

Approche statistique bayésienne pour l'estimation d'un risque de cancer radio-induit après exposition aux examens scanner durant l'enfance

Anaïs Foucault, Marie-Odile Bernier, Sophie Ancelet

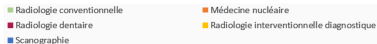
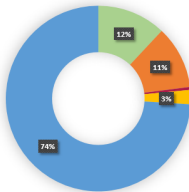
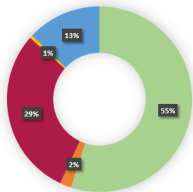
Institut de Radioprotection et de Sécurité Nucléaire (IRSN), PSE-SANTE/SESANE/LEPID
email: sophie.ancelet@irsn.fr

54èmes Journées de Statistique de la SFdS



Nombres d'actes

Dose efficace collective



- Increasing number of computed tomography (CT) examinations in France:
 - Frequency between 2012 and 2017 : +17% [IRSN,2017]
- 13% of diagnostic procedures but 74% of the annual effective dose
- Higher radio-sensitivity of children [UNSCEAR,2013]
 - 1% of children received at least one CT scan during the year over the 2012-2018 period [IRSN,2017]

Cancer risks following CT scans during childhood

- Many epidemiological studies are underway
 - **Identify** the existence of an association between the organ (radiation) doses following CT scans during childhood and a given cancer \Rightarrow **If it exists, low risk!**
 - **Estimate** the magnitude of this association (**and its uncertainty**)

Study	Patients	CNS	Leukaemia
France [Foucault et al. (2022)]	100 560	Yes	Yes
United-Kingdom [Pearce et al. (2012)]	180 000	Yes	Yes
Australia [Mathews et al. (2013)]	680 000	Yes	Yes
Germany [Krille et al. (2015)]	45 000	Yes	No
Netherlands [Meulepas et al. (2019)]	168 000	Yes	No
Taiwan [Li et al. (2020)]	130 000	Yes	No
EPICT [Hauptmann et al. (2023)]	659 000	Yes	

Cancer risks following CT scans during childhood

- Many epidemiological studies are underway
 - **Identify** the existence of an association between the organ (radiation) doses following CT scans during childhood and a given cancer \Rightarrow **If it exists, low risk!**
 - **Estimate** the magnitude of this association (**and its uncertainty**)

Study	Patients	CNS	Leukaemia
France [Foucault et al. (2022)]	100 560	Yes	Yes
United-Kingdom [Pearce et al. (2012)]	180 000	Yes	Yes
Australia [Mathews et al. (2013)]	680 000	Yes	Yes
Germany [Krille et al. (2015)]	45 000	Yes	No
Netherlands [Meulepas et al. (2019)]	168 000	Yes	No
Taiwan [Li et al. (2020)]	130 000	Yes	No
EPICT [Hauptmann et al. (2023)]	659 000	Yes	

A recurrent criticism : **Dosimetric uncertainty** not accounted for in risk estimates

How are the organ doses estimated ?

Example : The NCICT software [Lee et al. 2015]

File Batch Help

Patient parameters

Age:

Gender: Male Female

Height:

Weight:

Scanner parameters

Manufacturer:

Model:

Head filter Body filter

nCTDI_{vol} (mGy/100mAs):

Total collimation (mm):

Pitch:

Tube potential (kVp):

Current x Time (mAs):

CTDI_{vol} (mGy):

DLP (mGy·cm):

Effective Diameter (cm):

SSDE (mGy):

Predefined Protocol:

Scan Start (cm):

Scan End (cm):

Bar Graph

Copy organ list to Clipboard

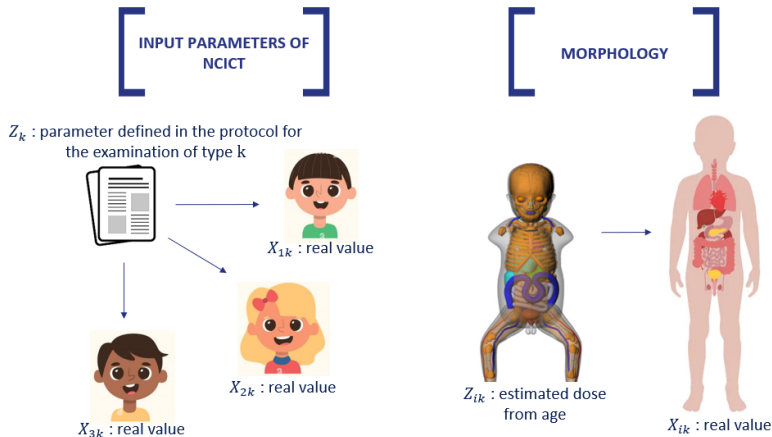
Organ absorbed dose (mGy)

Organ	Dose (mGy)
Brain	0.104
Pituitary gland	0.114
Lens	0.079
Eye balls	0.082
Salivary glands	0.459
Oral cavity	0.283
Spinal cord	3.806
Thyroid	9.187
Esophagus	5.745
Trachea	7.106
Thymus	8.109
Lungs	7.201
Breast	6.604
Heart wall	7.647
Stomach wall	7.798
Liver	7.553
Gall bladder	7.189
Adrenals	6.794
Spleen	7.533
Pancreas	7.059
Kidney	7.804
Small intestine	8.04
Colon	8.321
Rectosigmoid	5.846
Urinary bladder	5.658
Prostate	1.604
Uterus	0
Testes	0.169
Ovaries	0
Skin	2.931
Muscle	3.17
Active marrow	4.268
Shallow marrow	3.339
E60	5.387
E103	5.87

Developed by the National Cancer Institute (NCI)/National Institute of Health (NIH)

Different sources of dose uncertainty

- The input parameters of dosimetric models are uncertain
- The structure of the dosimetric models is uncertain



⇒ **The estimation of organ doses is uncertain**

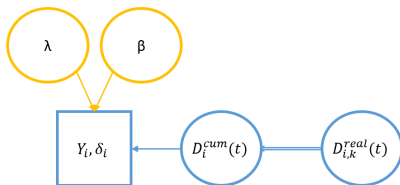
If not or poorly accounted for, dose uncertainty may cause [Carroll et al, 2006; Keogh et al, 2020] :

- a **bias** in risk estimates (generally towards the null);
- an **inadequate estimation** of the associated confidence/credible intervals;
- a **distortion** of the dose-response relationship [Hoffmann et al, 2017];

⇒ **Misleading conclusions** about the effect of organ (radiation) doses on the disease risk

- **Step 1:** Simulate plausible organ doses values for each CT examination from a dosimetric model like NCICT using a 2-dimensional Monte-Carlo algorithm [Simon et al. (2015); Lee et al. (2019); Thierry-Chef et al. (2021)]
- **Step 2:**
 - **Regression calibration** : averaged dose vectors (derived from all organ doses realisations) are used as plug-in estimates for the true organ doses when estimating disease risks [Little et al. (2014); Hauptmann et al. (2023)]
 - **Monte-Carlo Maximum Likelihood** : Estimate the risk coefficient by maximizing an approximation of the likelihood integrated over all organ doses realisations over a grid of fixed values for this coefficient [Stayner et al. (2007)]
- Asymptotic confidence intervals

- To propose an alternative approach :
 - **Bayesian learning of a model made of modules (e.g., multilevels/hierarchical models)** including the "black-box" dosimetric model provided by the NCICT software
⇒ Extension of the works by Richardson and Gilks, 1993; Hoffmann et al. 2017.
- To account for the difference between the patient's morphology and that of the NCICT phantom when modelling dose uncertainty



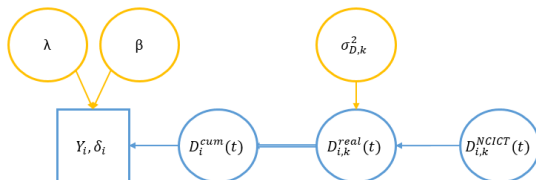
- Modelling the hazard rate of first diagnosis of a given cancer for patient i at age (in days) $t \in [0, +\infty[$

$$h_i(t; \beta) = \lambda \times \exp(\beta \times D_i^{cum}(t))$$

- $D_i^{cum}(t) = \sum_{u \leq t-x} D_{i,k}^{real}(u)$: : 2 or 5-years lagged cumulative radiation dose to the brain or the red bone marrow (RBM) (in Grays)
- $D_{i,k}^{real}(t)$: the true (unknown) organ dose received by patient i at age t for a CT examination of type k
- $\lambda > 0$: Baseline hazard rate **supposed to be constant over time**
- β : **Unknown risk coefficient**
- Assumption : Non-informative censoring

Accounting for dose uncertainty due to the patient's morphology

Let $D_{i,k}^{NCICT}$ be the organ dose estimated by the NCICT model for a CT examination of type k and a typical morphology (i.e., phantom selected from the age of patient i).



A multiplicative Berkson error model: $D_{i,k}^{real}(t) = D_{i,k}^{NCICT} \times B_k^D \times U_{i,k}^D(t)$

- $U_{i,k}^D(t) \stackrel{iid}{\sim} \text{logNormal} \left(-\frac{\sigma_{D,k}^2}{2}, \sigma_{D,k}^2 \right) \Rightarrow E(U_{i,k}^D(t)) = 1$
- Homoscedastic : $\sigma_{D,k}^2 = \sigma_D^2, \quad B_k^D = B^D, \quad \forall k$
- Heteroscedastic : $\sigma_{D,k}^2 = \sigma_{D,C(k)}^2, \quad B_k^D = B_{C(k)}^D, \quad \forall k \in C(k)$ (age, scanned body region)

A restricted version of the NCICT model

$$D^{NCICT}(organ, age, gender, model, kVp, filter, start, end, mAs, pitch) \\ = \sum_{z=start}^{end} DC(organ, age, gender, kVp, filter, z) \times \frac{nCTDI_w(model, kVp, filter)}{pitch} \times \frac{mAs}{100}$$

- $DC(organ, age, gender, kVp, filter, z)$: dose coefficients estimated by Monte Carlo simulations. **A table of previously estimated values is available.**
- $nCTDI_w(model, kVp, filter)$ is the CT Dose Index per volumetric unit, normalised to 100 mAs and **selected from the CTDI library [Lee and others, 2014].**

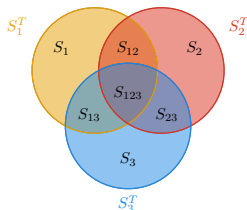
A preliminary step: Identify the NCICT software input parameters that are most influential when estimating organ doses

A variance-based global sensitivity analysis

Let $\mathbf{X} = (X_1, \dots, X_p)$ be the NCICT input parameters and Y the univariate outcome (a specific organ dose) of the (deterministic) dosimetric model f where $Y = f(\mathbf{X})$.

- Sobol' indice associated to input i [Sobol (1993)]:

$$S_i = \frac{\mathbb{V}[\mathbb{E}(Y|X_i)]}{\mathbb{V}(Y)}$$



⇒ When the input parameters are independent, Sobol' indices assign to each input parameter X_i a share of the variance of the organ dose estimate (Y).

- BUT some NCICT input parameters are dependent ⇒ We no longer have

$$1 = S_1 + S_2 + S_3 + S_{12} + S_{23} + S_{13} + S_{123}$$

Shapley effects [Owen (2014); Iooss et Prieur (2019)]

For any i in $D = \{1, \dots, p\}$:

$$Sh_i = \frac{1}{p} \sum_{A \subseteq D_{-i}} \binom{p-1}{|A|}^{-1} (S_{AU\{i\}}^T - S_A^T) \quad S_A^T = \frac{\mathbb{E}[V(Y|X_{\bar{A}})]}{V(Y)}$$

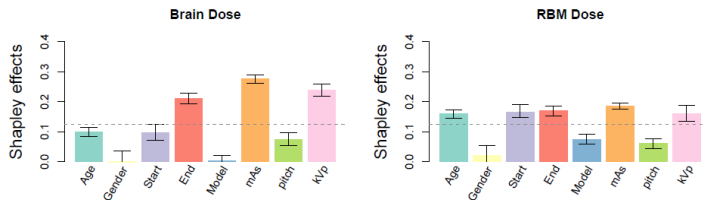
Two main properties:

- $Sh_i \geq 0$
- $\sum_i Sh_i = 1$

⇒ An interaction effect is equitably shared between each parameter involved in this interaction.

Application to the NCICT model (1/2)

- Estimation of the Shapley effects from a **single i.i.d input-output sample of PACS** (Picture archive and communication system) **data**
- Consistent "given-data" estimates of the total Sobol' indices using a **nearest-neighbor procedure** [Broto et al. (2020); Il Idrissi et al. (2021)]
- *sensitivity* R package, and in particular the *shapleysobol_knn* function, were used to estimate the Shapley effects
- Estimated Shapley effects for brain dose estimates associated with head examinations



- **Identification of most influential parameters for each examination category :**

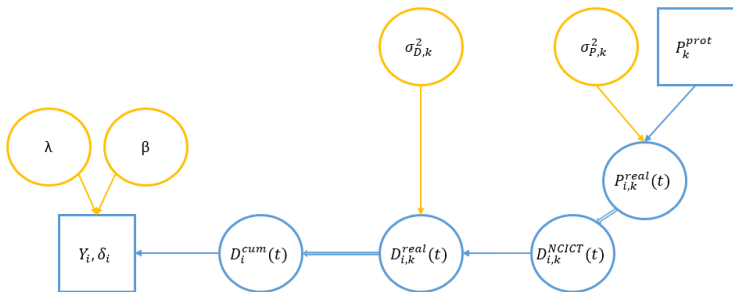
Anatomical area	Brain dose	Bone marrow dose
Head	mAs, kVp, end	mAs, kVp, start, end, age
Chest	mAs, kVp, pitch	mAs, pitch
Abdomen-pelvis	kVp	kVp
Multiple	start	start

- **Proportional Marginal Effects** [Herin et al. (2022)] were also estimated :
 - known to be less sensitive to correlations and to provide more pronounced hierarchy than Shapley effects in highly correlated cases;
 - tends to favor the inputs proportionally to their marginal contributions to every possible "coalitions" ⇒ Proportional redistribution of interaction effects

⇒ **Similar conclusions**

Continuous parameters uncertainty model

(mAs, pitch, $CTDI_{vol}$)

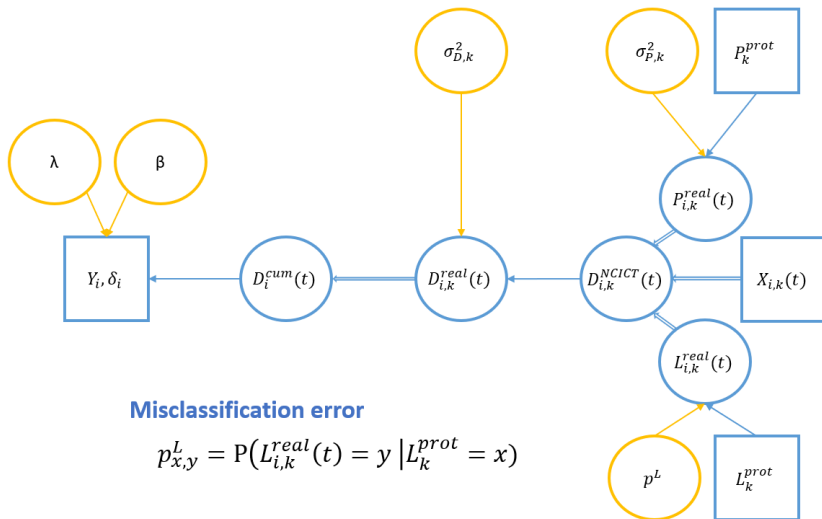


A multiplicative Berkson error model : $P_{i,k}^{real}(t) = P_k^{prot} \times B_k^P \times U_{i,k}^P(t)$

- $U_{i,k}^P(t) \stackrel{iid}{\sim} \logNormal\left(-\frac{\sigma_{P,k}^2}{2}, \sigma_{P,k}^2\right) \Rightarrow E(U_{i,k}^P(t)) = 1$
- Homoscedastic : $\sigma_{P,k}^2 = \sigma_P^2, B_k^P = B^P, \forall k$
- Heteroscedastic : $\sigma_{P,k}^2 = \sigma_{P,C(k)}^2, B_k^P = B_{C(k)}^P, \forall k \in C(k)$

Discrete parameters uncertainty model

(kVp, landmarks of the scanned body region)



Parameter	Prior distribution
Disease submodel	
$\lambda = \lambda' \times 10^{-7}$	$\lambda' \sim \mathcal{G}(1, 1)$
β	$\mathcal{N}(0, 10^6)$
Morphology uncertainty submodel	
σ_D	$\mathcal{HC}(0, 5)$ [Gelman (2006)]
$\sigma_{D,C(k)}$	$\mathcal{HC}(0, Z)$, $Z \sim U(0, 100)$
B^D	$\mathcal{LN}(-50, 100)$
$B_{C(k)}^D$	$\mathcal{LN}(-s^2/2, s^2)$, $s \sim U(0, 100)$
NCICT input parameters uncertainty submodel	
σ_P	$\mathcal{HC}(0, 5)$
$\sigma_{P,C(k)}$	$\mathcal{HC}(0, Z)$, $Z \sim U(0, 100)$
B^P	$\mathcal{LN}(-50, 100)$
$B_{C(k)}^P$	$\mathcal{LN}(-s^2/2, s^2)$, $s \sim U(0, 100)$
$p_{x, \cdot}^L$	$Dirichlet(0, 5, \dots, 0, 5)$

- **High-dimensional posterior distribution**
 - Over 500,000 latent variables to be estimated (but informed by external PACS data)
- Around 170,000 pseudo-observations
- **Hamiltonian Monte-Carlo algorithm** implemented in the R package *rstan*
 - Targeted Effective Sample Size : 4000
 - Importance of **model parameterization** ! (Ex : using standard normal distributions, 20 → 2 days of calculations for the most complex model)
- Comparison of the predictive performance of proposed models using the **Leave-one-out information criterion** (LOOIC) [Vehtari et al. (2017)]

- Set up in France in 2009 by IRSN
- 100 560 patients who:
 - Received at least one CT scan for a non-cancer pathology
 - Before the age of 10 years
 - Between 2000 and 2011
 - In one of the 21 participating university hospitals.
- Follow-up until a diagnostic of cancer, the death, the 18th birthday or the 31st December 2016



Some statistics

- **50 cases** of central nervous system (CNS) tumours
- **35 cases** of leukaemia
- Mean age at inclusion : 3 years / Mean follow-up : 9 years
- Patients **without cancer predisposing factors** : 97%

A transfer learning approach

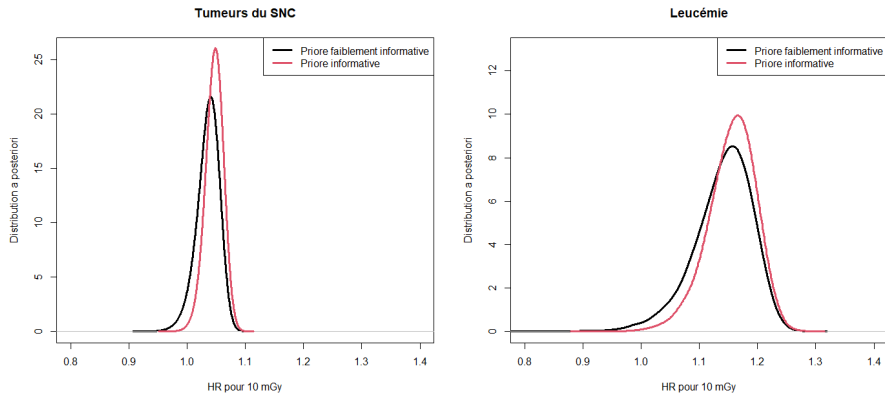
- No pair of values ($P_{i,k}^{real}(t), P_k^{prot}$) is available in the French CT cohort database
⇒ **Information required to learn about** $\sigma_P, \sigma_{P,C(k)}, B^P, B_{C(k)}^P$!
- Idem for $\sigma_D, \sigma_{D,C(k)}, B^D, B_{C(k)}^D, p_{x,y}^L$
- 8848 CT examinations collected in the PACS database ⇒ True technical parameters values associated with CT examinations + associated radiological protocols ⇒ **Pair of values** ($P_{i,k}^{real}(t), P_k^{prot}$) **were defined.**
- Different morphologies were simulated from growth curves: dose estimation based on age (typical phantom morphology) versus dose estimation based on weight and height ⇒ **Pair of values** ($D_{i,k}^{real}(t), D_i, k^{NCICT}$) **were defined**
- Assumption: These pairs of values follow **the same multiplicative Berkson error models** than in the French CT cohort
- **Joint Bayesian inference** of our multilevel models from both the French CT cohort data and the PACS data

Model	HR* per 10 mGy (95% CI)	LOOIC
CNS tumors (cases=50)		
Frequentist estimation	1,05 (1,01 - 1,09)	
No effect of radiation exposure ($\beta = 0$)		1381
Uncorrected	1,04 (0,97 - 1,08)	1382
Berkson homoscedastic	1,04 (0,97 - 1,07)	1611
Berkson + Bias heteroscedastic ^a	1,03 (0,97 - 1,06)	1383
Leukaemia (cases=35)		
Frequentist estimation	1,17 (1,07 - 1,27)	
No effect of radiation exposure ($\beta = 0$)		803
Uncorrected	1,15 (0,99 - 1,23)	1172
Berkson homoscedastic	1,15 (1,02 - 1,22)	789
Berkson + Bias heteroscedastic ^a	1,10 (1,01 - 1,16)	791

^a Latent structure with the smallest LOOIC on PACS data.

* Hazard Ratio: $\exp(\beta \times 10)$

Construction of informative prior probability distributions from the English cohort



Posterior probability distribution of the Hazard Ratio per 10 mGy

Aim :

- 1 Evaluate the ability of LOOIC to make a relevant model selection
- 2 Evaluating the effect of misspecification of Berkson error models on risk estimation

Aim :

- 1 Evaluate the ability of LOOIC to make a relevant model selection
- 2 Evaluating the effect of misspecification of Berkson error models on risk estimation

Organ doses were the ones from the French CT cohort :

- 1 Without adding a multiplicative Berkson error on the NCICT input parameters
 - 2 After multiplying the NCICT input parameters :
 - by an homoscedastic Berkson error component
 - by an heteroscedastic Berkson error component and a systematic bias component

⇒ True values of Berkson error model parameters = Posterior medians estimated from the French CT cohort
- Generate failure times for 10,000 individuals according to the Cox disease model to obtain cancer incidence rates similar to the ones observed in the French CT cohort.
 - Generate 100 replicated datasets for each scenario

Ability of LOOIC to make a relevant model selection

Estimation model	$\hat{\beta}^a$	Mean relative bias	95% coverage rate	Model selection ^b
		$\beta = 0,005$		
$\beta = 0$ Cox	0,005	0,01	0,98	0 99
		$\beta = 0$		
$\beta = 0$ Cox ($\beta > 0$)	10^{-3}			0 0

^aPosterior median averaged over 100 replicated datasets

^bProportion of simulated data sets for which the LOOIC of the simulation model is lower than the LOOIC of the other competing model with a difference of at least 8 points [Sivula et al. (2020)]

Impact of Berkson error model misspecification on risk estimation

Estimation model	$\hat{\beta}^a$	Mean relative bias	95% coverage rate
Berkson error (homoscedastic)			
Cox, $\beta = 0,02$			
Uncorrected	0,02	-0,01	98
Berkson error (homoscedastic)	0,02	0,03	100
Berkson error + bias (heteroscedastic)			
Cox, $\beta = 0,04$			
Berkson error (homoscedastic)	0,01	-0,65	0
Berkson error + bias (heteroscedastic)	0,04	-0,02	94

^aPosterior median averaged over 100 replicated datasets

Bayesian inference of the proposed multilevel models on the French CT cohort showed:

- A weak but statistically significant positive association between CT scans exposure during childhood and the risk of leukaemia
- No statistically significant association between the risk of CNS tumours and CT scans exposure
- A weak impact of dosimetric uncertainties on cancer risk estimates

Strengths:

- Use of two recent variance-based global sensitivity indices
- Use of a Bayesian multilevels approach :
 - Flexible modelling approach to describe different sources of uncertainty
 - Allows to integrate external information
 - Allows for the joint inference of all unknown quantities (e.g., true organ doses and risk)

Weaknesses :

- Using the LOOIC for model selection in case of multilevel models ? What else?
- Transfer learning : Possible underestimation of the variance of dose uncertainty. What about adding similarity parameter(s) ?

Thank you for your attention !

sophie.ancelet@irsn.fr