

# Regional Variation In $PM_{2.5}$ And Lung Cancer Incidence Across South Korea, 2015–2022

Sanghyo Elaine Jeong\*

1. Yongsan International School of Seoul, 23 Sinbanpo-ro 23-gil, Seocho-gu, Seoul, 06509, Republic of Korea

\*elainejeong1004@gmail.com

## ABSTRACT

Air pollution, or particularly fine particulate matter ( $PM_{2.5}$ ), has been linked to numerous adverse health outcomes, including lung cancer. South Korea has  $PM_{2.5}$  levels that greatly surpass World Health Organization guidelines, making it an important context for studying this relationship. However, few studies have examined this association at a nationwide or regional level while simultaneously accounting for smoking prevalence and population age structure. This study investigated whether regional variation in average  $PM_{2.5}$  concentrations from 2015-2022 was associated with lung cancer incidence across South Korea's 17 administrative divisions, adjusting for smoking prevalence and the percentage of the population aged 65 and older. Six linear regression models were used to evaluate these relationships, using regional  $PM_{2.5}$  concentrations, 2022 lung cancer incidence rates, smoking prevalence, and age structure data obtained from national databases. Results showed no statistically significant association between average  $PM_{2.5}$  concentrations and lung cancer incidence in any model ( $p > 0.05$ ), while smoking prevalence was a significant predictor ( $\beta = 1.104$ ,  $p = 0.039$ ), explaining 30.9% of regional variance. The final model, which included  $PM_{2.5}$ , smoking prevalence, and age structure, explained 40.3% of the variance in lung cancer incidence. These findings suggest that behavioral factors may play a larger role than regional  $PM_{2.5}$  differences in explaining lung cancer incidence across South Korea, though limited regional variation in  $PM_{2.5}$  and a small sample size may have constrained statistical power.

## KEYWORDS

$PM_{2.5}$ ; Air Pollution; Lung Cancer; Lung Cancer Incidence; South Korea; Smoking Prevalence; Environmental Health

## INTRODUCTION

Air pollution is a major environmental health issue that affects human health globally (1). According to the World Health Organization (WHO), more than 99% of the population lives in regions where the air pollution surpasses the WHO air quality limits, and approximately 4.2 million premature deaths are attributed to ambient air pollution each year (2). Particulate matter (PM) are particles that are inhalable (3). Within the realm of PM,  $PM_{2.5}$  and  $PM_{10}$  are the two most prevalent in studies regarding health. Fine particulate matter, or  $PM_{2.5}$  (particles that have a diameter of smaller than or equal to  $2.5\mu m$ ) are especially harmful because of their small size, allowing them to penetrate deep into the lungs and even enter the bloodstream (2). Studies have shown that long-term exposure to  $PM_{2.5}$  can cause problems ranging from atherosclerosis, respiratory and cardiovascular ill health, and possibly neurological development and diabetes (4). Acting both as a carrier of specific carcinogens and a carcinogen itself,

PM<sub>2.5</sub> is a cumulative result of vehicle emissions, industrial activities, power generation, and chemical reactions in the atmosphere (2). Long term exposure was linked to adverse perinatal outcomes and lung cancer, and it was determined as a cause of lung cancer in 2013 by WHO's International Agency for Research on Cancer (4) This study is especially relevant in the context of South Korea, as the PM<sub>2.5</sub> level there is 20.4µg per cubic meter, which significantly exceeds the WHO's recommended level of 5µg, by 4 times (5).

The relationship between PM<sub>2.5</sub> and lung cancer has been studied extensively in Korea primarily at the individual level. Despite growing evidence linking PM<sub>2.5</sub> exposure to lung cancer, few studies have conducted a nationwide ecological comparison across all 17 administrative divisions of South Korea while simultaneously incorporating regional PM<sub>2.5</sub> concentrations, lung cancer incidence, smoking prevalence, and population age structure within a single analytical framework. For example, a nationwide retrospective cohort analysis conducted by the Korea National Institutes of Health evaluated the associations between exposure to PM<sub>2.5</sub> and lung cancer risk, but used different variables like age and sex. They found that O<sub>3</sub> concentrations within were highly correlative with lung cancer incidence, while independent association with PM<sub>2.5</sub> was more modest (6). Similarly, another Korean study reported a significant increase in lung cancer risk with higher PM<sub>2.5</sub> exposure levels, with odds ratios ranging from 1.41 to 1.68, especially for individuals with obstructive lung function impairment, which means that the airways inside the lungs are either narrowed or blocked (7). These studies have provided varying findings, suggesting that additional research is needed to better understand the relationship between regional PM<sub>2.5</sub> exposure and lung cancer incidence.

This study investigates whether regional variation in average PM<sub>2.5</sub> concentrations is associated with regional lung cancer incidence across South Korea's 17 administrative divisions after accounting for smoking prevalence and age structure. We hypothesize that administrative divisions with higher average PM<sub>2.5</sub> concentrations will exhibit higher lung cancer incidence rates, even after accounting for differences in smoking prevalence and age structure. These findings aim to improve understanding of how environmental exposure may influence regional health disparities. Understanding how air pollution varies geographically may help identify populations at greater environmental health risk. The findings may also contribute to future public health policies focused on reducing exposure to harmful air pollutants.

## METHODS AND MATERIALS

This study was an ecological observational study examining the association between ambient fine particulate matter (PM<sub>2.5</sub>) concentrations and lung cancer incidence across the 17 administrative regions of South Korea. Regional PM<sub>2.5</sub> concentrations from 2015-2022 were analyzed in relation to 2022 regional lung cancer incidence rates, while adjusting for two potential confounding variables: smoking prevalence and population age structure.

This included the following 17 administrative regions, which are government areas used to organize the country. Its top-level divisions include one special city, six metropolitan cities, one self-governing city, eight provinces, and three special self-governing provinces (8): Seoul Metropolitan City, Busan Metropolitan City, Daegu Metropolitan City, Incheon Metropolitan City, Gwangju Metropolitan City, Daejeon Metropolitan City, Ulsan Metropolitan City, Sejong Special Self Governing City, Gyeonggi do,

Gangwon do, Chungcheongbuk-do, Chungcheongnamdo, Jeollabuk-do, Jeollanam-do, Gyeongsangbuk-do, Gyeongsangnam-do, Jeju Special Self Governing Province.

Data for this study were collected from four publicly accessible national databases. Annual average PM<sub>2.5</sub> concentrations from 2015 - 2022 were obtained from the National Institute of Environmental Research (NIER) air quality database (9). Population age structure data were obtained from the City Population database using the 2020 Census, from which the percentage of the population aged 65 years and older was calculated for each of the 17 administrative regions (10). Regional lung cancer incidence data were obtained from the Korea National Cancer Incidence Database (KNCIDB), which reports regional cancer statistics including the number of cases (N), incidence rates, 95% confidence intervals, and long-term incidence trends measured by annual percent change (11). Because annual regional incidence data were not available for every year, this study used the 2022 regional lung cancer incidence rate for each administrative region as the dependent variable. Finally, regional smoking prevalence data for 2025 were obtained from the Korea Disease Control and Prevention Agency (KDCA) Community Health Survey (12).

To estimate long-term PM<sub>2.5</sub> exposure, the average PM<sub>2.5</sub> concentration for each administrative region was calculated by averaging the annual PM<sub>2.5</sub> concentrations from 2015 through 2022. Population age structure was represented by percentage of residents aged 65 years and older, which was calculated as:

$$\text{Percent population aged 65+} = (\text{population aged 65+} / \text{total population}) \times 100 \text{ (Eq. 1)}$$

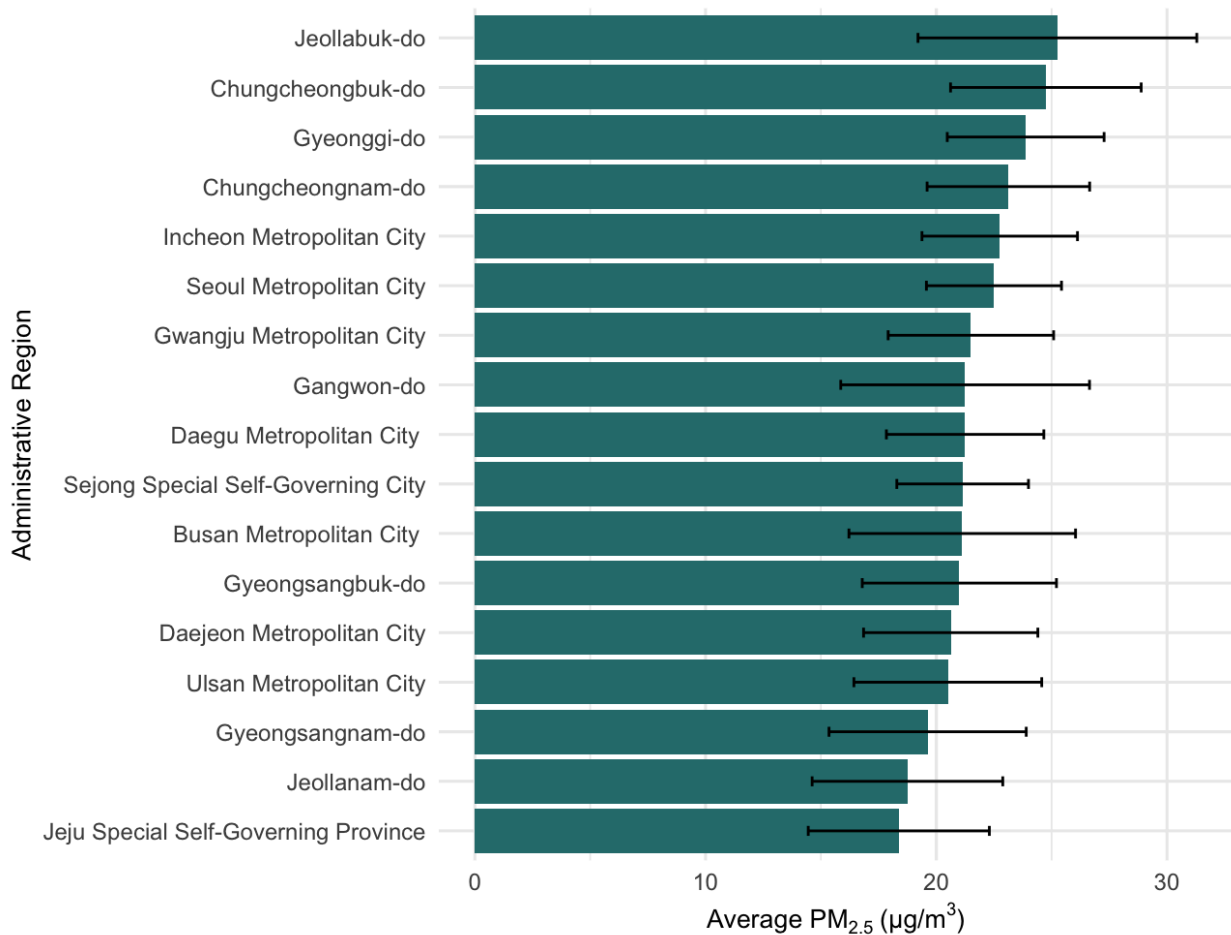
The data were analyzed using R, where multiple datasets were combined to account for potential confounding variables. Analytical graphs were initially created to organize the results. Linear regression was chosen for this study because the research question required examining the relationship between PM<sub>2.5</sub> concentration and lung cancer incidence rate while adjusting for smoking prevalence and age structure at the same time. Since the dependent variable, lung cancer incidence rate, is a continuous value rather than a binary outcome, linear regression is the appropriate method to model it. Regression also allows the effect of PM<sub>2.5</sub> on lung cancer incidence to be isolated from the effects of smoking prevalence and age structure, since these variables could otherwise influence the relationship being studied. All statistical analyses were conducted using R, which is commonly used in public health and environmental research for this type of regional-level regression analysis. Six linear regression models were fitted to evaluate the association between PM<sub>2.5</sub> concentration and regional lung cancer incidence. The models examined: 1) PM<sub>2.5</sub> concentration in 2022, 2) PM<sub>2.5</sub> concentration in 2015, 3) Average PM<sub>2.5</sub> concentration from 2015-2022, 4) Average PM<sub>2.5</sub> concentration adjusted for smoking prevalence, 5) Average PM<sub>2.5</sub> concentration adjusted aged 65 years and older, and 6) Average PM<sub>2.5</sub> concentration for both smoking prevalence and population age structure. Regression coefficients ( $\beta$ ), p values, and coefficients of determination ( $R^2$ ) were reported for each model. Statistical significance was defined as  $p < 0.05$ . The final multiple linear regression model was specified as:

$$\text{Lung Cancer Incidence Rate}_i = \beta_0 + \beta_1(\text{Average PM}_{2.5,i}) + \beta_2(\text{Smoking Prevalence}) + \beta_3(\% \text{Population Aged 65+}_i) + \epsilon_i \text{ (Eq. 2)}$$

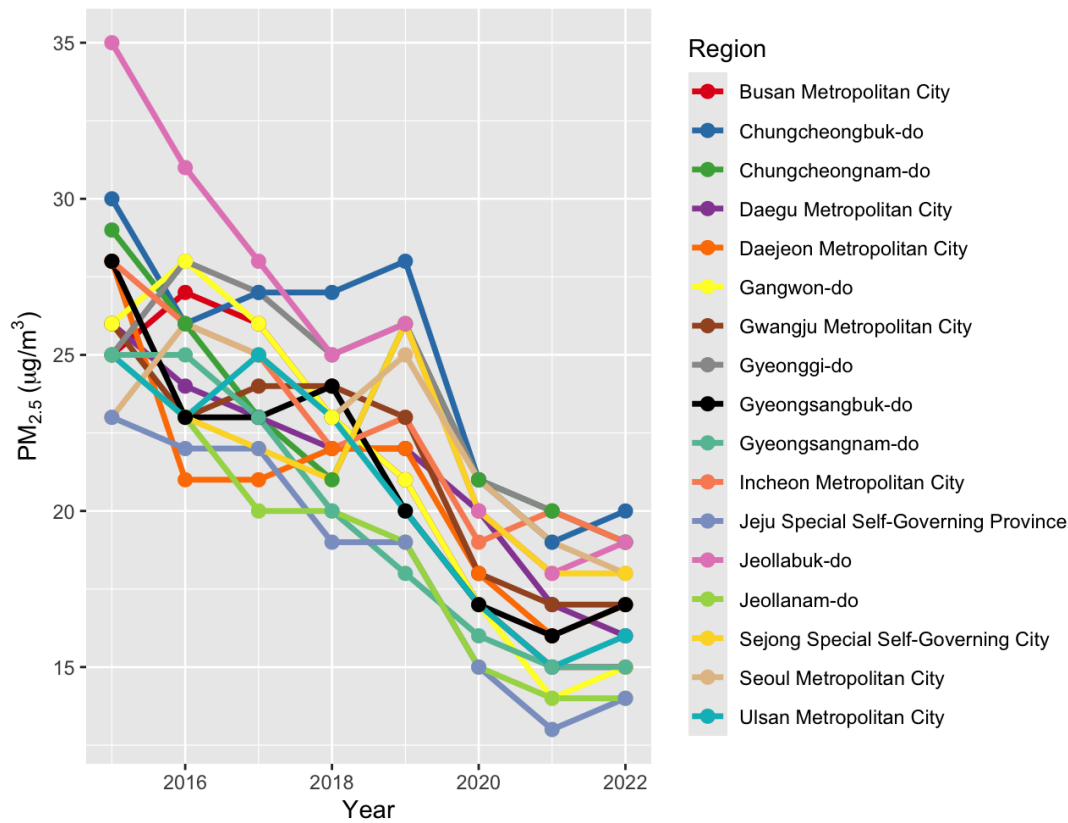
Lung Cancer Incidence Rate<sub>i</sub>: Lung cancer incidence rate for region *i* (dependent variable);  $\beta_0$ : intercept, or the predicted lung cancer incidence rate when all predictors equal zero;  $\beta_1$ : regression coefficient for average PM<sub>2.5</sub> concentration from 2015–2022;  $\beta_2$ : regression coefficient for smoking prevalence;  $\beta_3$ : regression coefficient for the percentage of the population aged 65 years and older;  $\varepsilon_i$ : random error, representing variation in lung cancer incidence that is not explained by the model; *i*: one of South Korea's 17 administrative divisions.

## RESULTS

Average PM<sub>2.5</sub> concentrations varied across South Korea's administrative regions between 2015 and 2022 (Figure 1). The highest was Jeollabuk do (35) and the lowest was the Jeju self governing province (11) (Figure 2). The figure displays a significant decrease in the average PM<sub>2.5</sub> levels for most of the 17 administrative regions in South Korea. In some regions, PM<sub>2.5</sub> concentrations reached their highest levels in 2019 before declining in subsequent years.



**Figure 1.** Average PM<sub>2.5</sub> concentration (2015–2022) by administrative region, ranked in descending order. Error bars represent the standard deviation in PM<sub>2.5</sub> concentrations across the study period. Jeollabuk-do had the highest average PM<sub>2.5</sub> concentration, while Jeju Special Self-Governing Province had the lowest.



**Figure 2.** Annual average PM<sub>2.5</sub> concentrations (µg/m<sup>3</sup>) from 2015 to 2022 across South Korea's 17 administrative divisions. Most regions show a general downward trend in PM<sub>2.5</sub> concentrations over the study period, with several regions peaking around 2015 or 2019 before declining in subsequent years.

Linear regression analyses were performed to examine whether regional PM<sub>2.5</sub> concentrations were associated with lung cancer incidence. Models were first evaluated without adjustment for potential confounding variables and then were repeated after adjusting for smoking prevalence and age structure

To examine the association between PM<sub>2.5</sub> levels and lung cancer incidence, six regression models were created. Linear regression analysis was done via six models to examine the link between the level of PM<sub>2.5</sub> and lung cancer incidence rate in South Korea's 17 administrative divisions (Table 1).

**Table 1.** Summary of six linear regression models examining the association between  $PM_{2.5}$  concentration and regional lung cancer incidence across South Korea's 17 administrative divisions, with and without adjustment for smoking prevalence and population age structure (% aged 65+).

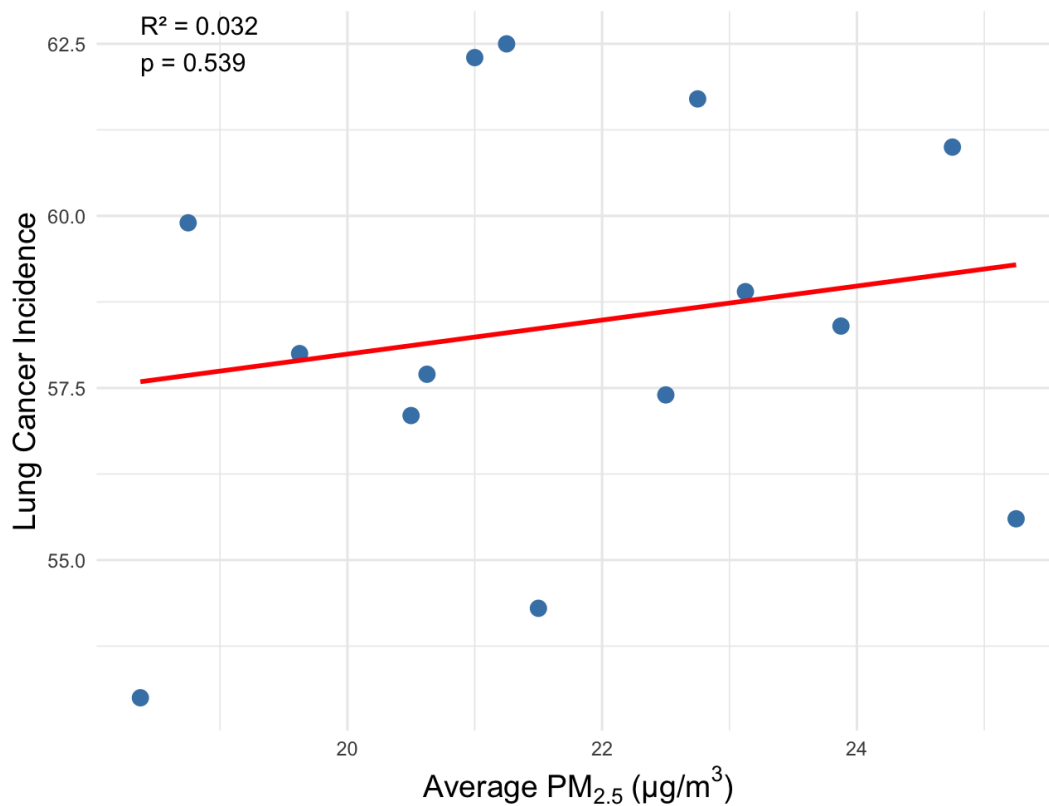
| Model | $PM_{2.5}$ Variable            | Covariates           | $\beta$ ( $PM_{2.5}$ ) | P-value ( $PM_{2.5}$ ) | $R^2$ |
|-------|--------------------------------|----------------------|------------------------|------------------------|-------|
| 1     | Smoking only                   | Smoking              |                        | 0.039                  | 0.309 |
| 2     | 2022 $PM_{2.5}$                | None                 | 0.272                  | 0.517                  | 0.036 |
| 3     | 2015 $PM_{2.5}$                | None                 | 0.146                  | 0.587                  | 0.025 |
| 4     | Average $PM_{2.5}$ (2015-2022) | None                 | 0.247                  | 0.539                  | 0.032 |
| 5     | Average $PM_{2.5}$ (2015-2022) | Smoking              | 0.319                  | 0.357                  | 0.363 |
| 6     | Average $PM_{2.5}$ (2015-2022) | Smoking + % aged 65+ | 0.325                  | 0.359                  | 0.403 |

$\beta$  = regression coefficient (slope).

$R^2$  = proportion of variation in lung cancer incidence explained by the model. (A higher number means that it is a closer fit)

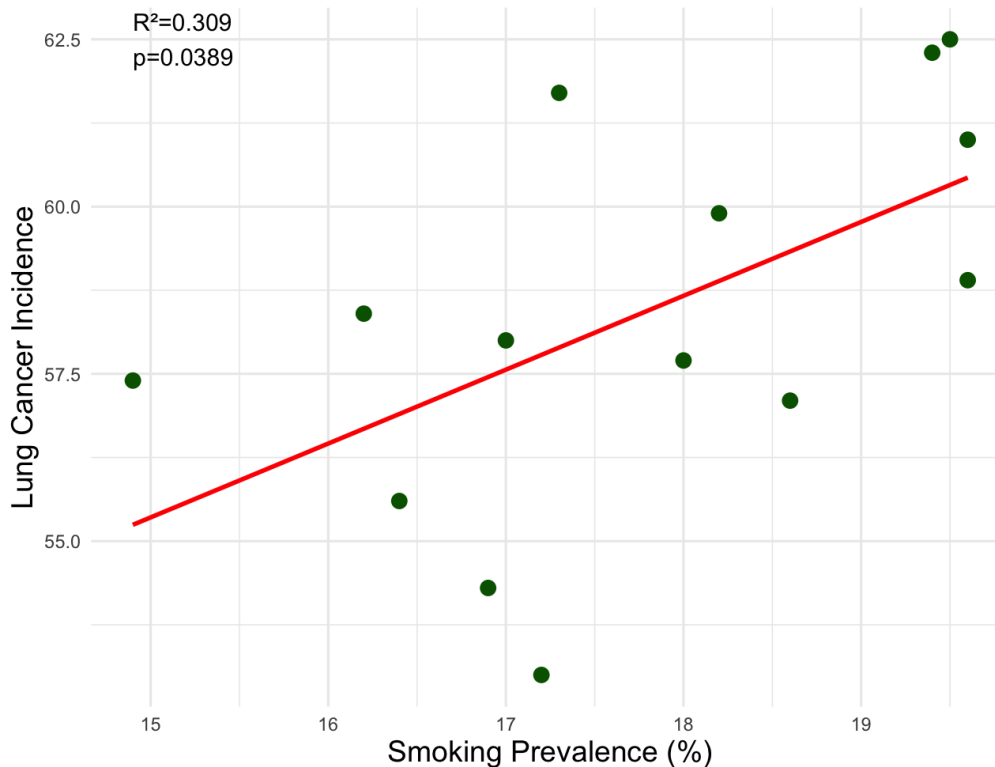
$p < 0.05$  indicates statistical significance.

Considering  $PM_{2.5}$  concentration independent of smoking prevalence, there was no statistically significant association with lung cancer incidence regardless of  $PM_{2.5}$  representation (Figure 3). The models that included  $PM_{2.5}$  2022 ( $\beta = 0.272$ ,  $p = 0.517$ ,  $R^2 = 0.036$ ),  $PM_{2.5}$  2015 ( $\beta = 0.146$ ,  $p = 0.587$ ,  $R^2 = 0.025$ ), and average  $PM_{2.5}$  2015-2022 ( $\beta = 0.247$ ,  $p = 0.539$ ,  $R^2 = 0.032$ ) all showed positive regression coefficients but were not statistically significant ( $p > 0.05$ ).



**Figure 3.** Scatterplot of average PM<sub>2.5</sub> concentrations (2015-2022) and 2022 lung cancer incidence rate across South Korea's 17 administrative divisions. The red line represents the linear regression fit ( $\beta = 0.247$ ,  $p = 0.539$ ,  $R^2 = 0.032$ ), indicating no statistically significant association between average PM<sub>2.5</sub> concentration and lung cancer incidence.

There is a statistically significant effect of smoking prevalence on the incidence rate of lung cancer ( $\beta = 1.104$ ,  $p = 0.039$ ), which explains 30.9% of differences between the regions ( $R^2 = 0.309$ ) (Figure 4).



**Figure 4.** Scatterplot of smoking prevalence (%) and 2022 lung cancer incidence rate across South Korea's 17 administrative divisions. The red line represents the linear regression fit ( $\beta = 1.104$ ,  $R^2 = 0.309$ ,  $p = 0.039$ ), indicating a statistically significant positive association between smoking prevalence and lung cancer incidence.

Adding smoking prevalence to the average  $PM_{2.5}$  concentration, its positive association with lung cancer incidence rate was shown again, but it was not statistically significant ( $\beta = 0.319$ ,  $p = 0.357$ ). Smoking prevalence showed itself as a statistically significant predictor ( $p = 0.036$ ), and the adjusted  $R^2$  improved to 0.363.

The final regression model included average  $PM_{2.5}$  concentration, smoking prevalence, and the percentage of people who are over 65 years old in the region's population. This model explains more variance in lung cancer incidence than any previous ones ( $R^2 = 0.403$ ), but the individual predictor variables are still not statistically significant. Average  $PM_{2.5}$  concentration is associated with an increase in lung cancer incidence positively ( $\beta = 0.325$ ,  $p = 0.359$ ), while smoking prevalence ( $p = 0.085$ ) and age structure ( $p = 0.432$ ) are not statistically significant.

Thus, no statistically significant associations between regional  $PM_{2.5}$  concentration and lung cancer incidence were found in any of the regression models. But including smoking prevalence and age structure increased the predictive ability of the models, as the final model accounted for approximately 40.3% of the variance in regional lung cancer incidence.

## DISCUSSION

This study found that there is no statistically significant relationship between the mean PM<sub>2.5</sub> concentration level and the lung cancer cases per region in South Korea's 17 administrative regions. This is because even though there are positive regression coefficients in all the regression models, none of the PM<sub>2.5</sub> levels were statistically significant ( $p > 0.05$ ). Therefore, the hypothesis that regions with higher PM<sub>2.5</sub> concentrations would have higher lung cancer incidence was not supported.

One possible explanation for this result is the small variation in PM<sub>2.5</sub> concentrations among the 17 administrative regions in South Korea. According to Figure 2, the average level of PM<sub>2.5</sub> varied by only 18 to 25  $\mu\text{g}/\text{m}^3$  for all years in question (2015–2022). Given that most regions experienced a relatively equal level of air pollution, there may not be enough variation in exposure in order to reveal statistically significant effects on the number of lung cancer incidences. In other words, when the concentration of PM<sub>2.5</sub> is relatively stable in different regions, it may be hard to establish a connection with the risk of developing a disease.

This study initially focused on hexavalent chromium (Cr(VI)) and its potential relationship with health outcomes. However, the limited availability of reliable and regionally comparable data made it difficult to conduct a nationwide analysis. The research, therefore, shifted its focus to PM<sub>2.5</sub>, for which more comprehensive regional and historical data were available. This allowed for a consistent comparison across all 17 administrative divisions and enabled the analysis of regional air pollution and lung cancer incidence.

Figure 1 shows that the level of PM<sub>2.5</sub> across the years has a significant downward trend. The overall decline in PM<sub>2.5</sub> concentrations across South Korea may be partly attributed to the implementation of increasingly stringent air-quality policies during the study period. South Korea introduced measures to reduce fine dust beginning in 2014, followed by the Comprehensive Action Plan on Fine Dust in 2017, which aimed to reduce PM<sub>2.5</sub> emissions by 30% from 2014 levels by 2022. Seasonal Particulate Matter Management was also introduced in 2019 to further reduce pollution during periods of high PM<sub>2.5</sub> concentrations. These policies included reductions in major emission sources such as coal use, diesel vehicles, and industrial activities, which may have contributed to the decreasing PM<sub>2.5</sub> concentrations observed across the 17 administrative regions (13).

This study has several limitations. First, the sample size was small because the analysis included only 17 administrative divisions in South Korea, which limits statistical power and makes it more difficult to detect relatively small associations. Second, PM<sub>2.5</sub> concentrations showed relatively limited variation across regions, which may have reduced the ability of the analysis to identify a clear relationship with lung cancer incidence. This limited variation may partly explain why the association between average PM<sub>2.5</sub> and lung cancer incidence was not statistically significant. Another important limitation is the potential lag between PM<sub>2.5</sub> exposure and lung cancer development. Lung cancer can develop over many years, while this study compared regional PM<sub>2.5</sub> concentrations from 2015–2022 with lung cancer incidence data from 2022. Therefore, current lung cancer incidence may reflect exposure to air pollution from earlier periods rather than only the PM<sub>2.5</sub> levels measured during the study period. Future studies could incorporate historical PM<sub>2.5</sub> exposure over a longer period to better account for this potential lag. Finally, this study used an ecological design, meaning that the analysis was conducted using regional averages rather than

individual-level exposure and health data. Therefore, the observed regional relationships cannot be interpreted as evidence that individual exposure to  $PM_{2.5}$  causes lung cancer. Future research using longer-term exposure data and individual-level information could provide a more precise assessment of the relationship between  $PM_{2.5}$  and lung cancer risk.

## CONCLUSION

This study found no statistically significant association between regional  $PM_{2.5}$  concentrations and lung cancer incidence across South Korea's 17 administrative divisions, although smoking prevalence showed a stronger relationship with lung cancer incidence. These findings highlight the importance of considering long-term air pollution exposure and multiple demographic and behavioral factors when investigating regional differences in lung cancer.

## ACKNOWLEDGMENTS

I would like to thank my research mentor, Kate, for her guidance and feedback throughout the development of this study. Next, I would like to thank Jacqueline for her support throughout the publishing process. I would lastly like to thank Polygence for providing the mentorship program and resources that made this research possible.

## CONFLICT OF INTEREST

The author declares no conflicts of interest related to this work.

## REFERENCES

1. Wang M, Kim RY, Kohonen-Corish MRJ, Chen H, Donovan C, Oliver BG. Particulate matter air pollution as a cause of lung cancer: epidemiological and experimental evidence. *British Journal of Cancer* [Internet]. 2025 Apr 4;132(11):986–96. Available from: <https://www.nature.com/articles/s41416-025-02999-2>
2. Types of pollutants [Internet]. Available from: <https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants>
3. Particulate Matter (PM) Basics | US EPA [Internet]. US EPA. 2026. Available from: <https://www.epa.gov/pm-pollution/particulate-matter-pm-basics>
4. World Health Organization: WHO. Outdoor air pollution a leading environmental cause of cancer deaths [Internet]. 2013. Available from: <https://www.who.int/europe/news/item/17-10-2013-outdoor-air-pollution-a-leading-environmental-cause-of-cancer-deaths>
5. Moon S. South Korea's annual PM2.5 levels four times higher than WHO guideline. *DongA Science* [Internet]. 2023 Aug 30; Available from: <https://dongascience.com/en/news/61411>
6. Kim NE, Lee JY, Lee JY, Lee GY, Lee CY, Park S, et al. Long-term exposure to fine particulate matter, ozone, and greenness and the risk of lung cancer: a retrospective cohort analysis within a national sample cohort. *Frontiers in Public Health* [Internet]. 2025 Sep 30;13:1661937. Available from: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12518234/>
7. Lee J, Choi JY, Lim JU. Gradual increase in lung cancer risk due to particulate matter exposure in patients with pulmonary function impairments: A Nationwide Korean Database analysis. *Journal of Korean Medical Science* [Internet]. 2025 Jan 1;40(42):e270. Available from: <https://doi.org/10.3346/jkms.2025.40.e270>
8. Introduction to Korean regions | Life in Korea | Study in Korea | Run by Korean Government [Internet]. Available from: <https://www.studyinkorea.go.kr/ko/life/regionsIntro.do>
9. Environmental Satellite Center [Internet]. Environmental Satellite Center. Available from: <https://nesc.nier.go.kr/en/html/datasvc/data.do?pageIndex=1&outputInnb=153&atrb=PM25-SURF ACE>
10. City population: Check for humans [Internet]. Available from: <https://www.citypopulation.de/>
11. Cancer Statistics > National Cancer Control Programs > NCCI | NATIONAL CANCER CENTER [Internet]. Available from: [https://www.ncc.re.kr/main.ncc?uri=english/sub04\\_Statistics](https://www.ncc.re.kr/main.ncc?uri=english/sub04_Statistics)
12. K2WebWizard [Internet]. Available from: <https://www.kdca.go.kr/bbs/kdca/42/215513/artclView.do>
13. Ryoo I, Ren L, Li G, Zhou T, Wang M, Yang X, et al. Effects of seasonal management programs on PM2.5 in Seoul and Beijing using DN-PMF: Collaborative efforts from the Korea-China joint research. *Environment International* [Internet]. 2024 Aug 22;191:108970. Available from: <https://www.sciencedirect.com/science/article/pii/S0160412024005567>