèles dose-risque sur l'estimation du risque de cancers radio-induits est examiné depuis peu en épidémiologie des RI. Divers comités

scientifiques ont proposé d'utiliser des mélanges de modèles en ERR et EAR dans les évaluations quantitatives de risque mais en attribuant des poids subjectifs aux modèles i.e., définis à partir de dires d'experts [1]. Suivant les travaux de Burnham et Anderson (2002)[2], Walsh et Kaiser [3] proposent plutôt d'utiliser une procédure quantitative fréquentiste de combinaison de modèles probabilistes, appelée inférence multi-modèles (MMI). Celle-ci est basée sur le calcul de poids de modèles reposant sur le critère d'information d'Akaike (AIC). Bien que l'utilisation de tels poids AIC peut en effet permettre de réduire les biais potentiellement induits par la sélection d'un modèle unique, aucune théorie existe quant à l'optimalité de l'utilisation de tels poids dans une procédure de combinaison de modèles.

Dans ce contexte et comme suggéré par ailleurs [3][4][5], nous souhaitons promouvoir l'utilisation alternative du Bayesian-Model Averaging (BMA), approche bayésienne bien connue pour ses bonnes propriétés théoriques en termes d'estimation et de prédiction dans le contexte de la combinaison de modèles. Celle-ci a également l'avantage d'intégrer une prise en compte de l'incertitude sur les paramètres inconnus des modèles considérés. A notre connaissance, cette approche n'a jamais été utilisée en épidémiologie des RI. Il s'agit de comparer les performances respectives des approches BMA et MMI à partir de cas d'étude réels et de simulations et de proposer des outils opérationnels aux épidémiologistes.

2 Cas d'étude

Ce travail est motivé par la question d'intérêt suivante: "Existe-t-il un impact statistiquement significatif de l'exposition aux RI sur l'excès de risque de mortalité par leucémie infantile?". De par leur grande richesse, les données de mortalité de la cohorte Life Span Study (LSS) des survivants des bombardements d'Hiroshima Nagasaki, fournies par la "Radiation Effects Research Foundation" (RERF, Japon), sont les données de référence principalement utilisées jusqu'à présent pour répondre à cette question. A noter qu'elles servent de base au système de radioprotection international.

Dans ce travail, nous utiliserons les données de mortalité par leucémie (tout type) disponibles dans la dernière mise à jour de la LSS [6]. Elles correspondent au suivi de 86611 survivants aux bombardements d'Hiroshima et Nagasaki sur la période 1950-2003, soit 3294×10^6 personnes-années. A la fin de l'année 2003, 318 survivants sont décédés d'une leucémie. Les données disponibles ont pour particularité d'être stratifiées par groupe de personnes-années homogènes en fonction d'un certain nombre de critères.

Treize modèles de régression de Poisson distincts mais non emboîtés (incluant notamment des structures en EAR et ERR) ont été considérés dans ce travail. Il s'agit de modèles dose-risque proposés dans la littérature pour estimer le risque de mortalité par leucémie radio-induite à partir des données de la LSS.

3 Description de la méthode

Soient M_k ($k=1, \dots, 13$) les 13 modèles candidats de risque de leucémie radio-induite. Chaque modèle est défini par un jeu de paramètres θ_k . Soit Δ une quantité d'intérêt à estimer ou prédire, i.e., ici, l'excès de risque de leucémie radio-induite pour un âge, sexe et dose de RI donnés. L'approche BMA [7] peut se résumer par les deux équations suivantes:

$$[\Delta|y] = \sum_{k=1}^K [\Delta(\theta_k)|y, M_k] \omega_k \quad (1)$$

où $[\Delta|y]$ désigne la loi a posteriori de Δ intégrée sur les différents modèles candidats. Elle s'écrit comme le mélange des lois a posteriori de Δ sous chaque modèle M_k ($k=1, \dots, 13$) pondéré par ω_k , le poids du modèle M_k . Ce dernier est défini comme la probabilité a posteriori de M_k sachant les données observées y :

$$\omega_k = [M_k|y] = \frac{[y|M_k][M_k]}{\sum_{l=1}^K [y|M_l][M_l]} \quad (2)$$

où $[M_k]$ désigne la probabilité a priori du modèle M_k et $[y|M_k]$ la vraisemblance marginale, encore appelée évidence, du modèle M_k . Celle-ci intègre l'incertitude sur les paramètres du modèle M_k . Par la suite, une loi uniforme est assignée à l'espace des modèles candidats.

Dans notre cas d'étude, deux difficultés principales se posent pour la mise en oeuvre du BMA: (1) le calcul de la vraisemblance marginale (notée ML_k par la suite) de chaque modèle M_k , qui s'écrit comme une intégrale en grande dimension (entre 10 et 23 paramètres) (2) les temps de calculs importants nécessaires pour obtenir, pour chacun des 13 modèles candidats, un échantillon représentatif de la loi a posteriori de l'excès de risque de leucémie radio-induite à partir des données de la LSS via un algorithme Monte-Carlo par Chaînes de Markov (MCMC).

Pour parer à ces deux difficultés, nous proposons d'utiliser une approche par échantillonnage préférentiel. En effet, celle-ci fait partie des techniques connues permettant d'obtenir un estimateur sans biais de ML_k pour tout k [7] à partir d'un échantillon de tirages pondérés, appelés particules, selon une fonction d'importance choisie. Par ailleurs, moyennant une étape de ré-échantillonnage, elle peut permettre d'accélérer (par rapport au MCMC) le temps nécessaire à l'obtention d'un échantillon approché de la loi a posteriori de l'excès de risque de leucémie radio-induite selon chaque modèle.

Afin de capter des formes de dépendances a posteriori spécifiques entre paramètres, nous proposons d'utiliser une fonction d'importance originale à savoir une copule Gaussienne basée sur des lois marginales de Student décentrées (appelée NCT-GC par la suite). Une fonction d'importance standard de type Student multivariée (appelée MVT par la suite) a également été considérée.

Par ailleurs, afin d'optimiser les performances de l'algorithme d'échantillonnage préférentiel proposé, nous avons implémenté (en langage Python) un algorithme d'échantillonnage

préférentiel adaptatif et multiple (AMIS) [8][9][10]. Pour chaque modèle M_k , l'idée est tout d'abord de construire des fonctions d'importance qui s'adaptent séquentiellement à la loi cible i.e., la loi a posteriori jointe des paramètres θ_k . À chaque étape de l'algorithme, il s'agit d'ajuster la loi d'importance choisie à l'ensemble des échantillons pondérés passés concaténés. Puis, afin de recycler les simulations obtenues au cours du processus adaptatif en vue de réduire les temps de calculs, de rendre la répartition des poids des particules générées la plus uniforme possible, et ainsi de réduire au mieux la variance de l'estimateur de Monte-Carlo résultant pour la vraisemblance marginale ML_k ($k=1, \dots, 13$), nous avons considéré la stratégie de post-traitement des poids des particules générées proposée par Owen et Zhou [8][9], qui considèrent l'ensemble des particules générées comme étant issues d'un mélange déterministe de toutes les lois d'importance considérées.

4 Résultats

Nous avons mis en oeuvre notre algorithme AMIS, basé sur une fonction d'importance NCT-GC, à partir des données de mortalité de la LSS et ce, pour chacun des 13 modèles candidats de risque de leucémie radio-induite. Ceci nous a permis de calculer le poids a posteriori de chaque modèle et d'obtenir un échantillon a posteriori approché de l'excès de risque de leucémie radio-induite pour chaque modèle et pour l'ensemble des modèles combinés (BMA).

Nous avons tout d'abord comparé les performances de notre algorithme AMIS à 3 stratégies plus simples de post-traitement des poids des particules générées. Dans 8 modèles sur 13, l'algorithme AMIS permet d'obtenir un gain en terme de taille d'échantillon équivalent (ESS), témoignant d'un recyclage plus efficace des particules. Dès lors qu'une approche par échantillonnage préférentiel adaptatif ne semble pas fonctionner, l'algorithme AMIS n'apporte aucune amélioration notable en terme d'ESS (cas de 5 modèles sur 13). Nous mettons également en évidence les avantages à l'utilisation d'une fonction d'importance NCT-GC par rapport à une fonction d'importance MVT: gain en terme d'ESS, fonction d'importance permettant de capter des corrélations a posteriori "en banane" non captées par la MVT. En revanche, le choix d'une fonction d'importance complexe telle que la NCT-GC ne semble pas recommandé dans le cas de modèles en trop grande dimension.

Dans notre cas d'étude, les approches BMA et MMI fournissent des poids de modèle contrastés. Néanmoins, les deux approches attribuent le poids le plus élevé au même modèle basé sur une relation dose-risque purement quadratique, modifiée par une fonction puissance en l'âge atteint. Les modèles de risque de leucémie radio-induite en ERR ont des poids plus élevés que les modèles en EAR. Bien que les intervalles de crédibilité à 95% obtenus diffèrent, un excès de risque de mortalité par leucémie infantile statistiquement significatif est observé pour chaque modèle ainsi que les approches BMA et MMI, pour une dose reçue supérieure à 0.10 Sv à un âge à l'exposition de 2 ans et un âge atteint de

7 ans. Les conclusions sont un peu plus contrastées pour des doses inférieures à 0.10 Sv.

Une étude par simulations est actuellement en cours en vue de renforcer la comparaison des approches BMA et MMI. Nous considérons un sous-ensemble des 13 modèles précédemment considérés et cherchons à comparer les performances en termes d'estimation mais également de prédictions des approches BMA et MMI dans la situation où le "vrai" modèle appartient ou pas à la liste des modèles candidats.

Bibliographie

- [1] Committee to Assess Health Risks from Exposure to Low Levels of Ionizing Radiation (2006), *Health Risks from Exposure to Low Levels of Ionizing Radiation*, BEIR VII Phase 2, The National Academies Press, Washington DC.
- [2] Burnham, K.P et Anderson, D.R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd Edition, Springer, New-York.
- [3] Walsh, L. et Kaiser, J.C. (2011), Multi-model inference of adult and childhood leukaemia excess relative risks based on the Japanese A-bomb survivors mortality data (1950-2000), *Radiation and Environmental Biophysics*, 50(1):21-35.
- [4] Little, M. et Wakeford, R. (2013), How is the risk of radiation-induced cancer influenced by background risk factors? Invited commentary on "A method for determining weights for excess relative risk and excess absolute risk when applied in the calculation of lifetime risk of cancer from radiation exposure" by Walsh and Schneider (2011). Letter to the Editor. *Radiation and Environmental Biophysics*, 52(1):147-50.
- [5] Richardson, D.B. et Cole, S.R. (2012) Model averaging in the analysis of leukemia mortality among Japanese A-bomb survivors. Letter to the editor. *Radiation and Environmental Biophysics*, 51:93-95.
- [6] Ozasa, K. et al. (2012) Studies of the Mortality of Atomic Bomb Survivors, Report 14, 1950-2003: An Overview of Cancer and Noncancer Diseases. *Radiation Research*, 177(3):229-243.
- [7] Hoeting, J.A, Madigan, D., Raftery, A.E. et Volinsky, C.T. (1999) Bayesian model averaging: A tutorial. *Statistical Science*, 14(4):382-417.
- [8] Owen, A. et Zhou, Y. (2000), Safe and effective importance sampling, *Journal of the American Statistical Association*, 95, 135-143.
- [9] Cornuet, J.M., Marin, J.M., Mira, A. et Robert, C.P. (2012), Adaptive Multiple Importance Sampling, *Scandinavian Journal of Statistics*, 39(4), 798-812.
- [10] Liu, J., Liang F. et Wong, W. (2001). A theory of dynamic weighting in Monte Carlo computation. *Journal of the American Statistical Association*, 96:561-573.