



Original Article

Age or environmental radiation dose rate: Which is more correlated with cancer incidence rates in the Republic of Korea?



Han Young Joo, Jae Wook Kim, So Yun Jeong, Jin Sik Choi, Joo Hyun Moon*

Department of Energy, Dankook University, 119, Danda-e-ro, Dongnam-gu, Cheonan, Chungnam, 31116, Republic of Korea

ARTICLE INFO

Article history:

Received 14 February 2022

Received in revised form

22 March 2022

Accepted 23 March 2022

Available online 25 March 2022

Keywords:

Environmental radiation dose rate

Age

Cancer incidence rate

Big data analysis

Statistically significant relationship

ABSTRACT

Our study adopted a big data analysis approach to determine whether there was a significant relationship between environmental radiation dose rates or age and cancer incidence rates in the Republic of Korea. The data for this analysis included environmental radiation dose rates, number of cancer patients, and age distributions of the residents from 2009 to 2016 in the administrative districts where environmental radiation monitoring posts were located. For this analysis, the environmental radiation dose rates were obtained from 171 monitoring posts located in 113 elementary administrative districts in the Republic of Korea. The number of cancer patients and the age distributions were obtained from the Central Cancer Information Center of the National Cancer Center of Korea and the Ministry of the Interior and Safety, respectively. Our findings indicated that there was no statistically significant relationship between the environmental radiation dose rate and the cancer incidence rate. However, age had a considerable influence on the cancer incidence rate of the monitored regions.

© 2022 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Background radiation includes radiation from both naturally occurring sources and anthropogenic sources such as the fallout from nuclear weapon tests and release from the routine operation of nuclear reactors. As a result, all human beings are constantly exposed to background radiation [1–3].

In the context of radiation protection, the International Commission on Radiological Protection (ICRP) classifies the human health effects of radiation as either deterministic or stochastic. Deterministic effects, which are also referred to as non-stochastic effects, are direct health effects, the severity of which varies with the dose and for which a threshold is believed to exist. Radiation-induced cataract formation is an example of a deterministic effect. Stochastic effects are the health effects that occur by chance, generally occurring without a threshold level of radiation dose. The likelihood of these effects is proportional to the radiation dose but their severity is dose-independent. Cancer and genetic alterations are key examples of stochastic effects.

The linear zero-threshold dose-response model (also referred to as the LNT model) is often adopted to minimize stochastic effects.

This model assumes that the frequency of occurrence of a stochastic effect is linearly proportional to the radiation dose without a threshold [4–10]. Nevertheless, this model is only valid for a relatively high dose, ranging between 20 and 100 mGy or higher, whereas environmental radiation levels are generally quite low compared to the aforementioned range. Therefore, validating the accuracy of the LNT model at environmentally relevant radiation doses poses an important challenge [4].

To validate the effectiveness of the LNT model even at low radiation doses, such as those that occur in the environment, we examined whether the environmental radiation dose has a linear effect on the occurrence of cancer. Our Spearman correlation analysis results confirmed that there was no significant correlation between the environmental radiation dose and the number of relative cancer patients [11]. To supplement the findings of our previous study, additional analyses were conducted using the Pearson correlation method and including the age distribution as an analysis variable. Specifically, we determined which of the two variables, environmental radiation dose rate or age distribution, was more statistically correlated with cancer incidence rates in each administrative district in the Republic of Korea.

* Corresponding author.

E-mail address: jhmoon86@dankook.ac.kr (J.H. Moon).

2. Methods

2.1. Data

The variables for our analyses included environmental radiation dose rates, the proportion of the residents over the age of 40, and the number of cancer patients for each administrative district in the Republic of Korea from 2009 to 2016. The raw data were obtained through websites operated by Korean government organizations. All data were analyzed using the R language for statistical computing, a well-known programming language that is widely used in the field of big data analysis.

2.1.1. Environmental radiation dose rates

As of the end of 2016, the Korean central/local government and nuclear utility operators have installed and operated a total of 171 monitoring posts at 113 out of the 250 elementary administrative districts in the country, which monitor the environmental radiation dose rate of each district in $\mu\text{Sv/hr}$ [12–19]. The environmental radiation dose rate measured by each monitoring post was assumed to be a representative value for the entire administrative district in which the monitoring post was located.

2.1.2. Relative cancer incidence

The data for this variable was the number of annual cancer patients in each administrative district announced by the Central Cancer Information Center of the National Cancer Center of Korea (CCIC). The data were obtained through a request for information disclosure through a website operated by the CCIC. Since the population of each administrative district was different, the numbers of cancer patients could not be directly compared. Therefore, statistical comparisons were conducted using an equivalent variable. The variable was the relative cancer incidence per hundred thousand people in each district, which was obtained by considering the number of cancer patients and the resident registration number in each district. The relative cancer incidence per hundred thousand people in each district was obtained with Eq. (1):

$$\text{Relative Cancer Incidence} = \frac{x_i \times 100,000}{n_i} \quad (1)$$

where x_i is the number of cancer patients in the i -th year and n_i is the number of residents registered in each district in the i -th year.

2.1.3. Proportion of residents over the age of 40

Ageing is known to be among the most important risk factors for cancer, as well as for many individual cancer types. Our study sought to confirm whether this was also true in Korea. Specifically, we sought to confirm whether the cancer incidence rate increased significantly if there was a greater proportion of residents over the age of 40 in the administrative districts. The Korean Ministry of the Interior and Safety (MOIS) publishes the population of each administrative district by 10-year-old units every month [20]. We used these data and derived the portion of the residents over the age of 40 among all residents.

2.2. Statistical analyses

Correlation analysis was performed to confirm the degree of correlation between the two other variables and the cancer incidence rate using the R language. The Pearson's correlation coefficient (r) was used to identify negative or positive linear relationships between two variables [21]. The resulting r value was between -1 and 1 , and absolute values closer to 1 indicate an accurate correlation (either negative or positive). Conversely, 0 means

that there is no correlation between the two variables. The correlation coefficient was calculated using Eq. (2) [22]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2)$$

where x_i, y_i are the individual i -th data element of variables X and Y , respectively, and n is the total number of samples. \bar{x} and \bar{y} are the averages of variables X and Y , respectively. Eq. (2) allows for the calculation of both the Pearson's correlation coefficient and the p -value. As mentioned before, the degree of correlation according to the correlation coefficient values can be interpreted as shown in Fig. 1.

Correlation analysis is used to calculate the degree of correlation between two variables, whereas the p -value is used to evaluate statistical significance. If the p -value is greater than 0.05 , the difference between the two variables is deemed non-significant [25]. Depending on the p -value, the existing hypothesis of no correlation between two variables can be rejected; however, the p -value does not indicate the correlation between the variables. Therefore, to understand the correlation between variables, a correlation coefficient must be obtained in addition to the p -value.

Pearson's correlation coefficients between the environmental radiation dose rate, proportion of the residents over the age of 40, and the relative cancer incidence data from 2009 to 2016 in the Republic of Korea were obtained to determine the relationship between these variables. The raw data for cancer incidence and age distribution were obtained for each basic administrative district in Korea, whereas the raw data for environmental radiation dose rates were obtained exclusively for the area where the monitoring post was located. Therefore, the correlation coefficients between the variables were derived only for the areas where all the data were available. The null hypothesis (H_0) for this analysis was that there is no positive or negative correlation between the cancer rate and one of the two other variables in each administrative district in the Republic of Korea. The alternative hypothesis (H_1) was that there is a correlation between the cancer rate and one of the two other variables.

3. Results and discussion

The statistics for environmental radiation dose rate, relative cancer incidence, and proportion of residents over the age of 40 for each year are summarized in Table 1. Furthermore, these data are presented in the map in Fig. 2.

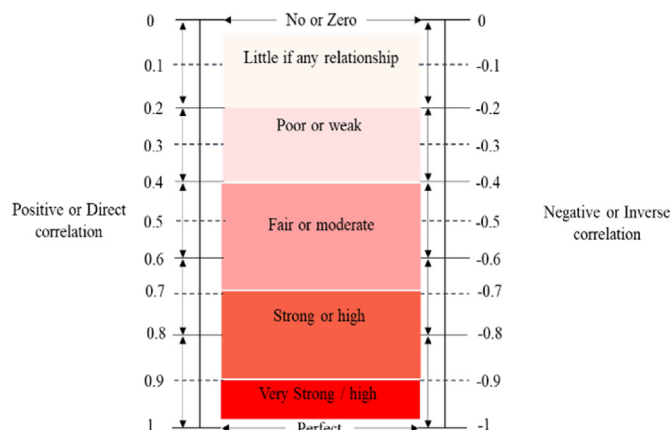


Fig. 1. Interpretation of Pearson's correlation coefficients [23,24].

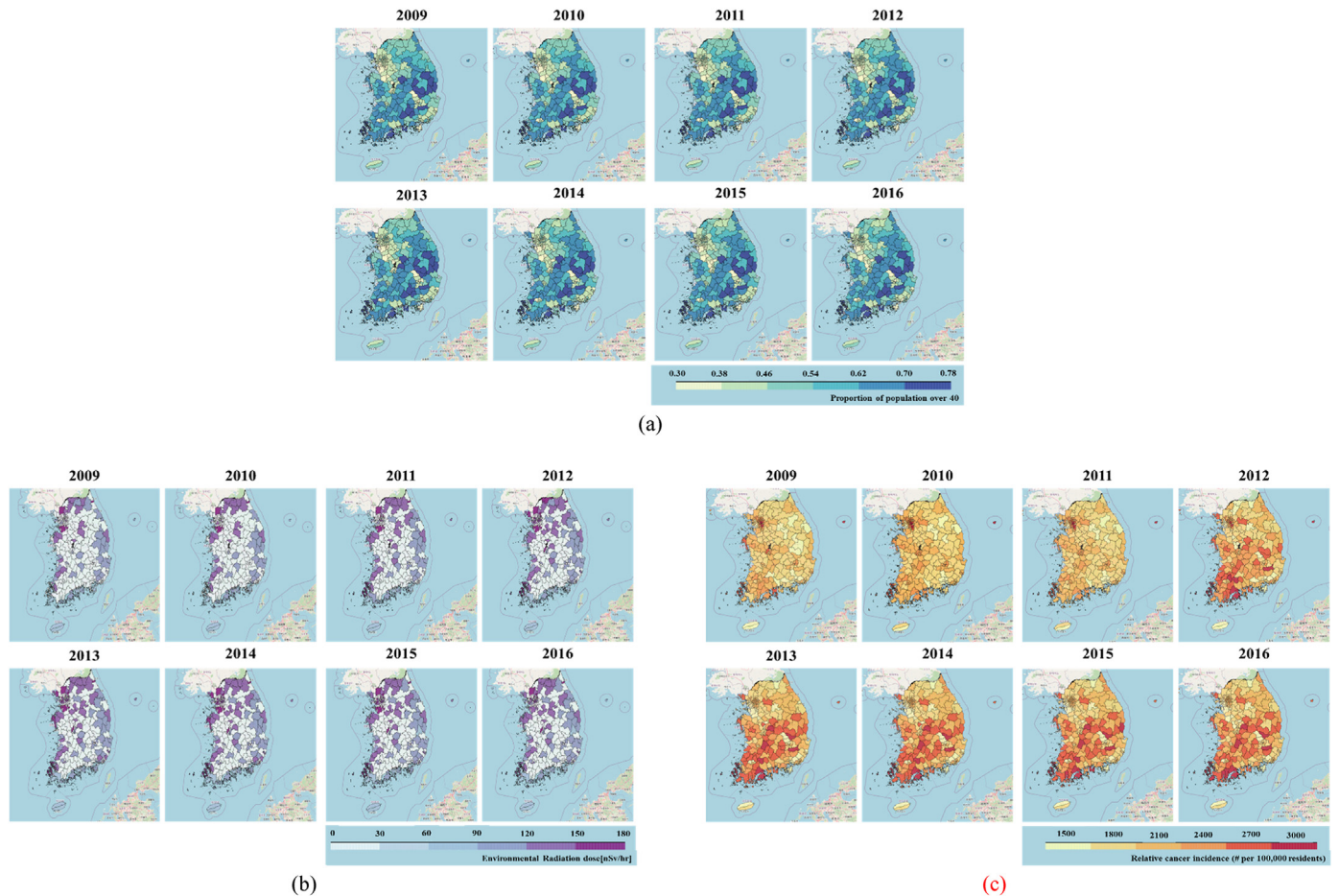


Fig. 2. Annual averages of (a) proportion of the population over 40, (b) environmental radiation dose rate, and (c) relative cancer incidence from 2009 to 2016.

The average environmental radiation dose rates in the monitored areas were between 118.5 and 122.2 nSv/h, which is approximately 1 mSv if these values were converted into an annual unit. This dose constitutes approximately 1/3 of the annual exposure of Korean citizens to natural radiation. Generally, environmental radiation dose rates fluctuate within the normal range except for temporary rises caused by an accident or radiation work and weather changes such as snow, thunderstorms, and rainfall.

The difference in the environmental radiation dose rates at different areas could be due to variations in radon concentrations, which in turn are caused by bedrock composition, cosmic radiation, or altitude. For instance, the environmental radiation rates in Paju, Daejeon, and Sokcho, where bedrock is mostly composed of granite, are higher than those of the other regions in the Republic of Korea.

Regarding environmental radiation doses, Sokcho, Paju, Hwa-seong, Dangjin, and Yeongtong exhibited the highest

Table 1

Statistics for environmental radiation dose rate (nSv/h), relative cancer incidence (# per 100,000 residents), and proportion of the population over 40 from 2008 to 2016.

Year	2009			2010			2011			2012		
	Radiation Dose Rate	Relative Cancer Incidence	Proportion of population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of population over 40
Mean	119.3	1561.3	51.0	117.7	1703.5	52.5	118.8	1994.8	53.8	122.5	2322.8	53.4
Standard variation	21.8	228.3	8.1	20.7	239.3	8.2	22.4	277.2	8.1	22.6	358.3	7.9
Min.	78.8	1251.6	34.1	77.4	1332.0	35.4	77.1	1559.0	36.9	77.1	1682.0	37.8
Max.	189.4	2449.9	66.7	181.1	2654.0	68.1	181.1	2929.0	69.3	183.3	3244.0	70.2
Year	2013			2014			2015			2016		
	Radiation Dose Rate	Relative Cancer Incidence	Proportion of Population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of Population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of Population over 40	Radiation Dose Rate	Relative Cancer Incidence	Proportion of Population over 40
Mean	121.6	2426.2	54.3	122.2	2415.2	55.2	122.2	2520.3	56.0	122.0	2661.8	57.0
Standard variation	21.4	430.4	7.7	21.8	478.1	7.8	21.9	490.3	7.8	21.9	516.2	8.0
Min.	76.6	1678.0	38.6	77.3	1579.0	39.3	77.1	1721.0	40.2	76.8	1762.0	41.2
Max.	179.9	3748.0	71.3	177.0	3738.0	72.3	181.4	3729.0	72.9	186.2	3967.0	73.9

Table 2
Pearson's coefficients between relative cancer incidence and other variables.

Year	Variable	Pearson correlation coefficient	p-value	Degree of correlation between the two variables
2009	Environmental radiation dose rate	-0.016	0.903	Little if any relationship
	Proportion of the population over 40	0.146	0.263	Little if any relationship
2010	Environmental radiation dose rate	0.094	0.466	Little if any relationship
	Proportion of the population over 40	0.127	0.327	Poor or weak
2011	Environmental radiation dose rate	0.074	0.571	Little if any relationship
	Proportion of the population over 40	0.448	2.348×10^{-4}	Fair or moderate
2012	Environmental radiation dose rate	-0.045	0.654	Little if any relationship
	Proportion of the population over 40	0.600	2.085×10^{-11}	Fair or moderate
2013	Environmental radiation dose rate	-0.139	0.154	Little if any relationship
	Proportion of the population over 40	0.755	5.915×10^{-21}	Strong or high
2014	Environmental radiation dose rate	-0.189	0.050	Little if any relationship
	Proportion of the population over 40	0.851	4.206×10^{-31}	Strong or high
2015	Environmental radiation dose rate	-0.217	0.025	Poor or weak
	Proportion of the population over 40	0.860	2.196×10^{-32}	Strong or high
2016	Environmental radiation dose rate	-0.265	0.005	Poor or weak
	Proportion of the population over 40	0.853	9.356×10^{-32}	Strong or high

Table 3
Pearson's coefficients between relative cancer incidence and other variables for the districts with the highest radiation dose rates.

Name of district	Variable	Pearson correlation coefficient	p-value	Degree of correlation between the two variables
Gangwon	Environmental radiation dose rate	-0.355	0.389	Poor or weak
Sokcho	Proportion of the population over 40	0.981	1.780×10^{-5}	Very strong or high
Gyeonggi	Environmental radiation dose rate	0.316	0.018	Poor or weak
	Proportion of the population over 40	0.795	0.446	Strong or high
Gyeonggi Hwaseong	Environmental radiation dose rate	-0.194	0.645	Little if any relationship
	Proportion of the population over 40	0.518	0.189	Fair or moderate
Chungnam	Environmental radiation dose rate	-0.746	0.034	Strong or high
Dangjin	Proportion of the population over 40	0.798	0.018	Strong or high
Suwon	Environmental radiation dose rate	-0.122	0.773	Little if any relationship
Yeongtong	Proportion of the population over 40	0.083	0.846	Little if any relationship

environmental radiation dose rates among all of the regions studied herein. In contrast, Ongjin, Seogwipo, Jindo, Jeju, and Geoje exhibited the lowest radiation levels. Using the methods described above, the five areas with the highest/lowest proportion of the population over 40 were identified. The five regions with the highest proportions were Goheung, Cheongsong, Bonghwa, Yeongdeok, and Namhae-gun, whereas the five regions with the lowest proportion were Yeongtong, Hwaseong, Geoje, Siheung, and Gimhae. Among the five regions with the lowest proportion, Geoje is a low-radiation region, whereas Hwaseong and Yeongtong are high-radiation regions. Correlation analyses were performed for each year and region.

Table 2 summarizes the correlation analysis results for the two comparison cases by year: (1) between relative cancer incidence and environmental radiation dose rate; (2) between relative cancer incidence and proportion of the residents over the age of 40. In 2009 and 2010, neither variable was confirmed to be correlated

with the relative cancer incidence. However, the proportion of residents over the age of 40 was correlated with the relative cancer incidence from 2011 onward, and the degree of correlation has been gradually increasing. Unlike the proportion of the residents over the age of 40, environmental radiation dose rates and relative cancer incidence did not appear to be correlated in most of the years analyzed herein.

The Pearson correlation results for the five highest environmental radiation dose rate regions are summarized in Table 3. Interestingly, there were many regions with a p-value of 0.05 or higher, and therefore the existing hypothesis could not be rejected. In the case of the regions where some null hypotheses could be rejected and alternative hypotheses could be selected, the correlation coefficients were negative for environmental radiation dose rate and positive for the proportion of the population over 40.

Table 4 shows the correlation analysis results for the lowest environmental radiation dose rate districts. Regarding the

Table 4
Pearson's coefficients between relative cancer incidence and other variables for the districts with the lowest radiation dose rates.

Name of district	Variable	Pearson correlation coefficient	p-value	Degree of correlation between the two variables
Incheon	Environmental radiation dose rate	-0.806	0.016	Strong or high
Ongjin	Proportion of the population over 40	0.902	0.002	Very strong or high
Jeju	Environmental radiation dose rate	-0.057	0.893	Little any if relationship
Seogwipo	Proportion of the population over 40	0.997	6.897×10^{-8}	Very strong or high
Jeonnam	Environmental radiation dose rate	0.437	0.279	Fair or moderate
Jindo	Proportion of the population over 40	0.978	2.556×10^{-5}	Very Strong or high
Jeju	Environmental radiation dose rate	0.219	0.603	Poor or weak
Jeju	Proportion of the population over 40	0.998	2.467×10^{-8}	Very strong or high
Gyeongnam Geoje	Environmental radiation dose rate	0.682	0.063	Fair or moderate
	Proportion of the population over 40	0.816	0.014	Strong or high

Table 5
Pearson's coefficients between relative cancer incidence and other variables for the districts with the highest proportion of the population over 40.

Name of district	Variable	Pearson correlation coefficient	p-value	Degree of correlation between the two variables
Jeonnam	Environmental radiation dose rate	-0.626	0.096	Fair or moderate
Goheung	Proportion of the population over 40	0.952	2.611×10^{-4}	Very strong or high
Gyeongbuk Cheongsong	Environmental radiation dose rate	-0.875	4.386×10^{-3}	Strong or high
	Proportion of the population over 40	0.988	4.149×10^{-6}	Very strong or high
Gyeongbuk Bonghwa	Environmental radiation dose rate	-0.899	2.408×10^{-3}	Strong or high
	Proportion of the population over 40	0.980	1.836×10^{-5}	Very strong or high
Gyeongbuk Yeongdeok	Environmental radiation dose rate	0.756	0.029	Strong or high
	Proportion of the population over 40	0.978	2.470×10^{-5}	Very strong or high
Gyeongnam Namhae	Environmental radiation dose rate	-0.805	0.016	Strong or high
	Proportion of the population over 40	0.984	1.089×10^{-5}	Very strong or high

Table 6
Pearson's coefficients between relative cancer incidence and other variables for the districts with the lowest proportion of the population over 40.

Name of district	Variable	Pearson correlation coefficient	p-value	Degree of correlation between the two variables
Suwon	Environmental radiation dose rate	-0.122	0.773	Little any if relationship
Yeongtong	Proportion of the population over 40	0.083	0.846	Little any if relationship
Gyeonggi Hwaseong	Environmental radiation dose rate	-0.194	0.645	Little any if relationship
	Proportion of the population over 40	0.518	0.189	Fair or moderate
Gyeongnam	Environmental radiation dose rate	0.682	0.063	Fair or moderate
	Proportion of the population over 40	0.816	0.014	Strong or high
Geoje	Environmental radiation dose rate	-0.780	0.353	Strong or high
Gyeonggi	Environmental radiation dose rate	-0.780	0.353	Strong or high
Siheung	Proportion of the population over 40	0.936	6.285×10^{-4}	Very strong or high
Gyeongnam Gimhae	Environmental radiation dose rate	-0.722	0.043	Strong or high
	Proportion of the population over 40	0.960	1.571×10^{-4}	Very strong or high

proportion of the population over 40, all of the five regions had p -values ≤ 0.05 , which allowed us to reject the existing null hypothesis and adopt an alternative hypothesis. The correlation coefficient indicating the degree of correlation was also 0.8 or higher in all the five regions. However, the correlation between the environmental radiation dose rate and the relative cancer incidence indicated that the existing null hypothesis could not be rejected in any of the regions except for Ongjin (i.e., the alternative hypothesis could only be applied to Ongjin).

For the districts with the highest proportion of the population over 40, the existing null hypothesis could be rejected in all the regions except for Goheung. For the environmental radiation dose rate, the three other cities except for Yeongdeok had negative correlations. The results are summarized in Tables 5 and 6.

Our analyses indicated that there was no correlation between the environmental radiation dose rate and the relative cancer incidence in the regions with the lowest proportion of the population over 40. The only exception was Gimhae, which had a negative correlation. For the proportion of the population over 40, our findings confirmed that three regions had a positive correlation, with Yeongtong and Hwaseong being the exceptions.

Fig. 3 illustrates the relationship between the environmental radiation dose rate and the relative cancer incidence, whereas Fig. 4 shows the relationship between the proportion of the population over 40 and the relative cancer incidence.

Finally, the population of the region with the largest or lowest number of cancer patients as of 2016 (i.e., the most recent data available) was examined. Table 7 shows the 10 regions with the

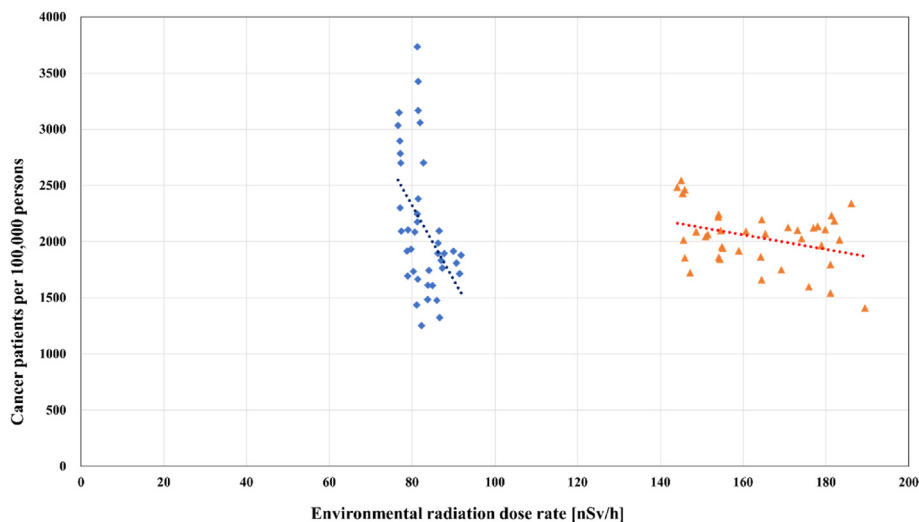


Fig. 3. Relative cancer incidences versus the environmental radiation dose rate for the districts with the five highest and the five lowest environmental radiation dose rates.

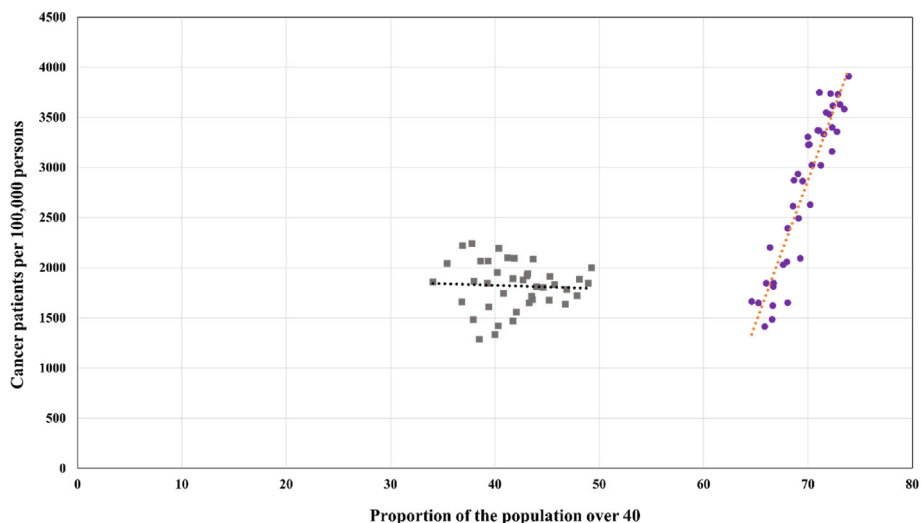


Fig. 4. Relative cancer incidences versus the proportion of the population over 40 for the districts with the five highest and the five lowest environmental radiation dose rates.

Table 7
Population in areas with the highest relative cancer incidence in 2016.

Name of district	Radiation Dose rate	Relative Cancer Incidence	Proportion of males over 40	Proportion of females over 40	Average Age
Jeonnam	134.58	3967	69.2	74.4	52.0
Shinan					
Jeonnam	104.06	3911	71.1	76.5	53.8
Goheung					
Jeonnam	81.28	3735	65.2	71.0	50.1
Jindo					
Chungnam Boeun	131.7	3723	67.8	72.7	50.9
Gyeongnam Namhae	112.61	3630	68.9	76.9	53
Gyeongbuk Yeongdeok	109.94	3616	69.0	75.5	52.3
Gyeongbuk	88.57	3582	70.5	76.4	52.7
Cheongsong					
Gyeongbuk	121.87	3570	65.9	69.3	48.2
Ulleung					
Jeonnam	99.23	3535	63.9	69.2	49.2
Wando					
Jeonnam	139.75	3527	66.2	71.1	50.2
Muju					

highest cancer incidence and Table 8 shows the 10 regions with the lowest cancer incidence. The average age in the areas with the highest relative cancer incidences was approximately 50 years or

older, whereas the areas where the average age was 40 years or younger had fewer cancer patients.

Table 8
Population in areas with the lowest relative cancer incidence in 2016.

Name of district	Radiation Dose rate	Relative Cancer Incidence	Proportion of males over 40	Proportion of females over 40	Average Age
Gangwon	130.8	2074	48.2	57.4	42.0
Yanggu					
Changwon	91.83	2066	47.1	50.5	40.0
Jinhae					
Incheon	162.49	2036	49.5	52.3	39.4
Gyeyang					
Cheongju	126.34	2033	43.7	47.6	38.9
Cheongwon					
Gwangju	107.28	2031	43.1	45.1	35.8
Gwangsan					
Incheon	136.37	2028	47.2	49.6	37.9
Yeonsu					
Gyeongnam Gimhae	119.71	2000	47.9	50.6	38.0
Gyeongnam Geoje	89.99	1914	44.3	46.3	36.8
Gyeonggi	113.34	1843	48.5	49.5	37.9
Siheung					
Gyeonggi	145.85	1762	49.6	52.6	40.0
Bucheon					

4. Conclusions

Our study sought to determine which of the two variables, environmental radiation dose rate or age distribution, was more statistically correlated with the cancer incidence rate in each administrative district in the Republic of Korea. To achieve this, data for environmental radiation dose rates, the proportion of residents over the age of 40, and the number of cancer patients for each administrative district in Korea from 2009 to 2016 were obtained from websites operated by Korean government organizations. The data were preprocessed and analyzed using the R language to investigate the relationships between the relative cancer incidence and the environmental radiation dose rate, as well as between the relative cancer incidence and the proportion of the population over 40. Our findings indicated that the environmental radiation dose rate had no statistically significant relationship or a negative correlation with the relative cancer incidence. However, the proportion of the population over 40 had a positive correlation with the relative cancer incidence. Therefore, age has a greater effect on cancer incidence than environmental radiation dose rate according to the data analyzed in our study. To examine the effects of age distribution in more detail, a follow-up study would be done to analyze the effect of age distribution by gender on the cancer incidence rate.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korean government (Ministry of Science and ICT; No. 2020M2D2A2062436) and by the Nuclear Safety Research Program through the Korea Foundation of Nuclear Safety (KoFONS) through financial resources granted by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea (No. 2003019).

References

- [1] UN Scientific Committee on the Effects of Atomic Radiation (UNSCEAR), Sources and Effects of Ionizing Radiation, 2000. New York.
- [2] Korea Institute of Nuclear Safety (KINS), Radiation Environment in Korea, vol. 48, Korea Institute of Nuclear Safety, 2016.
- [3] National Research Council, Health Risks from Exposure to Low Levels of Ionizing Radiation BEIR VII Phase 2, National Academies Press, Washington, 2006.
- [4] T.E. Johnson, Ch. 7 Biological basis for radiation safety, in: Introduction to Health Physics, fifth ed., McGraw-Hill Education, New York, 2017, pp. 285–337.
- [5] W.-L. Hsu, D.L. Preston, M. Soda, H. Sugiyama, S. Funamoto, K. Kodama, A. Kimura, N. Kamada, H. Dohy, M. Tomonaga, M. Iwanaga, Y. Miyazaki, H.M. Cullings, A. Suyama, K. Ozasa, R.E. Shore, K. Mabuchi, The incidence of leukemia, lymphoma and multiple myeloma among atomic bomb survivors: 1950–2001, *Radiat. Res.* 179 (2013) 361–382, <https://doi.org/10.1667/rr2892.1>.
- [6] D.L. Preston, E. Ron, S. Tokuoka, S. Funamoto, N. Nishi, M. Soda, K. Mabuchi, K. Kodama, Solid cancer incidence in atomic bomb survivors: 1958–1998, *Radiat. Res.* 168 (2007) 1–64, <https://doi.org/10.1667/rr0763.1>.
- [7] J.K. Lee, Ch. 9 radiation measurement, in: Principles of Radiation Protection vol. 1, Korean Association for Radiation Application, Seoul, 2016, pp. 381–450.
- [8] M. Tubiana, L.E. Feinendegen, C. Yang, J.M. Kaminski, The linear no-threshold relationship is inconsistent with radiation biologic and experimental data, *Radiology* 251 (2009) 13–22, <https://doi.org/10.1148/radiol.2511080671>.
- [9] F.R. Tang, K. Loganovsky, Low dose or low dose rate ionizing radiation-induced health effect in the human, *J. Environ. Radioact.* 192 (2018) 32–47, <https://doi.org/10.1016/j.jenvrad.2018.05.018>.
- [10] John D. Boice Jr., Jay H. Lubin, Occupational and environmental radiation and cancer, *Cancer. Cause. Control.* 8 (1997) 309–322.
- [11] Joo, Use of big data analysis to investigate the relationship between natural radiation dose rates and cancer incidences in Republic of Korea, *Nucl. Eng. Technol.* 52 (8) (2020) 1798–1806, <https://doi.org/10.1016/j.net.2020.01.015>.
- [12] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 40, KINS/ER-028, 2009.
- [13] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 41, KINS/ER-028, 2010.
- [14] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 42, KINS/ER-028, 2011.
- [15] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 44, KINS/ER-028, 2012.
- [16] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 45, KINS/ER-028, 2013.
- [17] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 46, KINS/ER-028, 2014.
- [18] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 47, KINS/ER-028, 2015.
- [19] Korea Institute of Nuclear Safety, Environmental Radioactivity Survey Data in Korea, vol. 48, KINS/ER-028, 2016.
- [20] MOIS, <https://jumin.mois.go.kr/>.
- [21] Nettleton David, Ch 6: selection of variables and factor derivation. *Commercial Data Mining: Processing, Analysis and Modeling for Predictive Analytics Projects*, Morgan Kaufmann, 2014, pp. 79–104.
- [22] Ch. 9 G.S. Kim, Correlation analysis, in: R Big Data Analysis for Value Creation, Hannarae, Seoul, 2017, pp. 217–240.
- [23] Ch. 7 M.K. Seo, Statistical analysis, in: R for Practical Data Analysis, Gilbut, Seoul, 2014, pp. 342–350.
- [24] Ch. 9 G.S. Kim, Correlation analysis, in: R Big Data Analysis for Value Creation, Hannarae, Seoul, 2017, pp. 217–240.
- [25] Ch. 8 J.H. Kim, Correlation analysis, in: Environmental Statistics & Data Analysis, Hannarae, Seoul, 2018, pp. 180–202.