On methodological problems in seeking for a correlation of lung cancer with radon level

J. Reszczyńska ı, M. Pylak ı/2, K.W. Fornalski ı/2, S.J. Mortazavi ı, L. Dobrzyński ı

1 National Centre for Nuclear Research (NCBJ), ul. A. Soltana 7, 05-400 Grodzisk-Swidnik, Poland
2 Institute of Physics, Polish Academy of Sciences (PAS), Al. Lotników 126/128, 02-668 Warsaw, Poland
3 Ex-Polon Laboratory, ul. Podziemia 6, 05-552 Łaz, Poland
4 Fox Chase Cancer Center, Philadelphia, PA, USA

Abstract

The relationship between lung cancer morbidity and radon concentration is an open subject of many researches all over the world. Secondary risk factors may limit classical statistical analysis, thus many authors do not support ecological results. To address this problem, the authors of this paper focused on the methodological aspects used for interpretation of such results when a rich collection of various explanatory parameters is available. This is e.g. the case of the county-level dataset collected by Simeonov and Himmelstein in 2015 [1], that contains geographically aggregated data on cancer risk factors, environmental features, demographics, collected for 3 142 U.S. counties and county-equivalents.

Apart from the Least-Square and Bayesian linear regression analysis, for the first time the Maximum Entropy Method (MEM) was used for interpretation of the influence of the correlated concentration of radon, altitude and UVB level on the lung cancer morbidity. An importance of binning the data is also shown. The binning of the data may be critical for final conclusions. For example, the negative slope of the fitted line can change to the positive one by a slight modification of the data binning.

It follows from both, LSQ and Bayesian methods of analysis that a decrease of cancer morbidity with increasing radon concentration is minute in the group of the lowest smoking prevalence, and it is statistically sound in the group of high smoking prevalence. This conclusion follows from both, LSQ and Bayesian, methods of analysis. The use of MEM provided much richer picture: a clear trend of decreasing morbidity of lung cancer with increasing radon density level independently of the altitude of the place of residence and the level of UVB. Moreover, this trend does not depend on the prevalence of smoking and the sex.

Results

Data for selected subgroups [1], after reduction of the altitude band to 200-275 m, were initially carefully analyzed using the classical least squares regression (LSQ) method and robust Bayesian regression methodology [2]. The only models taken into account for the dose-effect dependence were a linear dependence and constant - no dependence on radon concentration.

In all cases, except the population with the lowest smoking percentage, a negative slope is observed in both methods (see Fig. 3). One notes that in all 3 subgroups cancer incidence is substantially higher in men than in women. Nevertheless, reducing the data to a certain percentage of smokers, not essentially different for men and women, one can still observe a statistically significant stronger decrease in lung cancer with increasing radon concentration in men (Fig. 4) which indicates that women exhibit lower risks than men.

All those findings were supported by the analysis performed by the Maximum Entropy Method (Fig. 5).

The presented paper was submitted to the International Journal of Low Radiation (2019).

REFERENCES: