PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Multiple Air Pollutants Exposure and Leukemia Incidence in Tehran, Iran from 2010 to 2016: a retrospective cohort study
AUTHORS	khorrami, zahra; Pourkhosravani, Mohsen; Eslahi, Marzieh; Rezapour, Maysam; Akbari, Mohammad Esmail; Amini, Heresh; Taghavi-Shahri, Seyed Mahmood; Künzli, Nino; Etemad, Koorosh; Khanjani, Narges

VERSION 1 – REVIEW

REVIEWER	Mansoureh Dehghani Mashhad University of Medical Sciences, oncology
REVIEW RETURNED	18-Feb-2022

GENERAL COMMENTS	The manuscript is well written and organized, and a few changes are required. It is eligible to be published after following revisions: 1. Please mention the the study design type both in the manuscript and in the abstract (supposedly a retrospective cohort study). 2. It is more accurate to specify the objective as determining the correlation between "acute leukemia" (rather than "leukemia") and air pollutants. 3. Line 26 of the abstract page (Outcome measures): "of" is missing. 4. Page 3, strength and limitations: The sentence about confounders needs change to flow more smoothly and convey meaning. 5. Page6, line 26: Please provide more description on k class and k-1 class models. 6. Discussion: Please provide a comparison between the results of the mentioned studies and the present study, along with an analysis of the differences in the results. For example the Canadian study (reference number 30). 7. Strength and limitations are well discussed, though one limitation is missing. Is the length of residence in the study areas and the possibilities of life-time relocation or migration taken into consideration?

REVIEWER	LA Rodriguez-Villamizar Universidad Industrial de Santander
REVIEW RETURNED	27-Feb-2022

GENERAL COMMENTS	This study aimed “to examine the association between estimates of district level averages of multiple ambient air pollutants and leukemia incidence across Tehran districts”. The authors used data for Leukemia cases during 2000-2016 residing in 22 districts in Tehran and exposure data for PM10, SO2, NO, NO2, NOX, VOC from Land Use Regressions (LUR) available for the city. Latent profile Analysis (LPA) was used to classify air pollutants in two exposure profiles. Analysis was conducted at district level using

	covariables from census and previous surveys. The manuscript is clear and concise. Strengths and limitations of sources of data, epidemiological design and geographical unit of analysis are recognized by authors. The study adds important results in the topic of air pollution and leukemia effects, however, there are some comments and suggestions for the authors to consider: Abstract: The conclusion mention " that districts with higher air pollution have higher incidence rate of leukemia cancer ". The sentence might clarify that the association was with air pollutants mixtures or single-pollutants. Background The focus of the manuscript is the analysis of mixtures of air pollutants on leukemia incidence. Analysis of mixtures is a growing topic of research in environmental health, and it will be adequate for this section to introduce the readers in this topic and giving an overall explanation of methodologies used to address complex mixtures analysis, including LPA. In the methods section the authors briefly mention the comparison of LPA with cluster analysis, but an overall presentation of mixtures analysis techniques would be adequate to introduce the readers. Data Sources Exposure assessment: Sources of data are LUR for pollutants developed in 2010. The leukemia cases included a temporal range between 2010 and 2016, therefore, it is required that the authors explain whether a specific temporal adjustment was used or not for assigning the exposure to air pollutants during 2011-2016 at district level. If any temporal adjustment was used, it should be specified in this section. Statistical analyses: A brief explanation of selection of LPA and advantages of LPA over other methods for studying chemical mixtures (not only cluster analysis) might be desirable. Results: The outcome of interest is leukemia and although it was analyzed separately for myeloid and lymphoid leukemia, its epidemiology and potential causal effects differ in children and adults. Therefore, I would recommend the authors to explore a stratified analysis by age exploring results separately for children and adults. Discussion: LPA identified two profiles with clear separation of concentrations of pollutants (low and high), except for SO2, and all pollutants were included in the LPA even with high correlations among them (VOCs). A wider discussion is needed related to the implications of this analysis and the intersection between results of one-pollutant models and LPA. Strengths and limitations of sources of data, epidemiological design and geographical unit of analysis are recognized by authors. Conclusion: conclusion is given in terms of one-pollutant models despite the focus of the paper was the analysis of air-pollutant mixtures.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

The manuscript is well written and organized, and a few changes are required. It is eligible to be published after following revisions:
1. Please mention the study design type both in the manuscript and in the abstract (supposedly a retrospective cohort study).

We added this to the methods and abstract sections.

2. It is more accurate to specify the objective as determining the correlation between "acute leukemia" (rather than "leukemia") and air pollutants.
This was mentioned in the methods section, “Information about leukemia patients (**acute** lymphoblastic leukemia (ALL) or **acute** myeloid leukemia (AML)) residing in Tehran, diagnosed between 2010 and 2016 and their residential address (on a district basis) were obtained from the Ministry of Health’s Cancer Registry”.

3. Line 26 of the abstract page (Outcome measures): "of" is missing.
OK. We added that.

4. Page 3, strength and limitations: The sentence about confounders needs change to flow more smoothly and convey meaning.
We made some corrections.

5. Page 6, line 26: Please provide more description on k class and k-1 class models.
A significant p-value of the LMR LRT and VLMR LRT (i.e. $P < 0.05$) indicates a significant improvement in model fit in the k-class model compared to the (k - 1)-class model and thus rejects the (k - 1)-class model and suggests choosing a model with k classes.
In other words, each number of classes (or profiles, shown as k) is compared to the number of classes which is 1 unit less (k-1), and if the VLRT and LMRT values are not significant for higher classes, the model with less classes (k-1) is preferred and will get chosen. For example, in the present study, in more than 2 classes VLRT and LMRT were insignificant. Therefore, the model with 2 classes was selected.

6. Discussion: Please provide a comparison between the results of the mentioned studies and the present study, along with an analysis of the differences in the results. For example, the Canadian study (reference number 30).
OK. We added some text.

7. Strength and limitations are well discussed, though one limitation is missing. Is the length of residence in the study areas and the possibilities of life-time relocation or migration taken into consideration?
OK. We added that.

Reviewer 2

This study aimed “to examine the association between estimates of district level averages of multiple ambient air pollutants and leukemia incidence across Tehran districts”. The authors used data for Leukemia cases during 2000-2016 residing in 22 districts in Tehran and exposure data for PM10, SO2, NO, NO2, NOX, VOC from Land Use Regressions (LUR) available for the city. Latent profile Analysis (LPA) was used to classify air pollutants in two exposure profiles. Analysis was conducted at district level using covariables from census and previous surveys.
The manuscript is clear and concise. Strengths and limitations of sources of data, epidemiological design and geographical unit of analysis are recognized by authors. The study adds important results in the topic of air pollution and leukemia effects, however, there are some comments and suggestions for the authors to consider:

Abstract:
The conclusion mention “that districts with higher air pollution have higher incidence rate of leukemia cancer “. The sentence might clarify that the association was with air pollutants mixtures or single-pollutants.
This relation was seen in both single-pollutant models (NO_2 and NO_X) and multi-pollutant models. We added this to the conclusion.

Background
The focus of the manuscript is the analysis of mixtures of air pollutants on leukemia incidence. Analysis of mixtures is a growing topic of research in environmental health, and it will be adequate for this section to introduce the readers in this topic and giving an overall explanation of methodologies

used to address complex mixtures analysis, including LPA. In the methods section the authors briefly mention the comparison of LPA with cluster analysis, but an overall presentation of mixtures analysis techniques would be adequate to introduce the readers.
We added an explanation about LPA in the end of the introduction section.

Data Sources
Exposure assessment: Sources of data are LUR for pollutants developed in 2010. The leukemia cases included a temporal range between 2010 and 2016, therefore, it is required that the authors explain whether a specific temporal adjustment was used or not for assigning the exposure to air pollutants during 2011-2016 at district level. If any temporal adjustment was used, it should be specified in this section.
We did not assign the exposure to air pollutants during 2011-2016. Leukemia takes time to develop. The 2010 data is being used as an estimate of exposure levels that people faced before developing leukemia.

Statistical analyses: A brief explanation of selection of LPA and advantages of LPA over other methods for studying chemical mixtures (not only cluster analysis) might be desirable.
Latent profile analysis (LPA) as a person-centered approach can be used to examine the patterns of multiple air pollutants. LPA is a statistical method for identifying unobserved subgroups within populations based on observed indicators. In contrast to traditional methods, such as cluster analysis, LPA has several advantages. LPA does not require researchers to determine the number of profiles beforehand, which is more suitable to answer research questions that are exploratory in nature, and empirical indicators are available to determine the optimal number of profiles. In addition, LPA allocates individuals to subgroups probabilistically, taking into account the rate of classification uncertainty. Also, LPA used multiple statistical indices for determining the optimal number of subgroups.
We added this to the text (in statistical analysis).

Results: The outcome of interest is leukemia and although it was analyzed separately for myeloid and lymphoid leukemia, its epidemiology and potential causal effects differ in children and adults. Therefore, I would recommend the authors to explore a stratified analysis by age exploring results separately for children and adults.
Yes, that is a good suggestion. We re-ran the analysis for adults and children separately, and we added two tables.

Discussion: LPA identified two profiles with clear separation of concentrations of pollutants (low and high), except for SO2, and all pollutants were included in the LPA even with high correlations among them (VOCs). A wider discussion is needed related to the implications of this analysis and the intersection between results of one-pollutant models and LPA. Strengths and limitations of sources of data, epidemiological design and geographical unit of analysis are recognized by authors.
We added some text to the discussion. If the reviewer needs more information, he/she has to be more specific.

Conclusion: conclusion is given in terms of one-pollutant models despite the focus of the paper was the analysis of air-pollutant mixtures.
OK. We added that.

VERSION 2 – REVIEW

REVIEWER	LA Rodriguez-Villamizar Universidad Industrial de Santander
REVIEW RETURNED	28-Apr-2022

GENERAL COMMENTS	The authors have included all the suggestions and the results for children and adults resulted interesting. The manuscript is ready for publication.