Promouvoir le Bayesian Model Averaging pour améliorer l’évaluation quantitative du risque de cancers radio-induits. Application à la leucémie infantile

Sophie Ancelet\textsuperscript{1}, Merlin Keller\textsuperscript{2}
En collaboration avec : Olivier Laurent\textsuperscript{1}, Denis Hénon\textsuperscript{3}, Géraldine Ielsch\textsuperscript{1}, Jacqueline Clavel\textsuperscript{3}, Dominique Laurier\textsuperscript{1}

\textsuperscript{1}Institut de Radioprotection et de Sureté Nucléaire, Fontenay-aux-Roses
\textsuperscript{2}EDF R&D, Chatou
\textsuperscript{3}INSERM, CESP 1018, Villejuif

AppliBUGS, 20 Juin 2013
Outline

1. Context
2. Motivating Case Study
3. Method
4. Results
5. Discussion
1 Context

2 Motivating Case Study

3 Method

4 Results

5 Discussion
Epidemiology of ionizing radiation

- All people are exposed to ionizing radiation (IR)

Epidemiology of IR = Study of the stochastical effects of IR
  - Non-specific late effects with dose-dependent occurrence
    - Cancer diseases (solid, leukemia), cardiovascular diseases, cataract,...
  - Observational science (cohort studies, case-control studies) : confusing factors, bias, missing data, extrapolations from one population to another...
An important topic for radiation protection...

- At acute, medium-to-high level, external exposure to IR, excess risk of leukemia, breast, lung and thyroid cancer are clearly demonstrated.
- The excess risk of cancer diseases increases with the dose.
- Latency minimal period from a few years to decades.

May chronic exposure to low doses rate of IR result in adverse health effects?

Only a few results at low doses rates of IR:
- Lung cancer following radon inhalation (Darby et al., 2005; IARC, 2001)
- Excess risk of leukemia among children exposed *in-utero* (>10 mSv)
- **BUT mixed findings** for other potential health effects at doses <100 mSv.
Limits of current epidemiological studies

- Latency minimal period from IR exposure to cancer occurrence from a few years to decades
- Suboptimal designs (bias, non-observed confusing factors, ...)

→ Lack of statistical power to detect some potential small health effects of IR at low dose rates

Additional well-designed epidemiological studies are on progress BUT:

- **Years to decades of observations** required to reach an adequate statistical power to detect such potential health impacts
- **Some quicker replies** are legitimately called on the expected magnitude of a potential risk!

→ **Alternative approach**: Quantitative Risk Assessment (QRA) (NRC, 2009)
Step 1: Building one (several) probabilistic model(s) to describe the risk-exposure/dose-effect relationship of interest.

Step 2: Fitting the proposed model(s) to data observed in the so-called *evidentiary population*.

Step 3: Predicting some health impact indicators in a so-called *target population* (e.g., years of life lost, lifetime excess cancer deaths, attributable risk proportion,...) from some dose estimations and the information provided by the *evidentiary population*. 
In practice, the Life Span Study (LSS) of Hiroshima and Nagasaki A-bomb survivors is used as the evidentiary population.

- Main basis for setting international radiation protection standards (ICRP 2007)
- → The WHO report "Health Risk Assessment from the nuclear accident after the 2011 Great East Japan Earthquake and Tsunami based on preliminary dose estimation" (28/02/2013, available online)
Quantitative assessment of radiation-related risk: Necessary assumptions

**Transposition**: The fitted dose-response relationship is still valid for a target population different from the evidentiary population.

**Extrapolation**: The dose-response relationship is still valid in a range of exposures different from the one on which it has been estimated.

**Analogy**: All the types of IR exposure have a similar health impact.

Dose-response model fitted to an evidentiary population.

\[ RR = 1 + 0.0036 \times \text{Dose} \]
Quantitative assessment of radiation-related risk: Some methodological limits

Many statistical models describe the evolution over time of the Excess Relative or Absolute Risk (ERR or EAR) of cancer due to IR from LSS but usual practice of QRA ignores:

- **Model selection uncertainty**
- **Parameters uncertainty**
Walsh & Kaiser (2012) examine the impact of combining models for radiation-related leukemia risks assessments.


Considered as an objective basis for multimodel inference in many fields like epidemiology, biology and ecology (Zhang and Townsend (2009); Burnham et al. (2011); Walsh and Schneider (2013)).
Let $M_k (k = 1, \ldots, K)$ be $K$ competing risk models considered, each one defined by a set of parameters $\theta_k$. and $\Delta$ be a quantity of interest to estimate/predict and $y$ the observed data.

A model-averaged estimator of $\Delta$ is given by:

$$\hat{\Delta} = \sum_{k=1}^{K} \hat{\Delta}_k \omega_k$$

where $\omega_k (k=1,\ldots,K)$ are the Akaike weights defined by:

$$\omega_k = \frac{\exp(-0.5(\Delta AIC_k))}{\sum_{j=1}^{K} \exp(-0.5(\Delta AIC_j))}$$

where

$$AIC_k = -2\log[y|\theta_k] + 2p_k \quad n >> p_k$$

$$\Delta AIC_k = AIC_k - AIC_{min}$$
Burnham & Anderson (2004) propose the following Bayesian interpretation of the AIC weights.

Let $\pi_k$ be the prior probability placed on model $M_k$. Then the posterior probability for model $M_k$ given data $y$ is:

$$[M_k | y] \simeq \frac{\exp(-0.5(\Delta BIC_k))\pi_k}{\sum_{j=1}^{K} \exp(-0.5(\Delta BIC_j))\pi_j}$$

If the model prior probability $\pi_k$ are proportional to

$$\exp(0.5(\Delta BIC_k))\exp(-0.5(\Delta AIC_k))$$

then

$$[M_k | y] \simeq \frac{\exp(-0.5(\Delta AIC_k))}{\sum_{j=1}^{K} \exp(-0.5(\Delta AIC_j))} = \omega_k$$

"...traditional Bayesian thinking about the prior distribution on models has been that $\pi_k$, $k=1, \ldots K$ would also not depend on $n$ or $p_k$. This approach is neither necessary nor reasonable." (Burnham & Anderson (2004))
## Aim of the work

Investigate the use of Bayesian Model Averaging (BMA) to account for model and parameters uncertainties in cancer risk assessments due to IR
1 Context

2 Motivating Case Study

3 Method

4 Results

5 Discussion
Natural Background Radiation (NBR) constitutes the major source of exposure to chronic IR for most of the world population (UNSCEAR, 2008).

Three components contribute to 90% of the effective dose delivered:

- Radon gas (222Rn and 220Rn) and its decay products
- Terrestrial gamma rays (TGR)
- High energy cosmic ray particle

NBR & childhood leukemia: Why is it an important topic?

- During childhood, equivalent dose received by the red bone marrow (RBM) ranging from a few to several tens of mSv!!!
- Childhood leukemia
  - Relevant health indicator when studying the effects of NBR
  - Most strongly associated with exposure to external whole-body irradiation
  - Children more radiosensitive than adults (NRC, 2006)
- Childhood leukemia = the most frequent cancer in children but whose etiology remains widely unknown (Eden, 2010)
Positive association between radon exposure and leukemia incidence
  ▶ Some, though not all, ecological studies (Evrard et al., 2006; Laurier et al., 2001)
  ▶ A case-control study in Denmark (Raaschou-Nielsen et al., 2008) BUT limited statistical power (Little et al., 2010)

Positive association between exposure to TGR and directly ionizing cosmic radiation and childhood leukemia incidence
  ▶ A sufficient size case control study in the UK (Kendall et al., 2013)

← Quantitative Risk Assessment applied in Great Britain: 15 to 20% of leukemia cases might potentially be attributable to NBR over childhood (from 0 to 14 years old) (Little et al., 2009; Wakeford et al., 2009)
Data on the target population

Childhood acute leukemia incidence rates in metropolitan France by sex and mean attained age (0-14 years old), period 1990-2004 (French National Registry of Childhood Blood Malignancies (INSERM – RNHE))

6784 childhood leukemia cases recorded in France during the study period (around 250 cases per year among males and 202 among females)
Average red bone marrow doses (in mSv) received by fetuses, infants and children from radon, terrestrial gamma rays and cosmic rays in France

<table>
<thead>
<tr>
<th></th>
<th>Radon</th>
<th>Terrestrial gamma rays</th>
<th>Cosmic rays</th>
<th>All 3 exposures together</th>
</tr>
</thead>
<tbody>
<tr>
<td>In utero (9 month)</td>
<td>0.03&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.33&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.19&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.55</td>
</tr>
<tr>
<td>Infant (first year of life&lt;sup&gt;e&lt;/sup&gt;)</td>
<td>0.29&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.61&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.35&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.24</td>
</tr>
<tr>
<td>Child (yearly&lt;sup&gt;f&lt;/sup&gt;)</td>
<td>0.34&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.55&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.31&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1.21</td>
</tr>
<tr>
<td>Cumulated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(in utero - 12.5 years)</td>
<td>4,40</td>
<td>7,54</td>
<td>4,26</td>
<td>16,31</td>
</tr>
<tr>
<td>% of cumulated dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(in both cases)</td>
<td>27</td>
<td>46</td>
<td>26</td>
<td>100</td>
</tr>
</tbody>
</table>
Mortality dataset from the latest Life Span Study (LSS) cohort (provided by the Radiation Effects Research Foundation, Hiroshima, Japan)

- 86,611 survivors of the atomic bombings of Hiroshima and Nagasaki over period 1950-2000
- By the end of 2000, 284 had died from leukemia
- Stratified data by city, sex, age at exposure, weighted colon dose category (in Sv), attained age, calendar time period ... \( \rightarrow 31,422 \) strata

For each stratum:

- Number of deaths due to leukemia
- Number of person-years at risk
- \( PY \) —weighted average age at exposure
- \( PY \) —weighted average attained age
- Estimated stratum-average RBM doses (in Sv) corresponding to the most recent dosimetric system available for the cohort (\( DS02 \), established in 2002)
Additive and Multiplicative risk models

Let $Y_i$ be the number of leukemia deaths and $PYR_i$ the associated number of persons-years at risk in stratum $i$ of the LSS data.

\[
Y_i \sim \text{Poisson}(PYR_i \times \lambda_{\text{tot},i}^{LSS})
\]

\[
\lambda_{\text{tot},i}^{LSS} = \begin{cases} 
\lambda_{0,\xi}^{LSS}(s_i, c_i, a_i, e_i) + \text{EAR}(d_i, s_i, c_i, a_i, e_i) \\
\lambda_{0,\xi}^{LSS}(s_i, c_i, a_i, e_i) \times (1 + \text{ERR}(d_i, s_i, c_i, a_i, e_i)) 
\end{cases}
\]

\[
\text{ERR}/\text{EAR}(d_i, s_i, c_i, a_i, e_i) = (\alpha d_i + \beta d_i^2) \exp(\gamma d_i) \omega(\mu) (s_i, c_i, a_i, e_i)
\]

- $\lambda_{0,\xi}^{LSS}(s_i, c_i, a_i, e_i)$ is the LSS baseline risk in the absence of exposure
- $\text{EAR}$ is the Excess Absolute Risk / $\text{ERR}$ is the Excess Relative Risk
- Constraints must be assigned to the vector $\theta$ of unknown parameters
10 Poisson-disease models sharing common features have been found in the literature.

<table>
<thead>
<tr>
<th>ERR models</th>
<th>Np</th>
<th>EAR models</th>
<th>Np</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERR.Little (2008)</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAR.Schneider (2009)</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAR.Schneiderexp (2009)</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAR.Preston (2004)</td>
<td>23</td>
</tr>
</tbody>
</table>

Np = Number of parameters
How to assess the proportion of cases attributable to NBR in France?

In case of ERR transfer from the evidentiary population to the target population:

\[
h^F_{\text{tot}}[s, a, e, DNR(e)] = h^F_0(s, a) + \sum_{e=0.5}^{a-2} h^F_0(s, a) \cdot \text{ERR}(s, a, e, DNR(e))
\]

\[
AP_{\text{NBR}}[s, a, e, DNR(e)] = \frac{\sum_{e=0.5}^{a-2} \text{ERR}(s, a, e, DNR(e))}{1 + \sum_{e=0.5}^{a-2} \text{ERR}(s, a, e, DNR(e))}
\]

In case of EAR transfer from the evidentiary population to the target population:

\[
h^F_{\text{tot}}[s, a, e, DNR(e)] = h^F_0(s, a) + \sum_{e=0.5}^{a-2} \text{EAR}(s, a, e, DNR(e))
\]

\[
AP_{\text{NBR}}[s, a, e, DNR(e)] = \frac{\sum_{e=0.5}^{a-2} \text{EAR}(s, a, e, DNR(e))}{h^F_{\text{tot}}[s, a, e, DNR(e)]}
\]

Remark: Risk-free period (lag) of 2 years following exposure.
<table>
<thead>
<tr>
<th>1</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Motivating Case Study</td>
</tr>
<tr>
<td>3</td>
<td>Method</td>
</tr>
<tr>
<td>4</td>
<td>Results</td>
</tr>
<tr>
<td>5</td>
<td>Discussion</td>
</tr>
</tbody>
</table>
Let $M_k$ ($k = 1, \ldots, K$) be the $K$ competing risk models considered, each one defined by a set of parameters $\theta_k$. Let $\Delta$ be a quantity of interest (e.g., the percentage of leukemia cases attributable to NBR) to estimate/predict. **One main equation**:

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

where $\omega_k$ is the posterior probability for model $M_k$ given data $y$:

$$\omega_k = [M_k|y] = \frac{[y|M_k][M_k]}{\sum_{i=1}^{K} [y|M_i][M_i]}$$

**Remark** : Relies on the assumption that $\Delta(\theta_k)$ is *transferrable* across models.
Importance Sampling : Why ?

First tested approach :

- $[\Delta(\theta_k)|y, M_k]$ sampled using MCMC algorithms implemented in OpenBUGS

- $ML_k := [y|M_k]$ estimated using posterior-guided Importance Sampling (IS)

$$ML_k = \frac{1}{N} \sum_{i=1}^{N} \left[ y|\theta^{(i)}_k, M_k \right] \frac{\theta^{(i)}_k|M_k}{g(\theta^{(i)}_k)} \quad \theta^{(i)}_k \sim i.i.d. g(\theta^{(i)}_k)$$

**IS function $g$** : We propose a **product of univariate scaled noncentral Students distributions** fitted to the posterior samples.

- Due to high within-chain autocorrelations and large dataset ($\geq 30,000$ observations), approach is very computationally expensive! ($\approx 2$ days per model)

$\rightarrow$ Importance sampling enables to sample from posterior distribution **and** compute marginal likelihood all at once!
Importance Sampling : How ?

For each model $M_k$ :

1. Choose importance distribution $g(\theta_k)$ as a ‘good’ approximation of $[\theta_k|y, M_k]$

   Following Liu (2001), we use a multivariate Student distribution centered at MLE $\hat{\theta}_k$, with $df = 30$ and covariance matrix equal to inverse observed Fisher Information $\mathcal{I}(\hat{\theta}_k)^{-1}$

2. Draw $N$ i.i.d. realizations $\theta_k^{(i)}$ from $g$. Let : $\tilde{w}_k^{(i)} = \frac{[y|\theta_k^{(i)}, M_k][\theta_k^{(i)}|M_k]}{g(\theta_k^{(i)})}$ be the non-normalized importance weights

3. Estimate marginal likelihood (without bias) as : $\overline{ML}_k = \frac{1}{N} \sum_{i=1}^N \tilde{w}_k^{(i)}$, and any posterior expectation $\mathbb{E}[\Delta(\theta_k)|y, M_k]$ by

\[
\hat{\mathbb{E}}[\Delta(\theta_k)|y, M_k] = \frac{\sum_{i=1}^N \tilde{w}_k^{(i)} \Delta(\theta_k^{(i)})}{\sum_{i=1}^N \tilde{w}_k^{(i)}}
\]
Importance Sampling Resampling

- Posterior distribution $[\theta_k^{(i)}|y, M_k]$ is approximated by:
  $$\frac{\sum_{i=1}^N \tilde{w}_k^{(i)} \delta_{\theta_k^{(i)}}}{\sum_{i=1}^N \tilde{w}_k^{(i)}}$$
  where $\delta_{\theta_k^{(i)}}$ is the Dirac mass in $\theta_k^{(i)}$

  $\mapsto$ Approximate posterior sample can be obtained by resampling the $\theta_k^{(i)}$ with probability $w_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum_{j=1}^N \tilde{w}_k^{(j)}}$ (normalized importance weight)

- Quality of the importance sampling algorithm can be monitored by:
  - Equivalent Sample Size (Liu, 2001; Del Moral, 2004):
    $$ESS = \left( \sum_{i=1}^N w_k^{(i)} \right)^{-1}$$
  - Approximate weight variation coefficient (Oh and Berger, 1989):
    $$cv = \frac{\text{var}(\tilde{w}_k)}{N\text{mean}(\tilde{w}_k)^2}$$
Adaptive Importance sampling

- In practice, choosing a ‘good’ importance distribution $g(\theta_k)$ is a key issue
- **Idea**: Run Importance sampling iteratively to continually update $g(\theta_k)$
- Following (Oh and Berger, 1989), we perform the following steps:
  1. Set $t = 0$ and define:
     $$g^{(0)}(\theta_k) = \text{MVT} \left( \theta_k | \hat{\theta}_k, I(\hat{\theta}_k)^{-1}, df = df^{(0)} \right)$$
     Then:
  2. Draw $N = 500$ i.i.d. realizations $\theta_{k,t}^{(i)}$ from $g^{(t)}$ and associated weights $\tilde{w}_{k,t}^{(i)}$. Compute the corresponding ESS and cv values: $ESS_t, cv_t$.
  3. If $t \geq 1$ and $ESS_t < ESS_{t-1}$, discard weighted sample $(\theta_{k,t}^{(i)}, \tilde{w}_{k,t}^{(i)})$ and repeat previous step.
     While $ESS_t < 10000$ and $cv_t > 2.6 \times 10^{-5}$:
  4. Increment $t = t + 1$. Define
     $$g^{(t)}(\theta_k) = \text{MVT} \left( \theta_k | \hat{\mathbb{E}}^{(t)}[\theta_k|y, M_k], \hat{\mathbb{V}}^{(t)}[\theta_k|y, M_k], df = df^{(t)} \right),$$
     where $(\hat{\mathbb{E}}^{(t)}[\theta_k|y, M_k], \hat{\mathbb{V}}^{(t)}[\theta_k|y, M_k], df^{(t)})$ is fitted to the pooled posterior weighted sample. Then, go back to step 2.
1. Context

2. Motivating Case Study

3. Method

4. Results

5. Discussion
## Simple vs. Adaptive IS: ERR models

<table>
<thead>
<tr>
<th>$M_k$</th>
<th>$N$</th>
<th>$ESS$</th>
<th>$cv$</th>
<th>$\log \widehat{ML}_k$ prec.</th>
<th>Type of IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERR.UNSCEAR</td>
<td>100 000</td>
<td>52 847</td>
<td>$9.0 \times 10^{-6}$</td>
<td>0.012</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>10 500</td>
<td>8 319</td>
<td>$2.5 \times 10^{-5}$</td>
<td>0.020</td>
<td>Adaptive</td>
</tr>
<tr>
<td>ERR.Little</td>
<td>100 000</td>
<td>53 931</td>
<td>$9.0 \times 10^{-6}$</td>
<td>0.011</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>14 500</td>
<td>10 534</td>
<td>$2.6 \times 10^{-5}$</td>
<td>0.020</td>
<td>Adaptive</td>
</tr>
<tr>
<td>ERR.Littleexp</td>
<td>100 000</td>
<td>7 495</td>
<td>$1.2 \times 10^{-4}$</td>
<td>0.004</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>25 500</td>
<td>10 190</td>
<td>$5.9 \times 10^{-5}$</td>
<td>0.030</td>
<td>Adaptive</td>
</tr>
<tr>
<td>ERR.BEIR7</td>
<td>100 000</td>
<td>16</td>
<td>$6.3 \times 10^{-2}$</td>
<td>0.079</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>63 841</td>
<td>10 492</td>
<td>$8.0 \times 10^{-5}$</td>
<td>0.035</td>
<td>Adaptive</td>
</tr>
</tbody>
</table>

- Adaptive IS reaches a stable precision over models, with much less particles than simple IS.
- However, keep in mind that many importance draws have been discarded in the adaptive scheme.
# Simple vs. Adaptive IS: EAR models

<table>
<thead>
<tr>
<th>$M_k$</th>
<th>$N$</th>
<th>$ESS$</th>
<th>$cv$</th>
<th>$\log \hat{ML}_k$ prec.</th>
<th>Type of IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAR.BEIR7</td>
<td>100 000</td>
<td>71</td>
<td>$1.4 \times 10^{-2}$</td>
<td>0.475</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>75 500</td>
<td>10 030</td>
<td>$8.6 \times 10^{-5}$</td>
<td>0.036</td>
<td>Adaptive</td>
</tr>
<tr>
<td>EAR.Littleexp</td>
<td>100 000</td>
<td>252</td>
<td>$4.0 \times 10^{-3}$</td>
<td>0.248</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>49 500</td>
<td>10 140</td>
<td>$7.8 \times 10^{-5}$</td>
<td>0.035</td>
<td>Adaptive</td>
</tr>
<tr>
<td>EAR.Preston</td>
<td>100 000</td>
<td>111</td>
<td>$9.0 \times 10^{-3}$</td>
<td>0.376</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>101 500</td>
<td>10 031</td>
<td>$9.0 \times 10^{-5}$</td>
<td>0.037</td>
<td>Adaptive</td>
</tr>
<tr>
<td>EAR.Schneider</td>
<td>100 000</td>
<td>51 421</td>
<td>$9.0 \times 10^{-6}$</td>
<td>0.012</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>17 000</td>
<td>11 904</td>
<td>$2.5 \times 10^{-5}$</td>
<td>0.020</td>
<td>Adaptive</td>
</tr>
<tr>
<td>EAR.Schneiderexp</td>
<td>100 000</td>
<td>262</td>
<td>$3.8 \times 10^{-3}$</td>
<td>0.243</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>66 000</td>
<td>24 427</td>
<td>$2.6 \times 10^{-5}$</td>
<td>0.020</td>
<td>Adaptive</td>
</tr>
<tr>
<td>EAR.UNSCEAR</td>
<td>100 000</td>
<td>40 814</td>
<td>$1.5 \times 10^{-5}$</td>
<td>0.015</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>14 500</td>
<td>10 085</td>
<td>$3.0 \times 10^{-5}$</td>
<td>0.022</td>
<td>Adaptive</td>
</tr>
</tbody>
</table>
Convergence Issues : EAR.Preston

- **left**: Inverse of observed Fisher Information matrix
- **right**: Posterior distribution correlation matrix
- Fisher Info matrix is singular: \( \epsilon \) added to diagonal before inversion results in grossly overestimated variances and correlations
Convergence Issues: EAR.Preston model (contd)

- sorted log of normalized importance weights for simple Importance sampling
- As a result of overdispersed, overcorrelated importance distribution, unbalanced weights, with many zeros
Convergence Issues: EAR.Preston model (contd)

- Histogram of particles resampled according to importance weights
- Unbalanced weights result in spikes and noisy aspect of posterior approximation
Adaptive IS : EAR.Preston model

- sorted log of normalized importance weights for adaptive importance sampling
- As a result of adaptation and filtering, importance weights are much less dispersed
Adaptive IS : EAR.Preston model (contd)

- Histogram of particles resampled according to importance weights
- much better posterior approximation, reveals skewed marginals for certain parameters
AIC vs. BIC vs. posterior probabilities

| $M_k$           | AIC  | BIC  | $p(M_k|y)$ | AIC  | BIC  | $p(M_k|y)$ |
|-----------------|------|------|------------|------|------|------------|
| ERR.UNSCEAR     | 0.608| 1.0  | 0.988      | 0.612| 1.0  | 0.988      |
| ERR.Little      | 0.126| 0.0  | 0.011      | 0.127| 0.0  | 0.011      |
| ERR.Littleexp   | 0.259| 0.0  | 0.001      | 0.261| 0.0  | 0.001      |
| ERR.BEIR7       | 0.0  | 0.0  | 0.0        | 0.0  | 0.0  | 0.0        |
| EAR.BEIR7       | 0.0  | 0.0  | 0.0        | 0.0  | 0.0  | 0.0        |
| EAR.Littleexp   | 0.0  | 0.0  | 0.0        | 0.009| 0.0  | 0.0        |
| EAR.Preston     | 0.0  | 0.0  | 0.0        | 0.0  | 0.0  | 0.0        |
| EAR.Schneider   | 0.004| 0.0  | 0.0        | 0.572| 0.0  | 0.0        |
| EAR.Schneiderexp| 0.003| 0.0  | 0.0        | 0.396| 0.0  | 0.0        |
| EAR.UNSCEAR     | 0.0  | 0.0  | 0.001      | 0.023| 1.0  | 1.0        |

- **Left**: Weights normalized over all models
- **Right**: Weights normalized over ERR and EAR models separately
### ERR vs. EAR models

| $M_k$            | $AIC$ | $BIC$ | $p(M_k|y)$ |
|------------------|-------|-------|------------|
| ERR.UNSCEAR      | 1.0   | 1.0   | 0.999      |
| EAR.UNSCEAR      | 0.0   | 0.0   | 0.001      |
| ERR.BEIR7        | 0.179 | 0.0   | 0.0        |
| EAR.BEIR7        | 0.821 | 1.0   | 1.0        |
| ERR.Little       | 0.999 | 0.999 | 0.948      |
| EAR.UNSCEAR      | 0.001 | 0.001 | 0.052      |
| ERR.Littleexp    | 1.0   | 1.0   | 1.0        |
| EAR.Littleexp    | 0.0   | 0.0   | 0.0        |

- as expected, ERR models strongly outperform EAR models...
- ...except for BEIR7 models...
- ...which however have zero global weights compared to the other models!
Posterior predictive medians (in blue) and associated 95% credible intervals (in grey) of the percentage of cases of childhood leukemia over period 1990-2004 in metropolitan France for the 10 models Female, Total exposure to NBR
Posterior predictive medians (in blue) and associated 95% credible intervals (in grey) of the percentage of cases of childhood leukemia over period 1990-2004 in metropolitan France for the 10 models Male, Total exposure to NBR
BMA vs. MMI vs. BIC averaging estimates of the percentage of cases of childhood leukemia over period 1990-2004 in metropolitan France for French women by attained age [0-14 years-old]
BMA vs. MMI vs. BIC averaging estimates of the percentage of cases of childhood leukemia over period 1990-2004 in metropolitan France for French men by attained age [0-14 years-old]
Model-averaged percentages (and 95% CI) of cases of childhood leukemia potentially attributable to radon, terrestrial gamma and cosmic rays over period 1990-2004 in metropolitan France and over childhood (from 0 to 14 years old)

<table>
<thead>
<tr>
<th>Components of natural radiation</th>
<th>Radon</th>
<th>terrestrial gamma rays</th>
<th>cosmic rays</th>
<th>all 3 exposures together</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Males</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of attributable cases</td>
<td>5.5</td>
<td>11.3</td>
<td>6.9</td>
<td>20.5</td>
</tr>
<tr>
<td>Posterior predictive median</td>
<td>(0-36.1)</td>
<td>(0-53.6)</td>
<td>(0-42.0)</td>
<td>(0-67.6)</td>
</tr>
<tr>
<td><strong>Females</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of attributable cases</td>
<td>5.3</td>
<td>11.4</td>
<td>6.9</td>
<td>20.4</td>
</tr>
<tr>
<td>Posterior predictive median</td>
<td>(0-36.2)</td>
<td>(0-54.6)</td>
<td>(0-43.2)</td>
<td>(0-68.0)</td>
</tr>
</tbody>
</table>
1. Context

2. Motivating Case Study

3. Method

4. Results

5. Discussion
Conclusions (1)

- Point predictions suggest that a sizeable proportion (20%) of childhood leukemia cases might be attributable to radon, TGR and cosmic rays in France
  - So far, consistent with UK findings (Wakeford et al 2009)
  - BUT 95% credible intervals for predictions appear to be very large (95%CI=[0,68])
  - Results only valuable provided that radiation-related leukemia risk models can be transferred

→ Point predictions must be interpreted cautiously!

- Usual risk models uncertainty may be ignored to predict radiation-related childhood leukemia rates in a current population from LSS data → UNSCEAR 2006 ERR model strongly recommended

→ Still no way to validate risk prediction for childhood leukemia due to NBR : Data acquisition in progress in France.
Conclusions (2)

- A first approach to Bayesian model averaging for quantitative radiation-related cancer risk assessment.
- A novel filtered adaptative Importance sampling approach based on multivariate Student proposals with varying degrees of freedom.
- Our approach allows to perform Bayesian inference and exact Bayesian model averaging for 10 radiation-induced leukemia risk models with many correlated parameters in a reasonable time frame.

**BUT**

- Convergence of our adaptive sampling scheme is not guaranteed
- For three models, our adaptive sampling scheme is not optimal
Perspectives

- Improve our adaptive sampling scheme (e.g., using non-central multivariate Student to account for skewed posterior marginals)
- Compare the strengths and weaknesses of the Bayesian Model Averaging and frequentist multimodel inference proposed by Burnham & Anderson (2004)