Short-term exposures to PM$_{10}$ and PM$_{2.5}$ and lung cancer hospitalization risk in Shenzhen, China: A Double Negative Controls Study

Zhao Yang (✉ yangz98@connect.hku.hk)  
The University of Hong Kong

He Liang  
National Cancer Center, National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College

Wei Deng  
Peking University Cancer Hospital & Institute

Yong Ji  
National Cancer Center, National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College

Article

Keywords: double negative control, lung cancer hospitalization, particulate matter, unmeasured confounding

Posted Date: June 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1660474/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Few studies assess the short-term exposure to particulate matter (PM), including inhalable particles with an aerodynamic diameter of 10 F or less ($\text{PM}_{10}$) and fine PM with an aerodynamic diameter of 2.5 e or less ($\text{PM}_{2.5}$), on lung cancer hospitalization risk, especially in developing countries. This study examines whether short-term exposures to $\text{PM}_{10}$ and $\text{PM}_{2.5}$ affect lung cancer hospitalization risk in Shenzhen, China. We evaluated short-term PM-lung cancer hospitalization associations using cancer hospital-based data during 2018-2019 in Shenzhen, China. Daily data on air pollutants and weather conditions were collected from publicly available online resources. We estimated the corresponding associations based on a double negative control design after adjusting for unmeasured confounding. Two-pollutant models were fitted to examine the robustness of associations. On average, per 10 w increase in weekly mean $\text{PM}_{10}$ concentration was associated with an 0.25 (95% confidence interval [CI] 0.04-0.45) higher lung cancer hospitalization risk. The corresponding increase in lung cancer hospitalization risk for the same increase in $\text{PM}_{2.5}$ was 0.36 (95% CI 0.13-0.61). These detrimental effects remained after adjusting for sulfur dioxide and carbon monoxide and became stronger among males and those aged <75 years. No delayed effects were observed. Significant unmeasured confounding effects were also observed.

Why Was This Study Done?

- Long-term exposure to particulate matter (PM) is associated with lung cancer risk. Still, few studies have shown the short-term PM-lung cancer risk associations, especially in developing countries (e.g., China).

- Previous studies using either time-series, case-crossover, or cohort studies are vulnerable to unmeasured confounding. Causal models (e.g., propensity score method/inverse probability weighting and the difference-indifference approach) derived the average causal effects using observational data at the cost of several stringent assumptions, which are often violated in practice, inducing biased estimates.

- Few studies investigate the unmeasured confounding when assessing the short-term PM-lung cancer risk associations, particularly in observational studies.

What did the researchers do and find?

- We examined whether short-term exposures to $\text{PM}_{10}$ and $\text{PM}_{2.5}$ affect lung cancer hospitalization risk using a double negative control design based on the hospital-based data in Shenzhen, China, between January 1, 2018, and December 31, 2019.

- We found that per 10 increase in weekly mean $\text{PM}_{10}$ concentration was associated with an 0.25 higher lung cancer hospitalization risk, and the corresponding risk for the same increase in $\text{PM}_{2.5}$ was 0.36.

- PM-lung cancer hospitalization associations derived from observational studies are vulnerable to unmeasured confounding.
What do these findings mean?

- Our findings suggest that short-term exposures to PM$_{10}$ and PM$_{2.5}$ appear to increase lung cancer hospitalization risk.
- The results derived from observational studies (e.g., time-series, case-crossover, and cohort studies) should be interpreted cautiously, primary due to unmeasured confounding.

Introduction

Particular matter (PM, including inhalable particles with an aerodynamic diameter of 10 $\mu$m or less [PM$_{10}$] and fine particle with an aerodynamic diameter of 2.5 $\mu$m or less [PM$_{2.5}$])-lung cancer risk associations, especially long-term effects, have been well characterized in developed countries over the past decades [1–13]. However, differences in chemical components and the subsequent toxic characteristics of particular matters between developed and developing countries limit the application of these findings in developing countries [14], especially in China. As the largest and fastest developing country, China has experienced severe air pollution due to rapid industrialization, urbanization, and transportation development in recent years [15–18], so it is urgently needed to assess the health effects of those ambient air pollution [19].

A small but growing body of literature using either case-crossover [20], time-series [21–23], or cohort [24–30] studies have examined PM-lung cancer risk associations in China, yielding inconsistent results, particularly for the long-term effects. Seldom do the short-term PM-lung cancer risk associations were assessed [20–23]. Furthermore, most evidence investigating PM-lung cancer risk associations are susceptible to bias from either unmeasured confounders (e.g., health-seeking behaviour and treatment strategies) or sample selecting, which are the major concerns in observational studies. Causal model approaches (e.g., inverse-probability weighting [31, 32], regression discontinuity design [33], and difference-in-difference method [34]), mimicking randomized controlled trials, have been proposed to estimate unconfounded estimates at the cost of several stringent assumptions. However, most of these assumptions are not verifiable and often violate in practice, inducing biased estimates. For instance, no unmeasured confounding assumption required for the inverse-probability weighting or propensity score method is not empirically verifiable. Violation of the parallel assumption (which refers to no variable changing differentially over time across space other than the exposure of interest) for the difference-in-difference method often happens in practice, particularly when the outcome of interest is ill-defined [31, 32, 34].

The double negative control design [35], exploiting a confounding bridge function identified by a negative control exposure to link the potential outcome and a negative control outcome distribution, provides unconfounded exposure-outcome estimates. Briefly, the negative control outcome associated with confounders but not causally affected by the exposure of interest, mimicking the unobserved potential outcome via a confounding bridge approach, unifies previous bias adjustment methods in negative control designs and diminishes the likelihood of unmeasured confounding [36–38]. Additionally, the
double negative control design provides a powerful toolbox to assess the magnitude of the unmeasured confounding effects using the association between the negative control exposure and the negative control outcome.

This study examined whether short-term exposures to PM$_{10}$ and PM$_{2.5}$ have independent effects on lung cancer hospitalization risk based on the hospital-based lung cancer cases ($n=1,511$) in an economically advanced and less polluted immigrant city of Shenzhen in China. Herein, we focused on lung cancer hospitalization risk because short-term exposure to air pollution is more likely to exacerbate lung cancer-related symptoms and increase hospitalization risk [39], rather than cause lung carcinogenesis or cancer directly.

**Methods**

This is a double negative control study [35]. A pair of negative control outcome and exposure variables were employed to estimate the short-term PM-lung cancer hospitalization associations based on the time-series data, as shown in Fig. 1. Briefly, a transformation of the negative control outcome identified by a negative control exposure, referred to as the confounding bridge, is used to capture and diminish unmeasured confounding under several key assumptions. (1) Conditional on a sufficient set of measured and unmeasured confounders, the exposed subjects would experience the same average outcome as the unexposed subjects. This is the fundamental assumption required for inferring causality in observational studies [40]. (2) The negative control outcome is associated with both the measured and unmeasured confounders but not causally affected by the exposure. To this end, we used the number of lung cancer cases in week $t-1$ as the negative control outcome because it could not be affected by the mean PM concentration in week $t$ unless the presence of unmeasured confounding. (3) There are some confounding bridge functions such that the unmeasured confounding effect on the outcome at each exposure level is identical to that on the confounding bridge function. In practice, this can be achieved by prespecifying a proper parametric or non-parametric model [35]. (4) The negative control exposure is independent of both the outcome and the negative control outcome after adjusting for the exposure and the measured confounders. Notably, in time-series studies, the mean PM concentration in week $t+1$ could not affect the number of lung cancer cases in week $t$, and thus would be a well-defined negative control exposure. It has also been used in previous time-series studies to test unmeasured confounding [37, 38]. Then, any non-zero association between the negative control exposure and the outcome or the negative control outcome is completely driven by unmeasured confounding. Furthermore, there are no restrictions on the association of the negative control exposure with the exposure and the association of the negative control outcome with the outcome. **Supplementary Material** provides more details on the double negative control design. The study is in accordance with relevant guidelines and regulations.

We did not have a formal prospective analysis plan because this is a hypothesis-driven study. This study only involved the record data and did not involve human subjects directly. It was exempt from the IRB review in Cancer Hospital Chinese Academy of Medical Sciences, Shenzhen center.
Study Area

Shenzhen, one of the typical immigrant cities of the Pearl River Delta region in China, has been experiencing the fastest urbanization and transportation development for the past four decades, with 13,026.6 thousand permanent population (refers to those who have lived in Shenzhen over half a year, including 4,547.0 thousand registered and 8,479.7 thousand non-registered population) living in an area of ~1,997.5 square kilometers [41]. Typically, Shenzhen has a subtropical maritime climate, and is economically advanced and less polluted.

Data Collection

Cancer-specific case certificates with de-identified personal information were extracted from Cancer Hospital Chinese Academy of Medical Sciences, Shenzhen center between January 1, 2018, and December 31, 2019. Types of cancer are recorded by the International Classification of Diseases 9th Revision (ICD-9) or 10th Revision (ICD-10), with lung cancer coded as 162 in ICD-9 or C34 in ICD-10. In this study, we only focused on those lung cancer cases with the initial record, excluding the recurrent cases mainly due to their treatment strategies and poor prognosis. Of these eligible cases, we further divided weekly aggregated cases into several strata by sex and age (i.e., <65, 65–74, and 75+ years) without sample size calculation.

Daily time-series data on air pollutants, including PM$_{10}$ ($\mu$g/m$^3$, 24-hour average), PM$_{2.5}$ ($\mu$g/m$^3$, 24-hour average), ozone (O$_3$, $\mu$g/m$^3$, maximum 8-hour average), nitrogen dioxide (NO$_2$, $\mu$g/m$^3$, 24-hour average), sulfur dioxide (SO$_2$, $\mu$g/m$^3$, 24-hour average), and carbon monoxide (CO, mg/m$^3$, 24-hour average), were obtained from the National Air Quality Real-Time Publishing Platform (http://106.37.208.233:20035/). This platform is administered by the Chinese Ministry of Environmental Protection and has displayed real-time concentrations of air pollutants from controlled monitoring sites since January 2013. The weekly mean concentration for those air pollutants was simply averaged from all monitoring sites in Shenzhen (Supplementary Material) across a week based on the ISO week date system. To account for the potential effects of weather conditions, including temperature (°C) and relative humidity (%), we also obtained daily mean temperature and mean relative humidity for Shenzhen from the National Climatic Data Center (NCDC available at https://www.ncei.noaa.gov/; Air force station BAOAN [ID 594930]). Moreover, the weekly mean temperature and relative humidity were computed by averaging the daily monitoring data across a week.

Statistical analysis

Under the double negative control analytical framework, we employed a generalized additive linear confounding bridge function, which has been widely used in previous studies [34, 42]. We obtained PM-lung cancer hospitalization associations via a modified two-stage least estimator [35]. In stage 1, we regressed the negative control outcome on the negative control exposure, observed confounders and the
exposure, and obtained the predicted value of the negative control outcome. In stage 2, we regressed the primary outcome on the predicted value of negative control outcome, observed confounders, and the exposure. Then, the coefficient between the exposure and the outcome in stage 2 is the causal estimate of interest. Furthermore, the corresponding standard errors were estimated using the heteroscedasticity and autocorrelation covariance method [35, 43, 44].

We estimated the short-term effects of PM$_{10}$ and PM$_{2.5}$ on lung cancer hospitalization risk with using the square root of the number of lung cancer cases for normalization and variance stabilization [35, 45]. We excluded the missing data for lung cancer cases in week 40 2018 because it involved a seven-day long holiday on National Day of the People's Republic of China. We explored the delayed effects (i.e., the exposure affects the outcome for a lapse of time beyond the event period) and quantified net effects over a predefined lag period. For the main model, we emphasized the estimated associations of PM level in the present week and included a discrete Fourier transform of time to control for the underlying time trends in lung cancer hospitalization risk, an indicator of the month to account for short-term monthly variations, and natural spline functions with nine $df$ for temperature and three $df$ for the relative humidity to control for the potentially non-linear effects of weather conditions. We reported the point estimate and 95% confidence intervals as the change in lung cancer hospitalization risk per 10 $\mu g/m^3$ increase in weekly mean PM$_{10}$ or PM$_{2.5}$ concentration.

We fitted two-pollutant models beyond the main model by adjusting for an additional gaseous pollutant of either O$_3$, NO$_2$, SO$_2$, or CO. The PM$_{10}$- and PM$_{2.5}$-lung cancer hospitalization associations were considered robust if consistent causal estimates were obtained from both single- and two-pollutant models, as determined by a paired z-test [46]. We also carried out a set of sensitivity analyses to explore the delayed effects by examining various lag structures (Supplementary materials). We then conducted stratified analyses according to sex and age group to investigate the possibly modified effects. Finally, we performed the confounding tests and estimated the corresponding unmeasured effects with an additional adjustment for the PM$_{10}$ or PM$_{2.5}$ concentration, time, month indicators, temperature, and relative humidity in the same approach included in the main model.

All statistical analyses were performed using R software, version 3.6.3 (R Foundation for Statistical Computing) via the AER package for the two-stage least estimator and the stats package for the ordinal least square estimator. Sample codes are available from the first author on request. A $P$-value $< 0.05$ was considered statistical significance. We adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guides for reporting observational studies reporting associations of weekly mean PM$_{10}$ and PM$_{2.5}$ concentrations with lung cancer hospitalization risk [47].

Results

Figure 2 shows the weekly mean concentrations of PM$_{10}$, PM$_{2.5}$, O$_3$, NO$_2$, SO$_2$, CO, and the weekly number of lung cancer cases between January 1, 2018, and December 31, 2019. Overall, 1,511 lung cancer cases were identified and included in the final analysis. On average, the weekly mean PM$_{10}$ concentration was
42.4 μg/m³ (median, 40.3; interquartile range [IQR], 28.0 to 53.3), and the corresponding value for PM$_{2.5}$ was 24.7 μg/m³ (median, 24.3; IQR, 16.1 to 31.7). PM$_{10}$ was strongly correlated with PM$_{2.5}$, with the Spearman correlation coefficient of 0.97. The Spearman correlation coefficients between PM$_{10}$ and four gaseous pollutants were 0.51 with O$_3$, 0.74 with NO$_2$, 0.59 with SO$_2$, and 0.57 with CO. The corresponding correlation coefficients between PM$_{2.5}$ and gaseous pollutants were 0.50, 0.73, 0.56, and 0.66.

Supplementary Figure S1 shows a summary description of the weekly mean PM concentration and four gaseous pollutants and weather conditions of temperature and relative humidity.

Table 1 shows changes in lung cancer hospitalization risk with an increase of 10 μg/m³ in weekly mean concentrations of PM$_{10}$ and PM$_{2.5}$ with and without adjusting for gaseous pollutants. On average, per 10 μg/m³ increase in weekly mean PM$_{10}$ concentration was associated with an 0.25 (95% confidence interval [CI], 0.04 to 0.45) higher lung cancer hospitalization risk. The corresponding effect in lung cancer hospitalization risk for the same increase in PM$_{2.5}$ was 0.37 (95% CI, 0.13 to 0.61). The change in lung cancer hospitalization risk per 10 μg/m³ increase in weekly mean PM$_{2.5}$ concentration was larger than in PM$_{10}$. The consistent PM$_{10}$-lung cancer hospitalization associations were obtained in two-pollutant models, especially after adjustment for gaseous pollutants of O$_3$ and NO$_2$. No heterogeneous estimates were observed for the changes in lung cancer hospitalization risk per 10 μg/m³ increase in weekly mean PM$_{10}$ concentration. Similar results were also observed for PM$_{2.5}$. Additionally, confounding effects-driven completely by the unmeasured confounders exists throughout the study (Table 1).

No delayed effects were observed per 10 μg/m³ increase in weekly mean PM$_{10}$ concentration on lung cancer hospitalization risk, as shown in Fig. 3. As expected, modified effects by sex and age for the same increase in PM$_{10}$ concentration on lung cancer hospitalization risk were observed (Fig. 4). On average, per 10 μg/m³ increase in weekly mean PM$_{10}$ concentration might cause an 0.18 higher lung cancer hospitalization risk among Chinese males with a range of 0.11–0.15 higher risk among those aged < 65 years, although some of these effects did not reach the significant level of 0.05. Similar results were also observed for the same increase in PM$_{2.5}$, as depicted in Figs. 3&5. Moreover, significant unmeasured confounding effects exist throughout all sensitivity analyses (Figs. 3–5).

Discussion

Main findings

On average, we observed overall short-term PM-lung cancer hospitalization associations of 0.25 (95% CI 0.04 to 0.46) per 10 μg/m³ increase in weekly mean PM$_{10}$ concentration and 0.36 (95% CI 0.08–0.64) for the same increase in PM$_{2.5}$. The two-pollutant models yielded consistent results for PM$_{10}$ and PM$_{2.5}$, especially after adjusting for SO$_2$ and CO. These results seem reasonable for lung cancer hospitalization risk but not for lung cancer incident risk as lung carcinogenesis is not an accurate adverse event. However, previous epidemiological and experimental studies suggested that ambient fine particular
matter (e.g., PM$_{2.5}$) promotes lung carcinogenesis by causing epigenetic and microenvironmental alternations and triggering the correspondingly systemic inflammation via an increased level of cytokines and oxidative stress [48–51].

Two plausible explanations may partly explain such inconsistent conclusions. One is that lung cancer cases caused by PM$_{10}$ or PM$_{2.5}$ occur among frail participants due to either chronic diseases or some transient condition; their diseases have presumably been advanced to some degree rather than accurate adverse events. This reason is further supported by the relatively delayed PM-lung cancer hospitalization risk associations (Fig. 3). This explanation is also used to illustrate the short-term health effects of exposure to air pollution by the WHO Working Group [52]. Another potential explanation is Berkson's bias due to the selection of study participants [53, 54]. Previous studies showed that people in the high economic cities in China (e.g., Shenzhen) are apt to be sensitive to air pollution and tend to be more conscious in seeking health services [55–58]. In such a case, participants with asymptomatic lung cancer may diagnose by accident due to health-seeking behaviours for other respiratory diseases caused by PM$_{10}$ or PM$_{2.5}$. From this point of view, such health-seeking behaviours may further induce spurious PM-lung cancer hospitalization risk associations, especially among those hospital-based lung cancer patients.

Our results also showed suggestive evidence for sex-modified effects of PM on lung cancer hospitalization risk. In particular, we found significant PM-lung cancer hospitalization associations for Chinese males but not for Chinese females (Figs. 4 and 5). These results were consistent with the previous study [23], in which per 10 $\mu g/m^3$ increase in current PM$_{2.5}$ was positively associated with an 0.68 (95% CI 0.27 to 1.09) higher lung cancer mortality risk among males but was not associated with lung cancer mortality among females. Similar results were also found for PM$_{10}$. One potential explanation is the differences in smoking prevalence and air pollutant exposure as well as physiological lung architecture between males and females. For example, the 2010 national smoking survey results in China showed that smoking prevalence in Chinese females was only 3.4%, much lower than that in Chinese males with 62.4% [59]. A previous cohort study in Hong Kong among people aged 65+ years has also reported that the PM-lung cancer associations were greater in smokers than non-smokers [24]. Furthermore, empirical evidence showed that Chinese females tend to undertake more indoor activities and may expose less to the outdoor air pollutant than Chinese males, especially for those used to wearing masks when undertaking outdoor activities [20]. Lastly, previous studies also showed that females have smaller lungs than males throughout life, which may also affect the exposure level of air pollutants [60, 61].

We also noted that the PM-lung cancer hospitalization associations were modified by age for PM$_{10}$ and PM$_{2.5}$, especially those aged 75+ years, as shown in Figs. 4–5. One possible explanation is the competing risk bias, which means that patients who died from other diseases cannot be diagnosed with lung cancer again [62, 63]. For example, the incidence rate of lung cancer in China was relatively low among those aged 40 years or younger, increased dramatically since then, and reached a peak age of
80–84 years in 2017 [64]. However, the leading causes of death ranged from ischemic heart diseases among those aged 20–39 years to stroke among those aged 40–84 years, and finally ischemic heart disease among those aged 85+ years. As such, study participants enrolled in our studies would be healthier than the general population since participants with poor health status may die from ischemic heart diseases or other cardio-pulmonary diseases early.

Furthermore, few studies in the literature investigate the short-term effects of PM on lung cancer hospitalization, incidence, and mortality risk, especially in developing countries. A recent study using a case-crossover design showed that short-term exposure to PM$_{10}$ and PM$_{2.5}$ appeared to increase lung cancer mortality risk (6.5%, 95% CI 1.2 to 12.0% per 10 $\mu$g/m$^3$ increase in current PM$_{2.5}$) in Shenyang, a very typical industrial city in China, and has been experiencing severe air pollution in recent years [20]. Consistent results were also found in three other cities in China (i.e., Beijing, Chongqing, and Guangzhou) from 2013 to 2015, with a pooled estimate of 0.52 (95% CI 0.06 to 0.99) per 10 $\mu$g/m$^3$ increase in current PM$_{2.5}$, but not for PM$_{10}$ [23]. However, another recent time-series study conducted in Hefei, China, showed limited evidence for the short-term effect of air pollution (1.0024, 95% CI 0.9971–1.0078 per 10 $\mu$g/m$^3$ increase in current PM$_{10}$) on lung cancer mortality risk [22]. Such discrepancies may lie in the differences in the exposure definition and the modeling approach that often fails to account for unmeasured confounding. For example, smoking, health-seeking behaviour, and the corresponding treatment strategies were well-acknowledged factors for lung cancer patients, but they were not considered in previous studies [20, 22, 23]. Furthermore, little is known about the associations between air pollution and cancer progression or survival, although a recent study including 350,000+ lung cancer patients in California reported an inverse association between air pollutants (including PM$_{10}$ and PM$_{2.5}$) concentration and cancer survival [65]. The multi-pollutant model, including the positively correlated ambient air pollutants (e.g., O$_3$, NO$_2$, SO$_2$, and CO), incudes collinearity in the regression-type analysis (e.g., distributed lagged non-linear method [42]) and might yield biased estimates [66, 67].

**Strengths And Limitations Of This Study**

Our current study has three notable strengths. First, our findings were estimated based on a pair of negative control outcome and exposure to diminish unmeasured confounding. Thus, they were less likely to be susceptible to unmeasured confounding bias. Second, study participants enrolled in this study with an average age of 54 years, in which selection bias due to those who have died before recruitment would be minimal. Third, we explored and adjusted for unmeasured confounding, which was extensively discussed in the limitation section in previous studies but was not investigated comprehensively.

Nevertheless, limitations remain. First, study participants were hospital-based lung cancer cases rather than cancer registration patients. Thus, a large number of lung cancer cases might not be included in this study, which potentially leads to underestimated PM-lung cancer hospitalization associations. Including those lung cancer cases would provide more accurate and reliable estimates. Furthermore, our study was likely to include more health-conscious patients, yielding underestimated causal estimates. Second, the
exposure level of air pollutants was proxied by the average concentration of monitoring results across various monitoring stations, which may result in inevitable measurement errors and bias the causal estimates towards the null [68, 69]. Moreover, air pollutant exposures were not assessed at the individual level, which may also lead to aggregation bias, even if aggregated measures were reasonable estimates of the population average of personal exposure values [70]. Third, it is clear that the causal estimates of air pollutants are sensitive to the method of modeling time, weather conditions, and modeling assumptions [71]. However, the double negative control design is strikingly robust to the misspecification of the confounding bridge function, primarily for the well-defined negative control exposure (i.e., mean PM concentration in week \(t + 1\)) and outcome (i.e., the number of lung cancer cases in week \(t - 1\)) [35].

Fourth, we did not assess the concentration-response relationships using localized average causal effects between air pollutants and lung cancer hospitalization risk primarily due to the limited study period. Further analyses with a long study period may perform when data are available. Fifth, as an immigrant city of Shenzhen, \(~74\%\) of lung cancer patients included in this study were non-registered inhabitants, so we could not examine the long-term PM-lung cancer hospitalization associations. Sixth, our findings should be interpreted carefully with limited generalizability because we only assessed the short-term PM-lung cancer hospitalization associations in a specific location of Shenzhen in China.

Conclusions

Our study provides suggestive evidence that short-term exposures to PM\(_{10}\) and PM\(_{2.5}\) might cause increased lung cancer hospitalization risk in China and indicates which population subgroups are more susceptible to air pollutants. Further investigations are needed to evaluate the concentration-response associations of short-term exposures to PM\(_{10}\) and PM\(_{2.5}\) with lung cancer risk, especially in different regions in China.

Declarations

Funding: This research was funded by Shenzhen High-level Hospital Construction Fund.

Acknowledgements: We acknowledge all patients included in this study.


Institutional Review Board Statement: This study only involved the record data and did not involve human subjects directly. It was exempt from the IRB review in Cancer Hospital Chinese Academy of Medical Sciences, Shenzhen center.

Data availability state: Daily time-series data on air pollutants were obtained from the National Air Quality Real-Time Publishing Platform (http://106.37.208.233:20035/). Daily mean temperature and mean
Relative humidity for Shenzhen were obtained from the National Climatic Data Center (NCDC available at [https://www.ncei.noaa.gov/](https://www.ncei.noaa.gov/); Air force station BAOAN [ID 594930]). All data relevant to the outcome (i.e., lung cancer cases) are not publicly available; we would welcome collaborations and research proposals (Dr. Yang, E-mail: yangz98@connect.hku.hk).

**Informed Consent Statement:** Not applicable.

**Conflict of interest:** All authors report no conflict of interest.

**References**


63. Yang, Z.; Schooling, C.M.; Kwok, M.K. Credible Mendelian randomization studies in the presence of selection bias using control exposures. Front Genet 2021, 12, 729326,
Table 1. Changes in lung cancer hospitalization risk in association with an increase of 10 in weekly mean concentrations of PM$_{10}$ and PM$_{2.5}$ as well as unmeasured confounding effects with and without adjustment for gaseous pollutants of ozone (O$_3$), nitrogen dioxide (NO$_2$), sulfur dioxide (SO$_2$), and carbon monoxide (CO)
<table>
<thead>
<tr>
<th>Models</th>
<th>PM_{10}</th>
<th>P-value for difference*</th>
<th>Unmeasured confounding effects</th>
<th>PM_{2.5}</th>
<th>P-value for difference*</th>
<th>Unmeasured confounding effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single-pollutant model of PM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Causal estimate</td>
<td>0.25</td>
<td>(0.04-0.45)</td>
<td>Ref.</td>
<td>0.20</td>
<td>(0.06-0.33)</td>
<td>0.37</td>
</tr>
<tr>
<td>Unmeasured confounding effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Causal estimate</td>
<td>0.37</td>
<td>(0.13-0.61)</td>
<td>Ref.</td>
<td>0.31</td>
<td>(0.09-0.53)</td>
<td></td>
</tr>
<tr>
<td>Two-pollutant model of PM with adjustment for</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O_{3}</td>
<td>0.18</td>
<td>(-0.14-0.49)</td>
<td>0.703</td>
<td>0.20</td>
<td>(0.06-0.34)</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO_{2}</td>
<td>0.03</td>
<td>(-0.27-0.34)</td>
<td>0.249</td>
<td>0.15</td>
<td>(0.05-0.26)</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SO_{2}</td>
<td>0.21</td>
<td>(-0.01-0.43)</td>
<td>0.801</td>
<td>0.17</td>
<td>(0.03-0.31)</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>0.30</td>
<td>(0.05-0.55)</td>
<td>0.773</td>
<td>0.20</td>
<td>(0.08-0.32)</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* P-value for the difference was obtained by using a paired z-test.

**Figures**

![Directed Acyclic Graphs](image)

**Figure 1.** Directed Acyclic Graphs for illustrating the double negative control design using time-series data. \( X \) is the exposure of interest, \( Y \) is the outcome, \( Z \) is the negative control exposure, \( W \) is the negative control outcome, \( C \) is the set of measured confounders, and \( U \) is the set of unmeasured confounders at time \( t \). Here, \( X \) and \( Y \) are respectively the weekly mean concentration of PM_{10} or PM_{2.5} and number of lung cancer cases in week \( t \). \( Z \) is the weekly mean concentration of PM_{10} or PM_{2.5} in week \( t+1 \). \( W \) is the lung cancer cases in week \( t-1 \). As such, the exposure \( X \) in week \( t \) cannot causally affect the negative control outcome \( W \) in week \( t-1 \), and the negative control exposure \( Z \) in week \( t+1 \) cannot causally affect the outcome \( Y \) in week \( t \). Then, any non-zero \( Z-Y \) or \( X-W \) association is solely caused by the unmeasured confounding.
Figure 1

See image above for figure legend.

Figure 2

Weekly air pollutant concentrations for particulate matter (PM$_{10}$ and PM$_{2.5}$), carbon monoxide (CO), nitrogen dioxide (NO$_2$), ozone (O$_3$), and sulfur dioxide (SO$_2$) in Shenzhen, as well as weekly hospital-based lung cancer cases between January 1, 2018, and December 31, 2019.
Figure 3

Delayed effects of short-term exposures to PM\textsubscript{10} and PM\textsubscript{2.5} on lung cancer hospitalization risks and the unconfounding effects. Lag 1, 2, and 4 represent the previous week, the week before the previous week, the three weeks before the previous week concentration of PM\textsubscript{10} and PM\textsubscript{2.5}; and lag 0-1 and lag 0-2 represent the two-week moving average of the present week and the previous week, and the three-week moving average of the present week and the previous two weeks concentration of PM\textsubscript{10} and PM\textsubscript{2.5}. 
Figure 4

The possible modified effects by sex and age on PM$_{10}$-lung cancer hospitalization risk associations
Figure 5
The possible modified effects by age and sex on PM$_{2.5}$-lung cancer hospitalization risk associations

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymaterialsUpdated20220516.docx