Author's response to reviews

Title: Ambient pollutants, polymorphisms associated with microRNA processing and adhesion molecules: the Normative Aging Study

Authors:

Elissa Wilker (ewilker@bidmc.harvard.edu)
Stacey Alexeeff (sackerma@hsph.harvard.edu)
Helen Suh (hsuh@hsph.harvard.edu)
Pantel Vokonas (pantel.vokonas@va.gov)
Andrea Baccarelli (andrea.baccarelli@unimi.it)
Joel Schwartz (jschwartz@hsph.harvard.edu)

Version: 2 Date: 23 November 2010

Author’s response to reviews: see over
Response to the editor’s comments:
Reviewer's report
Title: Ambient pollutants, polymorphisms associated with microRNA processing and adhesion molecules: the Normative Aging Study
Version: 1 Date: 4 September 2010
Reviewer: Ralph J Delfino
Reviewer’s report:
Review of Wilker et al. Environ Health “Ambient pollutants, polymorphisms associated with microRNA processing and adhesion molecules: the Normative Aging Study”

General Comments
This interesting study addresses a novel area of research on effect modification of air pollutant exposure-response relations by polymorphisms in genes that process microRNAs. Authors found that a SNP in an miRNA processing gene (rs1062923) modifies associations of sICAM-1 and sVCAM-1 levels with 7-day average PM2.5 and SO42-.

Major concerns:

*Comment 1.
There was no discussion in the limitations section about the limited number of repeated measures, including subjects with only one measurement (N unknown). What were the results dropping subjects with no repeated measures? In other words, if all subjects had only one measurement, then this would be a cross-sectional study with its potential biases. This is not the case of course, but the separation of the few measurements by months or years adds to the potential for temporal confounding, which should be more vigorously controlled for in the analysis. By chance, different subjects with high VCAM and ICAM could have come for an evaluation during periods of high air pollution (a sort of temporal / cross-sectional bias).

Response 1.
We agree with the reviewer that it is important to address potential temporal confounding. First we want to point out that a mixed model does not require repeated measurements on all subjects. An unbalanced design generally results in unbiased estimates of the fixed effects, although inefficient estimates of the random intercepts for
the subjects with only one measurement. Specifically, a linear mixed model with a random subject intercept can be applied to longitudinal data with unequal number of observations per subject, as long as the missingness mechanism is at least missing at random (MAR), which means that the probability that responses are missing might depend on the set of observed responses, but is unrelated to the missing values that should have been obtained. Hence, even if subjects are less likely to return four years later if they have higher VCAM, or body mass index, or other covariate measured previously, there is no bias. There is only bias if the likelihood of returning in 4 years depends on air pollution that day or sVCAM-1 that day, even after conditioning on the previous visits levels of those variables. We believe this is highly unlikely, and therefore the missing data in this case should be at least MAR. As long as there are multiple observations for some subjects, the variability can be partitioned into between- and within-subject variability, and therefore, the model is estimable.

The number of measurements for each subject is included in Table 1. Of the 723 participants on whom complete data are now available, only 185 (26%) had only one visit.

For the revised version of the manuscript, we performed an analysis in which we restricted to individuals with 2 or more measures and found that this did not significantly change our results. In addition, to more vigorously control for seasonal confounding, we conducted analyses using sine and cosine terms for month of the year. Since this did not alter our conclusions, we include the results adjusted for seasonality with sine and cosine in this revised manuscript. We made the following changes in the text to reflect this:

We have added the following text in the methods section, under the subheading “sensitivity analyses”:

Because all participants did not have the same number of follow-up visits, and those who had the most may have been healthier than average, we assessed whether this differential selection influenced our results. We performed sensitivity analyses restricting to participants with 2 or more visits. In addition to testing whether 5 and 9 day moving averages produced significantly different associations, we also examined whether adjusting for season using dichotomous variables for Winter, Spring and Summer (treating Fall as the referent category) altered results. Finally, we performed analyses excluding current smokers and individuals with high CRP levels.

And in the results:

We performed a number of sensitivity analyses to test the robustness of our results. In our analyses, restricting to participants with two or more sICAM-1 and sVCAM-1 did not change our findings. Using indicator variables for season instead of sine and cosine for month of year produced similar results and results. Finally, we observed that neither restricting to non-smokers, nor restricting to participants with CRP<10mg/dL altered the results of our pollutant effects for 7 day moving averages of PM$_{2.5}$.

*Comment 2.*
The paper gives the appearance of a selective choice of exposure averaging time even though there were statements suggesting that averaging times were to be explicitly tested in the analysis. There was insufficient direct evidence that the choice of a 7-day average was more strongly associated with outcomes than shorter times.

**Response 2.**

We agree with the reviewer that a 7 day average is not conclusively more strongly associated with the outcomes tested. However, we did deliberately choose to focus on longer moving averages than the more traditional 2 day average, given the evidence that we have observed from our previous work examining blood pressure, as well as the MONICA Augsburg results for CRP, which suggest averaging times in that range. We chose a 7 day average to capture longer term, but still acute effects, less because of strong evidence for seven (as opposed to five or nine) than from the canonical nature of a week average. We agree that is not a strong argument, but we wanted to avoid testing multiple moving averages, which risks overfitting and raises issues of multiple comparisons. We have now included the sensitivity analyses at other averages in Table 2. Results from our sensitivity analyses may suggest that even longer averaging times are relevant for PM$_2.5$, particularly in association with sICAM-1, where the magnitude of the effect was stronger at 9 days than at 7.

We have reworded the conclusion to avoid giving the appearance that a 7 day moving average is somehow more strongly associated with sICAM-1 and sVCAM-1 levels than other time points:

The conclusion has been restated:

This study provides novel evidence of effect modification of the relationship between exposure to particulate matter and biomarkers of inflammation endothelial function in a population of elderly community-dwelling men. It also suggests that air pollution from SO$_4^{2-}$ may play a role in these processes, especially over longer averaging times. For the 7 day moving averages examined, strong evidence of effect modification was found for the rs1062923 SNP in GEMIN4. Our results suggest that future work investigating epigenetic regulation should consider exposure to ambient pollution as a potential marker of susceptibility. By examining the association between such biomarkers and air pollution, this paper adds to the growing body of evidence that elevated levels of particulate air pollution may induce cardiovascular effects through an interrelated process of inflammation and endothelial dysfunction.

**Comment 3.**

A major influence of smoking on adhesion molecules suggests that the analysis should be tested excluding smokers (who would be differently affected by ambient air pollution anyway).

**Response 3.**

In this population, fewer than 5% of participants are current smokers (n=34). Given the small number of current smokers, it is unlikely that smoking would
significantly affect our results. To test this hypothesis, we performed a sensitivity analysis in which we excluded these participants and our results were unchanged and the interactions with rs1062923 remained highly significant. We have added text to the manuscript to indicate that this association was tested and did not change our conclusions (described above in Response 1).

**Specific Comments:**

**Abstract:**

*Comment 4.*

Reconsider using the term endothelial dysfunction as equivalent to ICAM and VCAM changes with which it is associated. Endothelial dysfunction denotes an imbalance between vasodilation and vasoconstriction.

**Response 4.**

We have reworded the abstract. The conclusions are now written as:

PM$_{2.5}$ 7 day moving averages are associated with higher sICAM-1 and sVCAM-1 levels. SO$_4^{2-}$ 7 day moving averages are associated with higher sICAM-1 and a suggestive association was observed with sVCAM-1 in aging men. SNPs in miRNA-processing genes may modify associations between ambient pollution and sICAM-1 and sVCAM-1, which are correlates of atherosclerosis and cardiovascular disease.

**Introduction:**

*Comment 5.*

P 4, line 112: These references describe not only cardiovascular function (blood pressure) but also blood biomarkers, which are not measures of function. Are there any other similar studies with evidence supporting longer averaging times?

**Response 5.**

With regards to 7 day moving averages and association with blood pressure, our group has published at least 3 papers examining 5-7 day moving averages [1-3]. Additionally, we note that our group has observed temperature associations over even longer moving averages to be associated with sICAM-1 and sVCAM-1 [4].

As for measuring this time frame with blood biomarkers, Dubowsky et al observed associations between PM$_{2.5}$ and white blood cells which reached significance only at a 7-day mean [5]. In addition, the MONICA Augsburg study reported an association of CRP with particles that was observed for 5 day moving averages [6].

We have added additional references and reworded this sentence:

Most studies have examined acute effects in the 24 to 48 hour range, but longer averaging times of approximately one week may be more relevant to understanding the associations between particles and cardiovascular function and processes [3, 5-8].

**Methods:**

*Comment 6.*
P 7: How far was the stationary site from subject residences (range and average)?

Response 6.

The median is 19 km and the IQR is 10 km to 33 km. Given that PM$_{2.5}$ has a high degree of regional homogeneity, we believe that the use of a stationary monitor is a reasonable surrogate here. We discuss the limitations in our discussion section:

We have utilized stationary measures of air pollution to represent personal exposures. Prior research indicates that when examining longitudinal exposures to air pollution, most error is of the Berkson type. To the extent that it is classical, simulation studies have shown that it is highly unlikely to bias away from the null even in the presence of covariates indicates that this exposure misclassification may lead to an underestimation of the health effects of air pollution [47]. In addition, several studies, including one conducted in the Greater Boston area, have found that longitudinal measures of ambient particulate concentrations are representative of longitudinal variation in personal exposures [48].

*Comment 7.

P 8: You say “Our a priori hypothesis was to examine 7 day moving averages.” However, in the Introduction and abstract a hypothesis referred to showing differences by shorter vs. longer averaging times. This requires multiple averaging times in the analysis to answer. I suggest a typical lag 0 and 2- or 3-day average. Later you mention 5 and 9-day averages were examined and were more weakly associated, suggesting that odd averaging times were tested including 1, 3, 5, 7 and 9 days. All of these results should be shown

Response 7.

In a previous analysis, we examined shorter moving averages [7]. Therefore, in this analysis we chose to look at longer moving averages. We observed significant associations with PM$_{2.5}$ at longer moving averages, but not with BC, whereas our previous paper showed associations with BC at shorter moving averages and not with PM$_{2.5}$.

We agree that sensitivity analyses should be shown, and we now include 5, 7 and 9 day moving averages for all 3 pollutants in Table 2.

*Comment 8.

P 9, top, analysis: Since subjects came into the clinic during different seasons and years on one or a few occasions, then this should be controlled for in the analysis. By chance, different subjects with high VCAM and ICAM could have come for an evaluation during periods of high air pollution (a sort of cross-sectional bias).

Response 8.

We have added the following text to the manuscript to address this point: as addressed in the first major comment (see above). Briefly, we have addressed this concern by determining that our conclusions remained unchanged when we restricted to individuals with 2 or more observations during the study period.
Season is controlled for by sine and cosine terms within our models. Also, we include a term for time since baseline visit. We found that adjusting for season using dummy variables produced very similar results.

Results:
*Comment 9.
P 11: a bit over half way down the sentence (An IQR …” needs rewording.
Response 9.
We have changed the sentence to the following:

An IQR change in $SO_4^{2-}$ (1.39 µg/m$^3$) was associated with 1.4 (95% CI: 0.04, 2.7) higher 
sICAM-1 levels, and 1.6 (95% CI: -0.4, 3.7) higher sVCAM-1. BC at this averaging time 
was not significantly associated with either measure.

*Comment 10.
P 11 end 1st paragraph: It would informative to show the results for BC and for 
other averaging times as discussed above since it is unlikely you could have 
predicted which one was the strongest.
Response 10.
As described above, we now show the results for all pollutants in table 2.

*Comment 11.
P 12: although it is unlikely that BC would interact with the SNPs being itself NS, 
it is possible. Since we cannot see BC results, it’s hard to evaluate this 
possibility.
Response 11.
Although it is certainly possible that an interaction exists, we chose to limit the 
number of associations tested for this analysis so that our approach to addressing false 
positives adequately reflects the number of tests performed.

*Comment 12.
Tables 4-5 results for SNPs appear selective. It would be useful to see results for 
other SNPs.
Response 12.
We now show the results for all SNPs in an online appendix in order to be 
complete. Adjusted P-values for all of the other SNPs were 0.3 or higher. Given that 
the effects for all of the other SNPs did not approach significance, we felt that it would 
ot be appropriate to include these results, as the effect estimates for PM$_{2.5}$ in carriers 
and non-carriers of the variants are clearly non-significant in our models.

Discussion:

*Comment 13.
P 13, 1st para: You say “In this repeated measures study” but in fact, some unknown number of subjects had only one measurement. To really claim that this is the design you would have to drop those subjects. What happens to the results when you do that?

Response 13.

The number of subjects with one measurement is not unknown, as it is included in table 1 (n=185). When we drop these subjects, the results change by <.10% and remain highly significant. Additionally, both the main effect of PM$_{2.5}$ and the interaction with rs1062923 remain significant. Therefore, we feel it is appropriate to present the results including these subjects.

*Comment 14.

P 13, 1st para: You say “Results from our sensitivity analyses may suggest that even longer averaging times are relevant …” but this analysis is not shown.

Response 14.

We thank the reviewer for this comment. In our Table 2 we now show the 5 and 9 day moving averages for completeness.

Comment 15. P 13, 1st para: You say “carriers of the variant had lower levels of the markers of endothelial function” but you did not measure endothelial function.

Response 15.

We have changed the wording of this sentence to clarify that we are referring specifically to changes in sICAM-1 and sVCAM-1 rather than endothelial function:

In both cases, carriers of the variant had lower levels of sICAM-1 and sVCAM-1.

*Comment 16.

P 13, 1st para: plural SNPs was used but results for only one SNP were shown.

Response 16.

This sentence now says,

We observed evidence of effect modification a single SNP involved in the processing of miRNA.

*Comment 17.

P 13, 2nd para: association of what with BC?

Response 17.

We apologize for the ambiguity and have clarified this sentence:

This contrasts with another recent publication from the Normative Aging Study which examined 2 day moving averages of particles and only observed an association between sICAM-1 and sVCAM-1 with BC [13]. We observed significant associations for sICAM-1 in our models of 7 day moving averages for both PM2.5 and marginally significant (p=0.05) results for SO$_4^{2-}$. Previous work from our group and others has suggested that traffic pollution may adversely affect cardiovascular health and many of these studies have observed stronger effects with the components of air pollution associated with traffic particles [8, 13].

*Comment 18.
Previous work from our group and others has suggested that traffic pollution may adversely affect cardiovascular health and many of these studies have observed stronger effects with the components of air pollution associated with traffic.

Data describing the effects of miRNA processing on cardiovascular processes and inflammation is relatively limited [27, 42]. Even less is known about the role these processing genes play in interacting with environmental exposures.

Results were not presented to evaluate the statement: “rs1062923-PM2.5 crossproduct predicting adhesion molecule level was consistently the strongest of all interactions tested”

We have added this information to a supplemental table, and we now mention in the text that the adjusted p-values for these associations were 0.3 or higher.

This sentence was intended to point out that findings from our study may differ from previous work examining different (shorter) averaging times and due to the limitations in generalizibility. We believe that the use of a mixed model given our data structure is appropriate to analyze the data here. We agree that there is a possibility of temporal confounding and have addressed that by using the linear time trend plus sinusoidal seasonal control. We also believe that the long intervals between repeated
measures gains us some power, since short term serial correlation between today’s air pollution and yesterday’s air pollution is no longer an issue.

**Level of interest:** An article of importance in its field  
**Quality of written English:** Acceptable  
**Statistical review:** Yes, and I have assessed the statistics in my report.  
**Declaration of competing interests:**  
I declare that I have no competing interests

**Title:** Ambient pollutants, polymorphisms associated with microRNA processing and adhesion molecules: the Normative Aging Study  
**Version:** 1  
**Date:** 31 August 2010  
**Reviewer:** Melanie Jardim  
**Reviewer’s report:**  
In this manuscript, Wilker et. al. link genetic variants in genes which regulate miRNA processing, with serum levels of ICAM and VCAM in response to air pollutant exposure. Although the results presented here have not previously been shown in the human population there are serious concerns regarding the actual science as it is presented in this manuscript.

**Major Concerns**

**Comment 1.**  
This is not a microRNA paper. Therefore, the word “microRNA” should not appear in the title. It seems as if the authors are merely using this as a ploy to get their research published knowing that the field of microRNAs is booming with popularity.  
**Response 1.**  
These SNPs were selected because they are located in microRNA processing genes. Because this is the criterion by which SNPs were selected, we believe that titling the manuscript ‘Ambient pollutants, polymorphisms associated with miRNA processing and adhesion molecules: the Normative Aging Study” is appropriate.

**Comment 2.**  
The very small amount of data presented here is quite weak. I am not convinced that the detection of ICAM and VCAM is appropriate in this case. It would have been nice to see a set of in vitro experiments designed to test this hypothesis to accompany the human data, which understandably can be limited. It just seems like the authors forgot to connect a few dots.  
**Response 2.**  
We agree with the reviewer that in vitro experiments provide a wealth of important information about mechanistic associations. However, this is an epidemiologic study undertaken with the purpose of examining associations in a population of community-dwelling individuals. We believe that the analysis of repeated
measures in an observational setting provides unique and relevant information about exposures in free-living individuals, which are more likely to be representative of health effects on a population level. We do not agree that epidemiology studies must be accompanied by in vitro experiments to be publishable, any more that in vitro findings should not be published unless they can demonstrate similar in vivo results in humans.

**Comment 3.**
This manuscript lacks novelty especially since this group has already published a paper on SNPs in microRNA processing genes with the same dataset. The authors also seem to forget that it is currently estimated that approximately 60% to 70% of coding mRNAs are thought to be regulated by microRNAs. How is it surprising then that deviation from appropriate processing of miRNAs leads to changes in protein levels (any protein)?? It is true that deviation from adequate processing may affect the biogenesis of miRNAs, but it is not currently clear how these SNPs do so.

**Response 3.**
While it may not be surprising that some changes in miRNA would be associated with ICAM and VCAM levels, this is not what we are examining. First, we are examining genes related to a specific set of miRNA processing functions, and second, we are not looking at the association of SNPs in these genes with the outcome, we are looking at them as modifiers of the association with air pollution. We think, prior to this analysis, it was not at all clear that they would modify the air pollution association. We agree that it is not clear how these SNPs modify the effects of air pollution, and think that would be a useful area of future research.

We did examine this issue in a previous paper looking at the association of air pollution with blood pressure. We believe this is a sufficiently different exercise, since (a) Blood pressure and sICAM-1 and sVCAM-1, while both markers of cardiovascular disease, are controlled by very different biological processes and (b) different air pollutants have different physiologic impacts, and so the pollutants associated with markers of endothelial activation, and the averaging times for those associations, and the pathways which might modify those associations, could well be different. Therefore we do not believe this finding is a minor addendum to the first. We do believe that our finding of overlap in the SNPs which were observed to significantly modify blood pressure and ICAM-1 and VCAM-1 adds credibility, and suggest that it would be interesting to see whether these associations are observed in other populations and in case control and in vitro settings as well.

**Comment 4.**
Overall, the manuscript just seems to lack information. There didn’t seem to be enough data there to warrant publication at this time. The authors should perhaps re-evaluate their story and take the time to adequately think through the experiments/data that are necessary to make this manuscript more complete.

**Response 4.**
We believe that our paper provides information on the cohort analyzed, the covariates controlled for, the exposures used, the outcomes examined, the methods of used to measure ICAM and VCAM, the methods used for genotyping, and the statistical
analysis used to examine the association. We believe this is the information required to
review, and publish, and epidemiology study. Since the reviewer does not state what
additional information about our study we should provide, we cannot respond further at
this time. In addition, while the interaction discussed in this manuscript is one
component we wish to highlight, we also hope to draw attention to the findings we
observed for associations between pollutants and levels of adhesion molecules. The
findings of associations with PM$_{2.5}$ at longer moving averages is novel, as is the finding
of an association with sulfates.

**Level of interest:** Too insignificant to warrant publication in any journal

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a
statistician.

**Declaration of competing interests:**
I declare that I have no competing interests

References


5. Dubowsky SD, Suh H, Schwartz J, Coull BA, Gold DR: **Diabetes, obesity, and hypertension may enhance associations between air pollution and markers of systemic inflammation.** Environ Health Perspect 2006, 114(7):992-998.