A Simulation Study of Lognormal Measurement Error Effect—Discrimination Problem of Radon and Thoron—

Kazutaka Doi*, Shinji Tokonami, Yosuke Kobayashi, Hidenori Yonehara, and Shinji Yoshinaga

aNational Institute of Radiological Sciences, Regulatory Research Sciences Research Group
bNational Institute of Radiological Sciences, Environmental Radioactivity Survey Section
4-9-1, Anagawa, Inageku, Chiba-Shi, 263-8555 Japan.

Abstract. Several case-control studies have indicated an increased risk of lung cancer linked to indoor radon exposure. For a precise evaluation of radon-related lung cancer risk, however, contribution of thoron should be considered. There are a lot of studies in which passive type radon detectors are used without thoron discrimination techniques. Therefore, these passive type radon detectors may be strongly affected to the presence of thoron in case they are installed near a wall or floor as potential thoron sources. The thoron effect we consider here is an increase of radon signal in the radon detectors without the discrimination technique. This problem is classified as a part of measurement error problem in statistical literatures and the thoron is considered as a possible source of measurement error. In general, concentrations of radon and thoron follow a lognormal distribution. The effects of measurement error following normal distribution have been studied well, but there are few studies on measurement error following non-normal distribution. In order to evaluate the effect of measurement error due to thoron, we conducted a simulation study. We assumed a case-control study of lung cancer and indoor radon with hypothetical data where radon concentrations were measured with and without discrimination of thoron concentrations. We also assumed that logistic regression was used, and that the concentrations of radon and thoron were correlated, following lognormal distribution. The thoron disturbance in radon measurement resulted in an approximately 90% downward bias in the effect of radon, and this bias was almost constant when the parameters were varied. The downward bias was consistent with results from previous studies taking measurement error following normal distribution into account. It was confirmed that the effect of lognormal measurement error is concordant with the normal measurement error in this case.

KEYWORDS: measurement error; lognormal distribution; logistic model; radon; thoron.

1. Introduction

In most countries, radon is the largest natural source of exposure to ionizing radiation for the general population, and several case-control studies have indicated an increased risk of lung cancer linked to indoor radon exposure [1]. Although most of them did not show a significant risk [2]. A lack of statistical power is one of the reasons for the non-significant results. However, there are other reasons, including bias and confounding related to exposure assessment, for these non-significant results in each case-control study.

Spatial and temporal variation in radon concentrations is one of the main sources of uncertainty in radon measurements. The dependence of lung cancer risk on an explanatory variable, i.e., radon concentration in the presence of several confounding variables is investigated using regression models. When an explanatory variable is measured with error, the estimated regression coefficient is usually biased toward zero. This problem is well known as “attenuation” in the statistical literature, and a lot of methods have been proposed for correcting this downward bias [3-4].

Recently it has become aware that the reading of passive radon detectors that do not employ thoron discrimination techniques is affected by thoron. Because the presence of thoron in houses can be another source of uncertainty in radon measurements, the contribution of thoron should be considered for a precise evaluation of radon-related lung cancer risk [5-6]. Passive radon detectors without thoron

* Presenting author, E-mail: kazutaka@nirs.go.jp
discrimination techniques have been used in many studies, indicating that radon concentrations were likely overestimated.

Because we usually stay away from the walls of our house, the concentration of thoron in the place where we live is considered to be low. The thoron effect we consider here is an increase of the reading of the radon detectors without the discrimination techniques. This problem is classified as a part of measurement error problem in the statistical literature, and thoron is considered as a possible source of measurement error in this case.

It is known that the effects of measurement error vary depending on various factors, including the statistical distribution of the measurement error, analysis model, and correlation between variables, and whether the measurement error is differential or nondifferential. In general, radon and thoron concentrations follow a log-normal distribution. The effects of measurement error that follows normal distribution have been well studied, but there are few studies on measurement error following non-normal distributions. The measurement error in the statistical literature usually indicates the increased variation in observed variables, but that of thoron causes not only the increased variation but also the overestimate of radon concentrations. Therefore, we conducted a simulation study to evaluate the effect of measurement error due to thoron.

2. Methods

2.1 Simulation Settings

We assumed a case-control study on lung cancer and indoor radon. First, we generated radon (\(Rad_i\)) and thoron concentrations (\(Ton_i\)), which followed log-normal distributions. For the radon concentration, we set the geometric mean (GM) at 30, 60, 120, and 240 Bq/m\(^3\), and geometric standard deviation at 1.5, 2.0, 2.5, and 3.0 Bq/m\(^3\). For the thoron concentration, we set the geometric mean at 50, 150, 300, and 600 Bq/m\(^3\), and the geometric SD (GSD) at 1.5, 2.0, and 3.0 Bq/m\(^3\). Using a single a single passive radon detector, the observed concentration (\(Obs_i\)) can be expressed by the following equation:

\[
Obs_i = Rad_i + \frac{CF_{\text{Ton}}}{CF_{\text{Rad}}} \cdot Ton_i
\]

where \(CF_{\text{Rad}}\) is the conversion factor for the radon concentration (track cm\(^{-2}\) kBq\(^{-1}\) m\(^3\) h), and \(CF_{\text{Ton}}\) is the conversion factor for the thoron concentration (track cm\(^{-2}\) kBq\(^{-1}\) m\(^3\) h) \([5]\). We assumed 5 different passive radon detectors: SSI/NRPB (\(CF_{\text{Ton}}/CF_{\text{Rad}} = 0.06/2.2\)), regular Radopod (\(CF_{\text{Ton}}/CF_{\text{Rad}} = 0.10/2.62\)), modified Radopod (\(CF_{\text{Ton}}/CF_{\text{Rad}} = 1.32/2.64\)), Radtrak (\(CF_{\text{Ton}}/CF_{\text{Rad}} = 1.88/2.81\)), and a KfK monitor (\(CF_{\text{Ton}}/CF_{\text{Rad}} = 0.70/0.85\)) \([7]\).

The simulation was based on 1000 replications. The radon and thoron concentrations were correlated in a logarithmic scale, and the correlation coefficients were varied from \(-0.8\) to \(0.8\). We assumed the logistic model and generated a disease status (\(Di\)) from the Bernoulli distribution with probability \(p_i\):

\[
p_i = \frac{\exp(\alpha + \beta Rad_i)}{1 + \exp(\alpha + \beta Rad_i)}
\]

where \(\exp(\alpha)/(1 + \exp(\alpha))\) is the probability of developing lung cancer without radon exposure (\(\alpha = \log(0.5/(1 - 0.5)), \log(0.1/(1 - 0.1)), \log(0.02/(1 - 0.02)), \log(0.01/(1 - 0.01))\)), i.e., the background prevalence of lung cancer, \(\beta\) is the true risk coefficient for radon (\(\beta = 0.0005, 0.001, 0.002, 0.004, 0.008\)), and \(Di = 1\) if the \(i\)-th subject developed lung cancer and \(Di = 0\) otherwise.

The \(\beta\) value of 0.001 roughly corresponds to the results from North American case-control studies (excess odds ratio 0.11/100 Bq/m\(^3\)) and European case-control studies (excess odds ratio 0.084/100 Bq/m\(^3\)) \([2, 8]\). The \(\beta\) value of 0.002 roughly corresponds to that from the LIH-NCI study (excess odds ratio 0.19/100 Bq/m\(^3\)) \([9]\). For example, the \(i\)-th subject, whose radon concentration is \(Rad_i\), develops lung cancer with the probability \(p_i\) obtained from the above equation, and does not develop lung cancer with the probability \((1 - p_i)\).
Then we generated a sufficient number of subjects to be divided into cases ($D_i = 1$) and controls ($D_i = 0$), and the controls were matched to each case. The number of cases was varied from 250 to 8000. The matching ratio was also varied from 1:1 to 1:5.

### 2.2 Analysis models

The data analysis was conducted using conditional logistic regression. Using this model, the data were analyzed in two ways. One was that the explanatory variable in the model was $\text{Rad}_i$, the true radon concentration (model 1), and the other was that $\text{Obs}_i$, the observed radon concentration (model 2). The simulation results were evaluated in terms of the bias, difference between the estimate and the true value, standard error (SE), and 95% coverage probability. All analyses were performed using the SAS statistical software package (SAS 9.1.3, SAS Institute Inc., Cary, NC).

### 3. Results

The results are presented in Table 1. Each row of Table 1 reports the Monte Carlo mean bias (bias), standard error (SE), and coverage probability of the nominal 95% large sample confidence intervals (Coverage) according to the combinations of sample size and parameter values. Bias was divided by the true value, and the bias value of -0.5 corresponding to the underestimation, whose magnitude is half of the true value and that of 0.5 corresponding to the overestimation, whose magnitude is half of the true value.

<table>
<thead>
<tr>
<th>Case</th>
<th>Control</th>
<th>Alpha</th>
<th>Beta</th>
<th>GM (radon)</th>
<th>GSD (radon)</th>
<th>GM (thoron)</th>
<th>GSD (thoron)</th>
<th>Corr</th>
<th>Bias</th>
<th>SE</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>500</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 0.022 2.768 0.969</td>
<td>-0.016 0.734 0.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 0.016 1.553 0.948 0.017 0.515 0.041</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td>2000</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 -0.017 1.377 0.951 -0.014 0.363 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>4000</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 -0.006 0.973 0.950 -0.013 0.256 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4000</td>
<td>8000</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 -0.004 0.688 0.947 -0.010 0.181 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8000</td>
<td>16000</td>
<td>0.002</td>
<td>60 1.5</td>
<td>150 2.0</td>
<td>0.1 -0.001 0.486 0.952 -0.009 0.128 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The performance of the analysis with and without discrimination of radon from thoron, where sample size, alpha, beta, GM and GSD of radon and thoron, and correlation between radon and thoron were varied.
Table 2: The performance of the analysis with and without discrimination of radon from thoron, where sample size (matching ratio), and \( \text{CF}_{\text{Rn}}/\text{CF}_{\text{Tn}} \) were varied.

<table>
<thead>
<tr>
<th>Case</th>
<th>Control</th>
<th>Matching</th>
<th>( \text{CF}<em>{\text{Tn}}/\text{CF}</em>{\text{Rn}} )</th>
<th>GM (radon)</th>
<th>GSD (radon)</th>
<th>GM (thoron)</th>
<th>GSD (thoron)</th>
<th>( \text{Gerr} )</th>
<th>Bias</th>
<th>SE</th>
<th>Coverage</th>
<th>Bias</th>
<th>SE</th>
<th>Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>250</td>
<td>500 1:2</td>
<td>0.06/2.2</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>0.022 2.768 0.960</td>
<td>-0.011 2.705 0.961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>500 1:2</td>
<td>0.10/2.62</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>0.022 2.768 0.960</td>
<td>-0.035 2.866 0.964</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>500 1:2</td>
<td>1.32/2.64</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>0.022 2.768 0.960</td>
<td>-0.063 0.947 0.579</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>500 1:2</td>
<td>1.88/2.81</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>0.022 2.768 0.960</td>
<td>-0.016 0.734 0.279</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>500 1:2</td>
<td>0.70/0.85</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>0.022 2.768 0.960</td>
<td>-0.941 0.606 0.106</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000 1:2</td>
<td>0.06/2.2</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>-0.016 1.953 0.948</td>
<td>-0.049 1.910 0.950</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000 1:2</td>
<td>0.10/2.62</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>-0.016 1.953 0.948</td>
<td>-0.072 1.882 0.953</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000 1:2</td>
<td>1.32/2.64</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>-0.016 1.953 0.948</td>
<td>-0.087 0.666 0.257</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000 1:2</td>
<td>1.88/2.81</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>-0.016 1.953 0.948</td>
<td>-0.017 0.315 0.041</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>1000 1:2</td>
<td>0.70/0.85</td>
<td>60 1.5 150 2.0</td>
<td>0.1</td>
<td>-0.016 1.953 0.948</td>
<td>-0.041 0.426 0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As a whole, the thoron disturbance in radon measurement resulted in downward bias in the radon-related lung cancer risk. The SE of risk estimates from model 2 was generally smaller than that of model 1, indicating uncertainty in radon measurement resulted in apparent smaller uncertainty in estimates. The magnitude of bias from model 2 became larger as radon GM and GSD became smaller and thoron GM and GSD became larger. Estimated coverage probabilities of the 95% confidence intervals from model 1 were all close to 95%, with except for the situations of small sample size. On the other hand, those from model 2 had very poor coverage probabilities.

When the sample size of the cases and controls was varied from 250 to 800 and 500 to 16000, respectively (rows 1-6 in Table 1), the magnitude of bias from model 1 decreased as the sample size increased. On the other hand, when radon was contaminated with thoron, the magnitude of bias slightly decreased.

The background prevalence of lung cancer did not influence the results from both model 1 and model 2 (rows 7-11 in Table 1).

The background prevalence of lung cancer did not influence the results from both model 1 and model 2 (rows 7-11 in Table 1).

The bias of results from both model 1 and model 2 decreased as the magnitude of the true risk coefficient increased (rows 12-16 in Table 1).

The magnitude of bias from model 2 decreased as the correlation increased (rows 33-39 in Table 1).
As expected, the results from model 2 approached those from model 1 when the value of $\frac{CF_{Tn}}{CF_{Rn}}$ was small (top rows in Table 2).

The bias from model 1 was decreased as the number of control increased when the sample size was small (rows 16-20 in Table 2).

4 Discussion

We conducted a simulation study of the log-normal measurement error effect to evaluate the radon and thoron discrimination problem. These results confirmed the underestimation of the risk coefficient, suggested by a previous study. Although the magnitude of the underestimation varied a little with changes in parameter conditions, the underestimation was consistently observed as a whole. The downward bias in the estimates was consistent with results from previous studies of measurement error that followed a normal distribution. We confirmed that the effect of lognormal measurement error is consistent with the normal measurement error in this case.

As for the correlation between radon and thoron, we adopted 0.1 as the default value in the simulation, based on data from the Shanxi and Shaanxi Provinces [5]. When houses composed of soils and rocks rich in uranium and thorium are compared with houses with fewer soils and rocks, there is a correlation between the radon and thoron concentrations. However, the correlation was weakened because a variety of factors affected radon and thoron concentrations. Therefore, in this simulation, a correlation value of 0.1 was considered to be appropriate. When there was no correlation, the magnitude of bias from model 2 was slightly greater than when the correlation coefficient was 0.1. When radon and thoron were negatively correlated (correlation coefficient of $-0.8$), then the bias value exceeded $-1$, and we obtained a negative risk coefficient. However, negative correlations between radon and thoron are unlikely, and in the range of a weak positive correlation, a positive risk coefficient will be obtained.

Because the meta-analyses conducted in Europe and North America did not report about discrimination of radon and thoron, it is possible that their radon measurement were affected by cross sensitivity of the radon detector to thoron. Recently, these problems with thoron have attracted attention and are now widely recognized [10-11]. Future studies using meta-analysis and radon detectors with thoron discrimination techniques will solve these problems.

The purpose of this simulation was to evaluate the effect of thoron interference in situations in which thoron concentrations were and were not discriminated from radon concentrations. We cannot evaluate the risk of substance whose quantity is unknown. Therefore, to consider thoron contribution to lung cancer, it is the precondition that thoron concentrations are measured. This simulation focused on the comparison of the situations with and without thoron discrimination. Future studies will consider the contribution of thoron to lung cancer.

5 Conclusion

The thoron disturbance in radon measurement resulted in an approximately 90% downward bias in the effect of radon, and this bias was almost constant when the parameters were varied. The downward bias was consistent with results from previous studies taking measurement error following normal distribution into account.

Acknowledgements

We would like to express gratitude to Dr. Quanfu Sun for providing the Shanxi and Shaanxi provinces data.
REFERENCES


