POOLED BAYESIAN META-ANALYSIS OF TWO POLISH STUDIES ON RADIATION INDUCED CANCERS

Krzysztof Wojciech Fornalski a

* PGE EJ Sp. z o.o., ul. Mokotowska 49, 00-542 Warszawa, Polska.
krzysztof.fornalski@gkpg.pl, krzysztof.fornalski@gmail.com

Abstract

The robust Bayesian regression method was applied to perform meta-analysis of two independent studies on influence of low ionizing radiation doses on the occurrence of fatal cancers. The re-analyzed data comes from occupational exposure analysis of nuclear workers in Świerk (Poland) and from ecological study of cancer risk from natural background radiation in Poland. Such two different types of data were analyzed and three popular models were tested: constant, linear and quadratic dose-response dependencies. The Bayesian model selection algorithm was used for all models. The Bayesian statistics clearly indicates that the popular linear no-threshold (LNT) assumption is not valid for presented cancer risks in the range of low doses of ionizing radiation.

Methods

The robust Bayesian regression analysis assumes that each experimental point $E_i$ with original uncertainty $\sigma_i$ is given by appropriate Gaussian distribution (Bayesian likelihood function) and a prior function for the uncertainty $p(\sigma) = \sigma \cdot \sigma^{-2}$ (Fornalski and Dobrzyński 2011). Thanks to this assumption all potential outlier points give insignificant input to the general posterior probability distribution $P = \prod P_i$:

$$P = \prod \frac{\sigma_i}{\sigma_i} \exp\left(\frac{(M_i - E_i)^2}{2\sigma_i^2}\right)$$

where $M$ corresponds to the model (curve) being fitted to existing points.

Having established the posterior probability distribution, one can adapt it into a model selection algorithm described in (Fornalski and Dobrzyński 2011) and calculate the plausibility of assumed model $M$ as:

$$P(M|Data) \propto P(Data|M) = \sum P_i(M_i|Data) \prod_{i=1}^{N} \frac{\sigma_i}{\sigma_i} \exp\left(\frac{(M_i - E_i)^2}{2\sigma_i^2}\right) \equiv N_M$$

With the plausibilities of two models (namely A and B), one can find the relative likelihood

$$W = \frac{N_A}{N_B} = \frac{P(M_A|Data)}{P(M_B|Data)}$$

which is higher than 1, when model A is more likely than B etc.

Results

The figures above present the natural gamma terrestrial radiation in Poland (upper left), the low- and high-radiation (radon incl.) counties in Poland (upper right), the nuclear center in Świerk (bottom left) and the relationship between cancer incidences and cumulative doses of Świerk’s employees (bottom right).

The model selection algorithm was used for three models and plausibility values $N_A$ are presented in Table (left). The constant model and linear model have the plausibility on the same level. The plausibility of quadratic model (parabola) is higher, however, the results of regression give an inverted parabola (c>0). Thus the quadratic model needs to be rejected as it contradicts physics.

Finally, the constant fit RR=98.9% as well as linear one RR=100.2% - 1.27% D mSv^{-1} should be treated with the same likelihood. However, the negative slope of linear model (-1.27%/mSv/year) is strongly statistically significant (p<0.12, Bayesian statistics) and differs from the LNT assumption used in radiation protection standards.

Data

In the presented analysis the study regarding cancer risk due to natural background radiation in Poland (Fornalski and Dobrzyński 2012) will be referred to as Study I. The second study regarding the occupational exposure in Świerk nuclear centre (Fornalski and Dobrzyński 2013) will be referred to as Study II.

The pooled data of both studies is presented in Figure below. The first point on left in Figure corresponds to the Study II while rest of points are Study I.

![Data Figure]

The relative risk of lung cancer risk for the excess effective dose in the pooled Polish studies. The dotted line corresponds to the linear no-threshold (LNT) hypothesis while solid line is a result of robust Bayesian linear regression analysis of the data. All presented uncertainties are given 68% CI.

![Results Figure]

The figures above present the natural gamma terrestrial radiation in Poland (upper left), the low- and high-radiation (radon incl.) counties in Poland (upper right), the nuclear center in Świerk (bottom left) and the relationship between cancer incidences and cumulative doses of Świerk’s employees (bottom right).

![Results Figure 2]

The model selection algorithm was used for three models and plausibility values $N_A$ are presented in Table (left). The constant model and linear model have the plausibility on the same level. The plausibility of quadratic model (parabola) is higher, however, the results of regression give an inverted parabola (c>0). Thus the quadratic model needs to be rejected as it contradicts physics.

Finally, the constant fit RR=98.9% as well as linear one RR=100.2% - 1.27% D mSv^{-1} should be treated with the same likelihood. However, the negative slope of linear model (-1.27%/mSv/year) is strongly statistically significant (p<0.12, Bayesian statistics) and differs from the LNT assumption used in radiation protection standards.

References