HEALTH EFFECTS INSTITUTE

STATEMENT

Synopsis of Research Report 183, Part 3

Modeling of Multipollutant Profiles and Spatially Varying Health Effects

INTRODUCTION

Although it is clear that people are exposed to complex mixtures of pollutants emitted by diverse sources of air pollution, air quality standards worldwide are geared toward control of individual, or small sets of, pollutants. Consequently, most epidemiologic studies of air pollution and health to date have focused on estimating the adverse effects associated with ambient exposure in single-pollutant models. Employing multipollutant models using conventional statistical approaches frequently produces results that are difficult to interpret because air pollutant levels are often highly correlated. Therefore, advanced statistical methods are needed to investigate the health effects of air pollution mixtures.

HEI issued Request for Applications (RFA) 09-1, "Methods to Investigate the Effects of Multiple Air Pollution Constituents," to solicit research proposals that would address these methodologic challenges through the development of innovative statistical methods. Three studies were funded under RFA 09-1 that represent a variety of statistical approaches and applications. The studies by Dr. Brent Coull and Dr. Eun Sug Park and their colleagues are described in Research Report 183, Parts 1 and 2. For the current study, Dr. John Molitor and colleagues proposed to develop and apply statistical methods to examine associations between spatial patterns of correlated air pollutants and outcomes of health and poverty.

APPROACH

The investigators built on their previous work to develop Bayesian clustering methods to identify spatial clusters of air pollution exposures — and of other covariates such as socioeconomic status — and to estimate the association of health outcomes with those clusters. They use the term *profile* to define a set of pollutants (or more generally exposures). Their approach has three components: a prior for cluster allocation, a profile assignment submodel to assign exposure profiles to clusters, and a health effects submodel to link clusters of exposure profiles to the health outcome. The Bayesian models described by Molitor and colleagues are mostly fit using Markov chain Monte Carlo techniques. Their Bayesian framework allows a supervised (joint) estimation (meaning that they allowed the relationship between health outcomes and exposures to inform the formation of the clusters).

An important feature of these clustering methods is that they are flexible. For example, the number

What This Study Adds

- Advanced statistical methods are needed to investigate health effects of air pollution mixtures. Molitor and colleagues extended their cluster methods to include continuous exposures and successfully implemented them to analyze multipollutant mixtures.
- Their approach was aimed at identifying spatial clusters of air pollution exposures — and other covariates such as socio-economic status — and estimating health outcomes associated with those clusters. The approach is flexible, for example, the number of clusters does not need to be predefined, and uncertainty related to cluster allocation is accounted for.
- Future work is necessary to fully evaluate the methods, including simulation studies, comparison to traditional statistical methods, application in other settings, and inclusion of more pollutants.

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. John Molitor at the College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, and colleagues. The complete report, *Development of Statistical Methods for Multipollutant Research, Part 3: Modeling of Multipollutant Profiles and Spatially Varying Health Effects with Applications to Indicators of Adverse Birth Outcomes* (© 2016 Health Effects Institute), can be obtained from HEI or our Web site (see next page).

of clusters does not need to be predefined. In addition, these methods quantify the uncertainty related to the clustering allocation and propagate it in the health analyses. To group exposure profiles into clusters, Molitor and colleagues used Dirichlet-process mixture modeling techniques and combined the resulting clusters with multilevel regression models to estimate health outcomes. Subsequently, they developed post-processing Bayesian model-averaging techniques to find clusters that best represent the data and to assess uncertainty in the cluster allocation.

The investigators conducted analyses using three applications to demonstrate these methods on measures of poverty and adverse birth outcomes in Los Angeles County using census and birth certificate data.

A maximum of four pollutants were considered, including $PM_{2.5}$ and NO_2.

MAIN RESULTS AND INTERPRETATION

In its independent review of the study, the HEI Review Committee concluded that the investigators extended their cluster methods to include continuous exposures and successfully implemented these methods to analyze multipollutant mixtures. Their analyses demonstrated that their approach can be applied to real-world data sets and that they produced results that were largely concordant with a priori expectations. Results indicate that the effects of pollutants, as well as socioeconomic status variables, vary spatially and that they vary in a complex interconnected manner. The Committee thought the difficult subject matter was made much more accessible through the investigators' approach to presenting their results. For example, the Committee liked the spatially-varying maps of effects, which they believe are a useful and effective tool to communicate results.

The Committee appreciated the flexibility of the clustering approach. The explicit inclusion of spatially-varying contextual factors (e.g., socioeconomic status variables) as inputs to the clusters, in a way similar to the treatment of air pollutants, was considered unique and can potentially provide new insight into understanding vulnerable and susceptible populations.

The methods developed by Molitor and colleagues are complex. The investigators have put their models in a unified Bayesian framework as one way to allow a supervised (joint) estimation. In general, there are several important practical features of supervised modeling approaches that are worth considering. For example, there is a potential for feedback due to unbalanced data and misspecification of the models. The clusters identified are dependent on the health outcome, and changing the health outcome will generally change the definition of the clusters to some extent.

In addition, typically, they are computationally demanding.

The Committee noted that effects of the various data simplifications were not studied, such as the aggregation of exposure from the individual to the census tract or census block group level.

Finally, the Committee thought it would have been worthwhile to understand how the methods perform under known conditions and to compare the methods to traditional statistical methods for which the research community has already developed a deep understanding of their properties and performance.

CONCLUSIONS

Dr. Molitor and colleagues developed methods to address an important question in multipollutant research, that is, what are the combined effects of various constituents of an air pollution mixture. The Committee concluded that the multipollutant methods developed show promise, but that the full extent to which they will be useful remains to be seen. Future work is necessary to fully evaluate these methods, including simulation studies, comparison to traditional statistical methods, application in other settings, and inclusion of more pollutants. Such analyses could help to determine the degree to which these new methods will lead to a better understanding of how pollutant mixtures contribute to health effects, and ultimately, to better decisions about how to control them.

HEALTH EFFECTS INSTITUTE

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

pubs@healtheffects.org
www.healtheffects.org

Recycled Paper