HEALTH EFFECTS INSTITUTE

STATEMENT

Synopsis of Research Report 200

Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1

INTRODUCTION

The levels of most ambient air pollutants have declined significantly in the United States during the last few decades. Recent epidemiological studies, however, have suggested an association between exposure to ambient levels of air pollution — even below the current U.S. National Ambient Air Quality Standards (NAAQS) — and adverse health

What This Study Adds

- This study is part of an HEI program to address questions regarding potential associations between air pollution exposure and health outcomes at low ambient air pollution levels, particularly at levels below the current U.S. national air quality standards.
- Dominici and colleagues developed hybrid, U.S.-wide models using machine learning to estimate outdoor fine particle (particulate matter ≤ 2.5 μm in aerodynamic diameter, or $PM_{2.5}$) and ozone (O_3) concentrations at 1 km × 1 km grids, by combining monitoring, satellite, transport modeling output, and other data.
- They obtained Medicare data for 61 million Americans, ages 65 years and older, who enrolled between 2000 and 2012. Using both cohort and case–crossover designs, they analyzed the association between long-term and short-term outdoor $PM_{2.5}$ and O_3 exposures and mortality.
- The investigators report positive associations between nonaccidental, all-cause mortality and $PM_{2.5}$ and O_3 at low concentrations, including below the U.S. National Ambient Air Quality Standards (annual 12 μg/m^3 for $PM_{2.5}$ and 8-hour 70 ppb for O_3).
- These associations were robust to most adjustments for potential confounding by a number of lifestyle and behavioral factors in the cohort analyses. Sensitivity analyses did not meaningfully impact the findings of association.
- HEI's Low-Exposure Epidemiology Studies Review Panel noted, however, that several important issues still need to be addressed by the investigators regarding these results during the remainder of this project. In particular, the potential for confounding by time and the complexities introduced by the use of different spatial scales for the exposure and health data need to be explored in more detail, and the causal inference methods need to be more fully applied.
- The Panel concluded that Dominici and colleagues have conducted an extensive and innovative set of initial analyses in these extraordinarily large air pollution and health data sets. While initial conclusions may be drawn from these analyses, the Panel awaits the further analyses that are underway before reaching full conclusions on the air pollution and public health implications of this important research.

This Statement, prepared by the Health Effects Institute, focuses on the Phase 1 results of a research project funded by HEI and conducted by Dr. Francesca Dominici at the Harvard T.H. Chan School of Public Health, Boston, Massachusetts, and colleagues. The Statement is a synopsis of the Commentary, which summarizes the study, its findings, and its strengths and limitations. The complete report, *Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution: Phase 1* (© 2019 Health Effects Institute), can be obtained from HEI or our website (see last page).

effects. In view of the importance of such research findings, the Health Effects Institute in 2014 issued a request for applications (RFA 14-3) seeking to fund research to assess the health effects of long-term exposure to low levels, particularly below the NAAQS, of ambient air pollution and to develop improved statistical methods for conducting such research. HEI funded three studies under this program; each study used state-of-the-art exposure methods and very large cohorts. The studies were based in the United States, Canada, and Europe, thus providing a comprehensive cross-section of high-income countries where ambient levels are generally low.

The low-exposure-level studies are scheduled to be completed in 2020. In 2018, in order to inform the ongoing review of the NAAQS for fine particles ($PM_{2.5}$) and ozone (O_3), HEI requested Phase 1 reports from the U.S. (Francesca Dominici) and Canadian (Michael Brauer) investigators. HEI formed a special panel, the Low-Exposure Epidemiology Studies Review Panel, to evaluate the studies' methods, results, conclusions, and their strengths and weaknesses. This Statement focuses on the study by Dr. Francesca Dominici, from the Harvard T.H. Chan School of Public Health, Boston, Massachusetts, and her colleagues, titled, "Assessing Adverse Health Effects of Long-Term Exposure to Low Levels of Ambient Air Pollution."

APPROACH

Aims: The aims of the Dominici study were to (1) develop hybrid, high-resolution, exposure-prediction models to estimate long-term exposures to $PM_{2.5}$ and O_3 levels for the continental United States; (2) develop and apply causal inference methods; (3) estimate all-cause mortality associated with exposure to ambient air pollution for all U.S. Medicare enrollees between 2000 and 2012 using a cohort (long-term) and a case–crossover (short-term) design; and (4) develop tools for data sharing, record linkage, and statistical software.

Data and Methods: Dominici and colleagues developed hybrid air pollution concentration models for the contiguous United States for the period 2000 to 2012, using data from a variety of sources, including satellite data, chemical transport models, land-use and weather variables, and routinely collected air monitoring data from the U.S. Environmental Protection Agency (EPA).

With this large amount of data and using multiple approaches and input variables, the investigators developed a hybrid model to estimate daily $PM_{2.5}$ and O_3 concentrations at 1 km × 1 km grids across the continental United States. Complex atmospheric processes were addressed using a neural network that modeled nonlinearity and interactions. The neural network was trained using data covering the study period, and the predictions were validated against 10% of the EPA air monitors left out of the model. A similar approach was used to estimate and validate a model to predict O_3 concentrations during the warm months (April through September) of each study year.

Health data were obtained from the Centers for Medicare and Medicaid Services for all Medicare enrollees for the years 2000 to 2012, which represents more than 96% of the U.S. population 65 years of age and older (see Statement Table). The team obtained

Statement Table. Key Features of the Dominici et al. Study

Overall	
Medicare study population	60.9 million
MCBS study population	57,200
Study period	2000–2012
Case–Control Study	
Follow-up period	460.3 million person-years
Deaths	22.6 million
$PM_{2.5}$ average concentration	11.0 $\mu g/m^3$
O_3 average concentration	46.3 ppb
Case–Crossover Study	
Case days	22.4 million
Control days	76.1 million
$PM_{2.5}$ average concentration	11.6 $\mu g/m^3$
O_3 average concentration	37.8 ppb

records for all Medicare enrollees (~61 million), with 460 million person-years of follow-up and 23 million deaths. They also obtained covariate information from the Medicare Current Beneficiary Survey (MCBS; ~57,000 people), an annual phone survey of a nationally representative sample of Medicare beneficiaries, with information on more than 150 individual-level risk factors, including smoking and body mass index.

Using the Medicare data and cohort and case–crossover designs, they investigated the association between exposure to $PM_{2.5}$ and O_3 and all-cause mortality in two-pollutant analyses, including separate analyses for low pollutant concentrations. For the cohort study, they performed survival analyses using the Andersen-Gill method, a variant of the traditional Cox proportional hazards model that incorporates spatiotemporal features by allowing for variation in covariates from year to year. The investigators developed concentration–response curves by fitting a log-linear model with thin-plate splines for both pollutants while controlling for important individual and ecological variables, including socioeconomic status and race. For the case–crossover study, the case day was defined as the date of death, with exposure defined as the mean of the ambient concentration on that day and the day before; this was compared to exposure on three predefined control days. They fitted a conditional logistic regression to all pairs of case and matched control days, thus estimating the relative risk of all-cause mortality associated with short-term exposure to $PM_{2.5}$ and O_3. They also performed subanalyses to explore the health effects at lower levels of exposure.

To assess whether any subgroups within the cohort study were at higher or lower risk of mortality associated with long-term air pollution exposure, the investigators fitted the same statistical models to certain population subgroups (e.g., male vs. female and white vs. black). To explore the robustness of the results from the cohort analysis, they performed sensitivity analyses and compared any changes in risk estimates with differences in confounder adjustment and estimation approaches. Finally, since Medicare data do not include information on many important individual-level covariates, the investigators utilized data from the Medicare Current Beneficiary Survey to examine how the lack of adjustment for these risk factors could have affected the risk estimates for the Medicare cohort.

RESULTS

Dominici and colleagues report overall good performance of the models for estimating $PM_{2.5}$ and O_3 concentrations, with overall R^2 values of 0.84 and 0.80, respectively. For $PM_{2.5}$, the average annual concentration was 11.0 μg/m^3 during the study period, 2000–2012. Performance of the model varied between different geographical regions and seasons; the highest $PM_{2.5}$ concentrations were predicted to be in California and the eastern and southeastern United States, and model performance was better in the eastern and central United States than in the western part of the country. And, the $PM_{2.5}$ model performed best during the summer. For O_3, the average of 8-hour daily concentrations during the warm season was 46.3 ppb during the study period. O_3 concentrations were highest in the Mountain region and in California and lower in the eastern states. The average concentrations of $PM_{2.5}$ decreased during the study period, but O_3 concentrations remained more or less the same. Annual $PM_{2.5}$ and warm-season O_3 concentrations were only weakly correlated.

The 2000–2012 cohort of Medicare beneficiaries provided a very large population for studying association with long-term effects of exposure to ambient air pollution. In two-pollutant analyses of long-term effects, Dominici and colleagues report a 7.3% higher risk of all-cause mortality for each 10-μg/m^3 increase in annual average $PM_{2.5}$ concentrations and a 1.1% higher risk of mortality for each 10-ppb increase in average O_3 concentrations in the warm season. At low concentrations — less than 12 μg/m^3 $PM_{2.5}$ and less than 50 ppb O_3 — the risk was 13.6% for $PM_{2.5}$ and 1.0% for O_3 for each 10-μg/m^3 and 10-ppb increase in concentrations, respectively. The concentration–response relationships from the two-pollutant models showed almost linear curves, with no suggestion of a threshold down to 5 μg/m^3 $PM_{2.5}$ and 30 ppb O_3.

In subgroup analyses for long-term $PM_{2.5}$ exposure, the investigators found larger estimates of effect among males and among Hispanics, Asians, and particularly African Americans, compared with whites. Individuals with low socioeconomic status, as indicated by eligibility for Medicaid, appear to have a slightly higher risk per unit of $PM_{2.5}$ exposure. For long-term O_3 exposure, the subgroup analysis showed that the effect estimates were higher for Medicaid-eligible enrollees and slightly higher for whites, but these analyses also produced puzzling hazard ratios of less than 1 for certain subgroups, including Hispanics and Asians, and particularly for Native Americans, than the overall population.

For short-term exposures, the investigators observed a 1.05% greater risk of mortality in two-pollutant models for a 10-μg/m^3 increase in $PM_{2.5}$ concentrations and a 0.51% greater risk for a 10-ppb increase in 8-hour warm-season O_3 concentration.

(Pollutant levels were averaged over the current and previous day.) At low concentrations (below 25 μg/m^3 of $PM_{2.5}$ and below 60 ppb of O_3), the associations remained elevated for both pollutants (1.61% for $PM_{2.5}$ and 0.58% for O_3). The concentration–response curves showed the relative risk increasing sharply for both pollutants at a relatively low concentration and then leveling out at higher concentrations. The investigators observed evidence of effect modification for several variables, including a higher $PM_{2.5}$–mortality risk for females than for males.

INTERPRETATION OF RESULTS

In its independent review of the study, HEI's Low-Exposure Epidemiology Studies Review Panel noted that the report by Dominici and colleagues summarizes an impressive amount of work completed in the first part of this HEI project. Particularly strong aspects of this work include the extremely large, national cohort, with high-resolution exposure assessment and development and application of state-of-the-art statistical techniques. The Panel also noted that additional research, including further development of causal methods that would properly allow for the complexities in the design of the studies and nature of the data, is currently ongoing.

Exposure Assessment: The use of large, diverse, and existing data sets to generate estimates of $PM_{2.5}$ and O_3 concentrations on a 1 km × 1 km grid for the entire continental United States (~8 million km^2) is impressive, and allowed the investigators to estimate concentrations in areas where air monitors are sparse. However, as with any exposure assessment, it is critical to consider the potential for exposure prediction errors.

Despite steps to correct for regional and compositional differences, both geographical and temporal variability in the errors of the concentration estimates persisted in the final estimates for $PM_{2.5}$ and O_3. The exposure model was trained by leaving out 10% of EPA air quality monitors. But because these monitors are generally located in areas with high population density, it is possible that the model is prone to larger error in areas with lower population density — which generally have lower $PM_{2.5}$ concentrations and therefore are of greater interest in the context of this study. And, based on earlier work by the researchers that provides the basis for the exposure models used in these studies, it appears that the model may systematically underpredict concentrations for unexplained reasons. The nature, sources, size, and potential impact of the potential errors discussed here are important to understand and deserve attention in future analyses.

Long-Term Health Effects, the Cohort Study: Using the massive database of all Medicare recipients during 2000 to 2012, and combining it with the equally large exposure predictions, Dominici and colleagues have performed a study with extraordinary statistical power to investigate the association between all-cause mortality and long-term exposure to a range of $PM_{2.5}$ and O_3 levels. That they observed an association between annual average concentrations and mortality at higher concentrations was not the new finding of this research, but the findings at low levels, particularly at levels below the current NAAQS, are novel and potentially important.

The greatest challenge to the internal validity of this study, as for all observational studies, is the potential for confounding, which can bias the results. To address such concerns, the investigators performed numerous analyses with some 20 covariates. They also utilized findings from a smaller Medicare cohort that had a much richer set of potential confounding variables to assess the likely impact of having only a limited number of covariates in the main cohort analysis. In addition, to allow for the effects of time-dependent covariates known to vary from year to year, they utilized a variant of the classic Cox proportional hazards model, the Andersen-Gill formulation.

However, this is a complex study. Health and personal characteristics are available for individuals, but ambient air pollutant exposure is estimated at the ZIP code level (averaged from the 1 km × 1 km spatial scale of the prediction model). Additionally, the ZIP code scale is the smallest spatial unit at which individual residential and other covariate information is available. These factors, coupled with confounders that can act at the level of the individual, the community, or the regional environment, result in a complex hybrid model. These issues pose important challenges for the next phase of the work planned by the investigators, and the causal inference methods under development will need to focus on these challenges.

Based on the current results, the Panel offers the following comments most relevant to the cohort analyses.

The investigators performed various analyses to explore the potential impact of confounding; however, the Panel noted several areas with a potential for residual confounding in the cohort study. For example, some results from the subgroup analyses are puzzling, particularly the dramatically higher effect of $PM_{2.5}$ exposure in African Americans and the negative (protective) effects of exposure to O_3 for Native Americans, Hispanics, and Asians.

Although the investigators have used the Andersen-Gill formulation to better model time-dependent variables, the Panel's biggest concern relates to the problem of potential for temporal confounding, with both overall nonaccidental mortality and $PM_{2.5}$ levels declining steadily over the period of the study, 2000–2012. Because this is an open cohort (new individuals enter the cohort as they enroll for Medicare), age — which is controlled in the analyses — is not necessarily strongly correlated with calendar time. As a result, confounding could occur because of the contributions of both age and calendar time. The Panel believes that without accounting for confounding by time, the findings of the long-term exposure study should be viewed with caution.

The Panel also has concerns about the impact of the likely exposure misclassification and confounding related to the hybrid nature of the study, but appreciates that exposure measurement error correction methodology for spatially varying pollutants and methods to address confounding in such a complex study setting are still in their infancy. Additionally, the Panel notes that data on individual health-related behaviors, such as smoking, diet, and exercise, do not capture the full extent of variability in the behaviors, such as geographical variability. Finally, the presence of other pollutants — such as NO_2 — may also confound the associations between $PM_{2.5}$ and O_3 and mortality.

Another important issue in interpretation of these results is related to the very large population studied here, and consequently the very high apparent precision of the results (i.e., the very small confidence intervals). Because the impact of bias and model misspecification is not reflected in standard uncertainty measures, one should be cautious about over-interpreting the narrow confidence intervals. The Panel's comments and concerns about the potential impacts of bias and of unmeasured confounding should be viewed in this broader context.

Short-Term Health Effects, the Case-Crossover Study: The second study in this report uses a case–crossover design — a variant of the time-series design — to evaluate short-term effects of low-level air pollution in the Medicare population. One advantage this study design has over the long-term design is that it is based on variation in exposure and mortality experienced by an individual over short periods of time (days, rather than years). Therefore, only confounding factors that vary over short periods of time, such as weather, are of potential concern, rather than the much larger array of potential confounders that either do not vary with time or have long-term trends. On the other hand, by design, time-series analyses only address the immediate impact of air pollution on mortality rather than the pollutants' role in the development of chronic morbidity and subsequent mortality.

Dominici and colleagues report a relative risk increase of 1.05% and 0.51% in daily mortality rate for each 10-μg/m^3 increase in $PM_{2.5}$ and 10-ppb increase in O_3, respectively. The concentration–response analyses for $PM_{2.5}$ and O_3 suggest a nonlinear relationship, with a steeper slope at low concentrations and flattening at higher concentrations. They have also investigated effect modifications for a range of variables. For example, they report that the mortality effect of short-term exposure to $PM_{2.5}$ is greater in women than in men, in contrast to the finding in the cohort study. The effects in other subgroup analyses were generally not significant, except for Medicaid eligibility. Also, NO_2 — another time-varying covariate — was not included in these analyses.

Causal Modeling: There is increasing interest in research on casual inference methods because of the challenges in accounting for confounding in the preceding analyses of observational data, and Dominici and colleagues are devoting significant effort to the development and extension of two such methods.

In the first method, the investigators have developed a generalized-propensity-score approach for confounding adjustment along with a regression calibration method to address exposure measurement error in health models. In the second approach, they have developed a new Bayesian causal modeling approach, known as *local exposure–response confounding adjustment,* to estimate exposure–response curves accounting for differential effects of confounders at different levels of exposure. Both of these approaches serve as potentially useful starting points, and the Panel notes that current applications do not address the concerns raised about the long-term and short-term studies — in particular, concerns about residual confounding and impacts of the complex hybrid nature of the study designs — and so it looks forward to the full development and applications of these methods to the health analyses.

Sharing of Models and Data: Dominici and colleagues have made a special effort to make available their data, workflows, and analyses, and have posted these at a secure high-performance computing cluster with the objective of developing an open science research data platform. Additionally, the codes and software tools are publicly available from another depository.

The investigators' work in these areas will continue. The Panel finds these efforts praiseworthy and encourages the Dominici team to continue sharing the unique resources they have developed.

CONCLUSIONS

Using very large air pollution model and health data sets, Dominici and colleagues have reported initial results using two types of analysis — a cohort analysis of long-term exposures and a case–crossover analysis of short-term exposures. They have found positive associations of both $PM_{2.5}$ and O_3 with all-cause mortality, with associations extending to concentrations below the current NAAQS and with little evidence of a threshold. The investigators also conducted a range of sensitivity analyses and controlled for many confounders; these did not meaningfully change the initial findings of associations. These initial analyses are thorough and comprehensive, and make a valuable contribution to the literature.

As extensive as these analyses are, as noted by the Panel and by the investigators, there are several key questions that need to be investigated further before firmer conclusions can be drawn. Particularly important among these are (1) issues around the potential for confounding by time trends and other variables, including other pollutants such as NO_2, and geographical patterns in exposure and health status; (2) impact of the different spatial scales of the variables in both the long-term and short-term analyses, and the resulting complex quasi-ecological (hybrid) nature of the models, with the potential for exposure misclassification and residual confounding; and (3) extension of their work by the development, testing, and application of causal inference methods in the full study population.

Dominici and colleagues have performed a set of extensive and creative analyses in the largest air pollution and health databases to date. While initial conclusions may be drawn from these first analyses, the Panel will wait for the planned extensive further analyses to be completed before reaching full conclusions on the air pollution and public health implications of this important research.

HEALTH
EFFECTS
INSTITUTE

75 Federal Street, Suite 1400
Boston, MA 02110, USA
+1-617-488-2300 phone
+1-617-488-2335 fax
www.healtheffects.org

Recycled Paper