HEALTH EFFECTS INSTITUTE

STATEMENT

Synopsis of Research Report 202

Dispersion and Bayesian Models of Traffic-Related Air Pollutants

INTRODUCTION

Traffic emissions are an important source of urban air pollution, and exposure to traffic-related air pollution has been associated with various adverse health effects. However, exposure assessment is challenging because traffic-related air pollution is a complex mixture of particles and gases that varies greatly by location and over time. This variability complicates the development of accurate models of traffic-related air pollution to assess exposure to air pollutants for epidemiological studies, in particular because of small-scale variations within cities. Dr. Stuart Batterman from the University of Michigan and his team aimed to improve estimates of traffic-related air pollution concentrations for use in health studies. They used a systematic approach to apply and test a dispersion model — RLINE — developed by the United States Environmental Protection Agency (U.S. EPA) and novel statistical approaches (called "Bayesian spatiotemporal data fusion models" by the investigators) that combine measurements with concentration estimates generated by RLINE. The long-term goal was to apply and improve existing models that could then be employed in other settings.

APPROACH

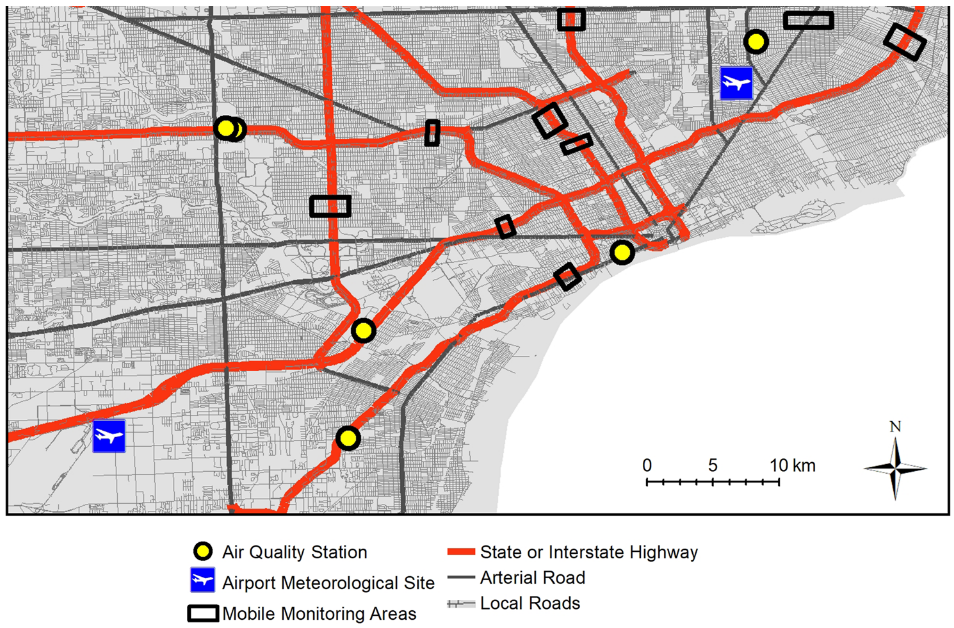
The study used data collected as part of NEXUS (Near-road EXposures and effects of Urban air pollution Study), a large study conducted in Detroit to evaluate health effects of air pollution in children with asthma living near major roads. All air pollution data were previously collected for the NEXUS study at central and near-road monitoring sites in 2011–2014 or by measuring concentrations at different distances from roads with a mobile monitoring platform during one week in December 2012 (see Statement Figure).

The investigators employed models of varying computational complexity — RLINE plus five different statistical methods — for particulate matter ≤ 2.5 µm in aerodynamic diameter ($PM_{2.5}$), nitrogen oxides (NO_x), carbon monoxide (CO), and black carbon (BC). RLINE was designed to model concentrations of air pollutants by including factors such as traffic volume, meteorology, and other factors that influence how those pollutants spread after being emitted by motor vehicles. First, the investigators evaluated the RLINE model by predicting daily average ambient NO_x, CO, and $PM_{2.5}$

What This Study Adds

- The investigators undertook to improve the estimation of air pollution exposure from traffic by applying and testing different statistical models of concentrations near major roads.
- Specifically, they evaluated whether inclusion of predictions from the RLINE model of traffic-related air pollution would improve sophisticated statistical models for potential use in exposure assessment.
- Each model provided different useful information, and inclusion of RLINE improved predictions of the increase in near-road concentrations of $PM_{2.5}$, but not of NO_x, relative to background levels.
- The application of the statistical models was an important contribution. However, the usefulness and generalizability of these models remain limited until they have been evaluated with long-term measurements.

This Statement, prepared by the Health Effects Institute, summarizes a research project funded by HEI and conducted by Dr. Stuart Batterman at the University of Michigan, Ann Arbor, and colleagues. The complete report, *Enhancing Models and Measurements of Traffic-Related Air Pollutants for Health Studies Using Dispersion Modeling and Bayesian Data Fusion* (©2020 Health Effects Institute), can be obtained from HEI or our website (*see next page*).



Statement Figure. Map of Detroit showing air quality monitoring stations, airport weather stations, and near-road mobile monitoring locations. (The map is based on Figure 1 of the Investigators' Report and Figure 6 in the Additional Materials, with background layers from Michigan GIS Open Data.)

concentrations at five U.S. EPA monitoring sites in the Detroit area. Second, the investigators systematically applied and evaluated the performance of a series of increasingly complex statistical models by including factors such as day of week, upwind versus downwind of the nearest major road, and traffic activity. They also developed a model that predicted both $PM_{2.5}$ and NO_x concentrations together instead of in separate models.

MAIN RESULTS AND INTERPRETATION

In its independent review of the study, the HEI Review Committee concluded that Batterman and colleagues had successfully evaluated the performance of the RLINE model, as well as the performance of universal kriging and sophisticated statistical models that combined RLINE output with measurements. The Committee agreed with the investigators that both RLINE and measurements contributed useful information to the concentration predictions from statistical models. The performance of the RLINE model depended on the pollutant as well as on spatial and temporal factors, such as distance from the nearest major road. In addition, statistical models with different sets of assumptions generally led to the same conclusions and provided complementary information on how the air pollutants were spatially distributed. Finally, adding RLINE to the statistical models or jointly modeling NO_x and $PM_{2.5}$ improved predictions only for $PM_{2.5}$ and not for NO_x.

The Committee thought the statistical models were state of the science and well executed and that the application of the statistical models was a novel and important contribution. They appreciated that the models were systematically compared using a number of performance statistics. On the other hand, the Committee thought that the report may have overstated the usefulness of the models for epidemiological studies for several reasons. First, the models appeared to have limited use over a broad geographic area. Second, the models performed better closer to roads than farther away, which might translate to biased health effect estimates because the exposure predictions would be more accurate for the most highly exposed people in an epidemiological cohort (with participants living at varying distances from major roads). In addition, the uncertainties in the predictions of air pollutant concentrations remained large, even for the most refined models.

There remains a need to further refine the models and distribute these new tools for wider use. In particular, these and similar models will need to be rigorously tested on large databases of measurements collected over long periods before they are used on a large scale in epidemiological studies.

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