Author’s response to reviews

Title: Lung cancer mortality of residents living near petrochemical industrial complexes: a meta-analysis

Authors:
Chengkuan Lin (chl309@mail.harvard.edu)
Huei-Yang Hung (feiyang1080@gmail.com)
David Christiani (dchris@hsph.harvard.edu)
Francesco Forastiere (f.forastiere@deplazio.it)
Ro-Ting Lin (roting@mail.cmu.edu.tw)

Version: 1 Date: 07 Aug 2017

Author’s response to reviews:

Manuscript Number: ENHE-D-17-00161

Manuscript Title: Lung cancer mortality of residents living near petrochemical industrial complexes: a meta-analysis

Reviewer #1

1. The paper provides a valuable, clear, carefully framed and I think a well executed large pooled meta-analysis of the topic that applies what look like appropriate methods and stats tests. It fills a useful gap in the literature and, in the process, raises some important points for discussion in any future studies.

I have a small number of questions that as far as I can see are not addressed or not addressed fully in the text. It would be useful if the authors could either explain why they are not covered or include some brief mention in the text. I also have a small number of suggested minor amendments that I think will strengthen the text. My comments relate primarily to the methods used to underpin the study, the selection of the studies for analysis that could have different
characteristics warranting a slightly fuller discussion, and the limitations of the study which also merit a little expansion.

===>[Response]: We would like to thank you for reviewing our manuscript and providing constructive suggestions, which help to improve the quality of this manuscript. We have carefully revised the manuscript and our responses are provided as below.

2. I assume an expert statistical reviewer has looked at the study. The 12 independently studies populations selected in the meta-analysis 'Areas' selected for the study are tabulated. Some give details on geographical zones and others do not. It would be helpful to have more discussion of this characteristic at the end.

===>[Response]: We’ve consulted a statistical expert, Prof. Hsieh, C. Chung of the Harvard Chan School of Public Health, on the meta-analysis of this study. We’ve already included him in the acknowledgements.

Our analysis has already incorporated the heterogeneity of “area” into the subgroup analysis (see “Location” in Table 2). The definition of petrochemical area has also been summarized in Table 1, column “Definition of petroleum area and reference area.” (new page 30) In response to your suggestion, we agree to address the characteristics of “areas” by two means: First, we summarized the details of industrial activities and substantial chemicals in the newly added column “Industrial activity/ Substantial chemicals” in Table 1. (new page 30) Second, we have revised the following sentence based on this comment in the text: “Most studies defined the exposure based on the geographical locations or distances of residencies from PIC [7, 8, 19, 21, 22], while one study compared the exposed group and reference group by matching job categories in PIC and non-PIC towns [18].” (new page 15, paragraph 3, line 20)

3. I am unclear as to the extent which occupation overlapping with residency might play a part in such analyses in the future and wondered if it was considered in the meta-analysis planning?
We agree that the residential exposure might be overlapped with occupational exposure. Most selected studies have considered this issue either by separating the environmental exposure from the occupational exposure or adjusting for job categories in the analysis. To address this issue, we have added following sentences in Discussion: “Fourth, certain portion of residents living nearby PICs might risk occupational exposure as well. Some studies have separated the environmental exposure from the occupational exposure (study ID=A, B, G, H) or at least adjusted for job categories in the analysis (study ID=M) to reduce the influence of occupational exposure.” (new page 18, paragraph 1, line 5)

4. WHO in 2011 considered 1 in 10 lung cancer cases to be work-related. I am aware that in some of the studies selected, workers may live close to petro-chemical and steel and other industrial complexes - the steel industry being one of a number of complicating factors? - where lung carcinogens are present. Was this also considered in setting up the study and, if so, how? If not, why not?

Thank you for addressing this point. Indeed, a PIC might involve other industrial process, such as steel manufacturing, and thus presented more than one source of pollutants. We extracted information about industrial activity and substantial chemicals in each PIC. To address the heterogeneity, we have summarized this information in Table 1. (new page 30) We agree to add following sentence in the Discussion: “Third, each PIC might involve other manufacturing processes (such as steel, cocking, and power plants) and the exposure level could also be affected by geographical factors across different countries. Limited by the lack of corresponding exposure data, our findings were not able to address the heterogeneity between PICs.” (new page 18, paragraph 1, line 2)

5. Also, were any study areas affected by radon? I believe in Sicily for example, it is present and would be a factor in lung cancer mortality? If so, was this factored into any analysis and does it merit a brief discussion at the end?

Thank you for this comment. We agree that other risk factors should be addressed. This part is now stated in the Introduction and Discussion of the revised manuscript, respectively: “Long-term exposure to the poor air quality, as well as radon, chemicals, and
arsenic compounds among residents living near petrochemical manufacturing complexes raised general awareness and the need to understand the possible adverse health effects among nearby residents [4, 14].” (new page 4, paragraph 2, line 17), “There are some limitations need to be addressed when interpreting our results. First, not all potential confounders were adjusted in the seven articles, such as smoking, radon exposure, meteorological factors, and socioeconomic status.” (new page 17, paragraph 1, line 1), and “The data on radon exposure, as one of the risk factors of lung cancer, were absent in all selected papers. However, there is no evidence of higher radon exposure in PIC areas than non-PIC ones [49, 50].” (new page 17, paragraph 1, line 12)

6. And was there any discussion of plume characteristics and prevailing winds that applied to the selected study areas that might affect exposure?

====> [Response]: Thank you for your remark. All selected papers didn’t consider meteorological factors into their analysis. However, since all included groups were from both upwind and downwind locations, subject-selection bias and meteorological effects due to location variance were less likely. We agree to address this issue in the Discussion of the revised manuscript: “Similarly, seasonal variations of wind directions might either increase or decrease the effect of PIC exposure on residents’ health. Since all studies have exposure and reference groups from both upwind and downwind locations, subject-selection bias and meteorological effects due to location variance were reduced.” (new page 17, paragraph 1, line 14)

7. Again, whilst using ambient air quality indicators on a regional or national basis makes good sense for the meta-analysis, the subject may merit some mention in the discussion which touches on the 'assumption' that 'most factories followed local regulations to some extent'. This statement is not referenced and it would help to provide some evidence about how air pollution regulations compare in Sicily, elsewhere in Italy, Taiwan, USA and UK if available?

====> [Response]: Thank you for addressing this point. We agree that the reason for this assumption should be presented. We have provided some evidences and clarified this issue by adding following sentence: “An effective air quality intervention involved a series of steps, including regulatory establishments, pollution reductions, and anticipated improvements in
health [44]. Although data on ambient pollution monitoring around PICs in the early periods were very limited and hard to obtain, previous studies have documented pollution reductions could be attributable to changing regulations [45, 46].” (new page 16, paragraph 2, line 11)

Methods.

8. Could the authors, for information, identify the other sources they used to identify the relevant studies beyond Pubmed? Pubmed is sometimes probably the least comprehensive source, Web of Science and Science Direct may often much better.

==> [Response]: We appreciate this comment on finding and searching more databases. We have expanded our search databases to PubMed, Cochrane, Web of Science, Science Direct, and others (e.g., Countway Library of Medicine and governmental reports) for an exclusive literature searching. The number of identified articles increased from 365 to 1,249, and one additional paper was included for meta-analysis. Our new results have been revised throughout the manuscript, including Abstract (new page 2, paragraph 3, line 17), Method (from new page 6, paragraph 1, line 3 to new page 7, paragraph 1, line 12), Results (from new page 10, paragraph 2, line 6 to new page 14, paragraph 1, line 2), Discussion (from new page 14, paragraph 3, line 20 to new page 18, paragraph 1, line 9), Figure 1, and all Tables.

9. With regard to use of the Newcastle-Ottawa QA scale for bias in non-RCT studies, it is notable that those who created the scale were relatively recently still looking to test some aspects of its validity. Others who have used the scale have some reservations too about the criteria used. It would therefore be useful to have this aspect briefly mentioned in the paper and referenced.

==> [Response]: Thank you. We agree that Newcastle-Ottawa scale is under developing. However, the reliability of the Scale is quite fair and it has been widely used in meta-analysis. We agree to add the following sentence to address the reason of adopting the Scale: “Although the discussion on the validity of the Scale remained inconclusive, the reliability of the Scale is quite fair and widely used in meta-analysis [32, 33].” (new page 9, paragraph 2, line 14)
Discussion

10. The last point in the methods section could also be picked up in the discussion section with a little expansion of the limitations section and comment on the sub-group analysis which addressed issues surrounding not just location but 'area' including radon and prevailing wind issues(?, along with the possibility of occupation of residents muddying the epi waters (?)

---[Response]: Thank you for addressing this point. In line with previous comments to address the heterogeneity of geographical areas in Comment #2, occupational exposure in Comment #3, and limitations and other confounders in Comments #5&6, we have revised the following sentences based on this comment in the text.

“Most studies defined the exposure based on the geographical locations or distances of residencies from PIC [7, 8, 19, 21, 22], while one study compared the exposed group and reference group by matching job categories in PIC and non-PIC towns [18].” (new page 15, paragraph 3, line 20)

“There are some limitations need to be addressed when interpreting our results. First, not all potential confounders were adjusted in the seven articles, such as smoking, radon exposure, meteorological factors, and socioeconomic status.” (new page 17, paragraph 1, line 1)

“The data on radon exposure, as one of the risk factors of lung cancer, were absent in all selected papers. However, there is no evidence of higher radon exposure in PIC areas than non-PIC ones [49, 50]. Similarly, seasonal variations of wind directions might either increase or decrease the effect of PIC exposure on residents’ health. Since all studies have exposure and reference groups from both upwind and downwind locations, subject-selection bias and meteorological effects due to location variance were reduced.” (new page 17, paragraph 1, line 12)

“Fourth, certain portion of residents living nearby PICs might risk occupational exposure as well. Some studies have separated the environmental exposure from the occupational exposure (study ID=A, B, G, H) or at least adjusted for job categories in the analysis (study ID=M) to reduce the influence of occupational exposure.” (new page 18, paragraph 1, line 5)
Appendices

11. Could these be slightly expanded to provide a little more information along the lines flagged above?

Answer: Yes, we agree to provide more information in Additional file 1 by revising the following paragraph: “b. Selection of non-exposed cohort: Michelozzi’s study reported the SMR, which compared the exposure group to the general population in Rome; so did Tsai’s study in Louisiana and the United States as well as Susana’s study in Wales and England. Instead of comparing to residents in Teesside, Bihopal’s study selected residences in Sunderland as the reference group. Pasetto selected “commuters” as the “non-residential exposed” group. Yang selected another 16 non-petrochemical industrial countries as compared group.” (Additional file 1, new page 1)

12. Minor 'stylistic' points.

p6 line 11 and 12 has an incomplete sentence and an orphan ‘were’.

Answer: Thank you. The ‘were’ has been deleted. (new page 6, paragraph 2, line 16)

p9 line 13 should read data ‘were’ limited and not ‘was’.

Answer: Thank you. The correction is made on new page 9, paragraph 3, line 21.

p16 line 12 should read ‘fewer’ not ‘less’

Answer: Thank you. The correction is made on new page 17, paragraph 1, line 9.
INTRODUCTION

1. Petrochemical plants are often the source of environmental contamination not only of air, but also of other environmental matrices. The authors should add something about the range of pollutants from petrochemical plants that can contribute to lung cancer risk for populations living close to them. Example of a reference with information on pollutants from petrochemical plants:


[Response]: Thank you for this comment. We agree that other source of environmental contamination should be addressed. This part is now stated in the Introduction of the revised manuscript: “Long-term exposure to the poor air quality, as well as radon, chemicals, and arsenic compounds among residents living near petrochemical manufacturing complexes raised general awareness and the need to understand the possible adverse health effects among nearby residents [4, 14].” (new page 4, paragraph 12, line 17) We also summarized the details of industrial activities and substantial chemicals in the newly added column “Industrial activity/ Substantial chemicals” in Table 1 (new page 30).

2. Even though the authors provide a formal meta-analysis of the association between residence near petrochemical industrial complexes and lung cancer mortality, they should report examples of -or review the- evidence from studies on lung cancer incidence that usually are focused on more recent periods.

[Response]: We appreciate this comment on expanding our study period to more recent years. We tried to expand our search databases to PubMed, Cochrane, Web of Science, Science Direct, and others (e.g., Countway Library of Medicine and government reports) for an exclusive literature searching. The number of identified articles increased from 365 to 1,249, and one additional paper was included for meta-analysis. However, for the purpose of the study, our interpretation of this study was focusing on the lung cancer mortality, not incidence. Therefore, one of the exclusion criteria of study selection has been stated in the text: “(3) studies that
reported lung cancer incidence only and lack of mortality data.” (new page 6, paragraph 2, line 16)

3. The authors should add something about the very heterogeneous nature of petrochemical complexes that can have inside refineries, some or several petrochemical plants with different productions, and also power plants.

===>[Response]: Thank you for addressing this point. The nature of petrochemical industrial complexes (PIC) might be very heterogeneous and involves other industrial process, resulting in more than one source of pollutants. We agree to address the heterogeneity of PIC by summarizing information about industrial activity and substantial chemicals in each PIC in Table 1. (new page 30) We also added following sentence in the revision: “These petrochemical manufacturing factories are usually clustered in an industrial area together with other manufacturing processes or industry, such as steel, coking, and thermoelectric plants [7, 8], and called petrochemical industrial complexes (PIC).” (new page 4, paragraph 2, line 13).

4. Furthermore, petrochemical complexes usually change their productions and technologies over the years, especially complexes that have been operative for decades. Something should be added on the possible variation of contamination in the ambient environment over time due to these changes.

===>[Response]: Thank you for your comments. Indeed, productions might vary over time in PICs. We agree that the possible variation of contamination should be presented. In addition to summarizing information about industrial activity and substantial chemicals in each PIC (see Table 1, new page 30), we also added the following sentence based on this comment in the text: “Pollutants emitted from PIC might vary over time, likely due to the change of manufacturing process and pollution control technology.” (new page 9, paragraph 3, line 19)
METHODS

Review process and data extraction

5. The authors declare that they selected the latest article if there were more than one for the same population.

They should consider to verify the selection strategy since, at least for what I know about the petrochemical plant in Sicily (ref 21), a more recent publication is available.


--->[Response]: We appreciate your suggestions to consider the 2012 publication of Pasetto et al. In our literature search, this paper has already been one of identified papers. Pasetto et al. published another study in 2008, which shared the same follow-up periods and study population. In our previous analysis, we chose the earlier study (Pasetto et al., 2008) because it contained more subjects than the 2012 one (5,911 vs. 5,627), although reasons were not well explained in both papers. We agree to adopt publications in later period. In response to this comments, we have replaced the 2008 one to the 2012 one to avoid overlapping issue (Table 1, study ID=G).

(new page 30) We also added one reference obtained from Comment #1 (WHO’s Human health in areas with industrial contamination). Even we conducted the meta-analysis using new data, our main findings remain the same direction (i.e., pooled RR=1.03). In response to the change, we have revised the manuscript, including Method, Results, and Discussion accordingly.

6. Some of the studies carried out to assess the risk for populations living close to industrial sources of contamination are focused on multiple diseases and have different study designs, therefore it is not simple to identify all the papers that report results on a specific disease. I suggest to check if there are more recent publications of the first and last authors of the selected papers that can be of interest.

--->[Response]: We followed the Cochrane handbook to conduct literature search. Searching publications of the first and last authors of the selected papers is beyond the standard procedure. The number of identified articles increased from 365 to 1,249. We included one additional report from other sources. Our new results have been revised throughout the manuscript, including
Abstract (new page 2, paragraph 3, line 17), Method (from new page 6, paragraph 1, line 3 to new page 7, paragraph 1, line 12), Results (from new page 10, paragraph 2, line 6 to new page 14, paragraph 1, line 2), Discussion (from new page 14, paragraph 3, line 20 to new page 18, paragraph 1, line 9), Figure 1, and all Tables. We also agree to address the study design of each study by adding new column “Study design” in Table 1 (new page 30) and the study design has been noted in the Additional file 1.

7. It is not clear if the authors focused their meta-analysis on lung cancer mortality risk for populations living close to petrochemical complexes only or if they considered industrial areas including petrochemical complexes. They should consider that the following two selected papers regard industrial areas with multiple sources of contamination including petrochemical complexes.

'Does living near a constellation of petrochemical, steel, and other industries impair health?' (ref 16)

'Small area study of mortality among people living near multiple sources of air pollution' (ref 17)

 ===> [Response]: We appreciate your comments on the substantial multiple sources of contamination. In this study, we only included residents living nearby the PICs, which inevitably might involve other industrial processes within the same industrial areas. We agree to add following sentences to address the limitation of this study: “Third, each PIC might involve other manufacturing processes (such as steel, cocking, and power plants) and the exposure level could also be affected by geographical factors across different countries. Limited by the lack of corresponding exposure data, our findings were not able to address the heterogeneity between PICs.” (new page 18, paragraph 1, line 2)

8. It is not clear if the meta-analysis is focused on the effect of air pollution in terms of particulate matter from industrial areas including petrochemical plants, or if it is focused on the effect of the complex milieu of pollutants from petrochemical plants only.

 ===> [Response]: Thank you for addressing this point. Rather than “particulate matter” or “petrochemical plants only,” our exposure of interest was “PIC” per se. In the Introduction, we
have defined our study focusing on PIC: “Petrochemical manufacturing industry, defined as petroleum refining (Standard Industrial Classification code [SIC] 2911) or industrial organic chemicals manufacturing (SIC 2869), involves processes that produce and potentially emit hazardous chemicals into the surrounding air, soil, and water. These petrochemical manufacturing factories are usually clustered in an industrial area together with other manufacturing processes or industry, such as steel, coking, and thermoelectric plants [7, 8], and called petrochemical industrial complexes (PIC).” (new page 4, paragraph 2, line 10)

STATISTICAL ANALYSIS

9. The authors should clarify why they included the characteristic of ethnicity for subgroup analysis.

==> [Response]: We consider ethnicity for subgroup analysis because previous study has reported that the lung cancer mortality might vary with location, gender, and ethnicity. Therefore, in this study, our subgroup analyzed those key factors to exam if the mortality risks of lung cancer are different among those characteristics. We agree to add the citation in the text. (new page 8, paragraph 3, line 19)

10. It is questionable if latency, as defined, can be useful for the subgroup analysis. The authors should clarify what is the rationale for the chosen cutoff. Also, they should consider the categorization by years of observation after a chosen cutoff from the ‘started operation year’.

==> [Response]: We agree with the comment on enhancing the definition of latency and conducting another subgroup analysis based on your suggestion.

The latency period was the starting year of subject follow-up minus the starting year of PIC operation. Lung cancer could take 10–20 years of latency after exposure to ambient carcinogens {1, 2}. Therefore, the cutoff at 20 years could be reasonable. We agree to clarify the rational for the chosen cutoff by adding following sentence in the text: “The estimations on latency period
for lung cancer diagnosis varied widely but usually required approximate years to decades [42, 43].” (new page 16, paragraph 1, line 7)

We further considered another subgroup analysis based on your suggestion of the categorization by years of observation after a chosen cutoff from the 'started operation year.' We defined this new categorization as “bona fide observation.” To be consistent, we chose 20 years after the starting year of operation as the cutoff point and subtracted it from the end year of subject follow-up (see attached response, Table R1 (E)). We further categorized the observation period “10 or more years” and “less than 10 years.” However, we didn’t find significant difference of risk among two groups.

In response to this comment, we have revised the following sentence in the Method: “We further did subgroup analysis by different characteristics [30], including sex, location, ethnicity, PM10 standard, latency period (first year of study period more than 20 years after operation year of PICs vs. less or equal to 20 years), and bona fide observation (defined as 10 or more years of observation after 20 years of PIC operations vs. less than 10 years)” (new page 8, paragraph 3, line 18) and also added the new result in Table 2. (new page 11, paragraph 2, line 10)

Reference:


ASSESSMENT OF DATA QUALITY

11. The assessment of data quality is not clear. The authors should clarify the meaning of the scores and how studies with different study design, as the ones selected for the meta-analysis (cohort, case-control, ecological/small-area), were classified.

===>[Response]: We used the Newcastle-Ottawa Scale (NOS) as a quality assessment tool in our study. The Scale is widely used for cohort and case-control studies. The total score of the
Scale is nine (presented using stars), which combined the scores on three main categories: (1) selection of participants, (2) comparability of study groups, and (3) ascertainment of exposure (case-control studies) or outcome of interest (cohort studies).

(1) Selection of participants (total score is four): The selection domain is subdivided to evaluate the selection of the exposed and non-exposed cohorts, the ascertainment of exposure, and whether the study demonstrated that the outcome of interest was not present at the start of the study.

(2) Comparability of study groups (total score is two): Study that adjusted most important factor received one star, and study that controlled additional factors received two stars.

(3) Ascertainment of exposure (case-control studies) or outcome of interest (cohort studies) (total score is three): Exposure of interest is made up of three questions as well: ascertainment of exposure, same method of ascertainment for cases and controls, and non-response rate. Outcome of interest is made up of three questions: the appropriateness of the methods used to evaluate the outcome, the length of follow-up, and the degree of the loss to follow-up.

The only difference of the study design will show up on the third category—the ascertainment of exposure (case-control studies) or outcome of interest (cohort studies). Study designs have been noted on Additional file 1. We also added column “Study design” in Table 1. (new page 30)

AIR QUALITY STANDARD

12. The criterion of air quality standard is questionable since regulations were not effective in the past in many countries, especially in industrial areas. Furthermore, other pollutants emitted by industrial sources present in petrochemical complexes and causally associable with lung cancer are not considered in this kind of classification. It is not clear why the authors report air quality standard for TSP, PM10 and PM2.5 while they consider only PM10 in the analysis, and why the cutoff of 150 µg/m3 was chosen.
Thank you for your comments. Although regulations might not be effective in the past, we still could reasonably assume that the air quality improved after the implementation of regulations to some extent [1, 2]. We agree to clarify this issue by adding following sentences in the Discussion: “An effective air quality intervention involved a series of steps, including regulatory establishments, pollution reductions, and anticipated improvements in health [44]. Although data on ambient pollution monitoring around PICs in the early periods were very limited and hard to obtain, previous studies have documented pollution reductions could be attributable to changing regulations [45, 46].” (new page 16, paragraph 2, line 11)

We agree that other pollutants might be emitted from the PIC. In line with Comment #1, we have extracted information about industrial activity and substantial chemicals in each PIC and added the information in Table 1. (new page 30)

We applied PM10 for sub-group analysis because PM10 is the most relevant to the study periods among all selected studies compared to TSP and PM2.5 (see below Table R2). Although, some countries proposed TSP to be the air quality indicator in the early years, they changed to PM10 years later. In addition, the TSP is regulated as primary and secondary TSPs, which reduce the comparability across countries and over time. As for PM2.5, the starting-follow-up year was earlier than the year of setting PM2.5 levels among all studied countries, which made the subgroup analysis undoable. As a result, the adoption of PM10 is a more feasible approach than TSP and PM2.5.

The following table (see attached response, Table R2) also shows the PM10 regulation levels at the starting-follow-up year. The cut-off point of 150 µg/m3 was probably most feasible point because four out of 13 subjects were larger than 150 µg/m3. In contrast, selecting a number of less than 150 µg/m3 made the subgroup analysis less meaningful. Therefore, we chose the cutoff of 150 µg/m3 in our analysis.
Reference:


RESULTS AND DISCUSSION

To be revised after addressing the above observations.

==> [Response]: All suggestions have been addressed and corrections have been made as mentioned above. Thank you very much for reviewing our manuscript. We appreciate the opportunity to improve the paper. We have revised our manuscript based on your comments throughout the text accordingly, which have been highlighted by yellow background in the revision.