ORIGINAL PAPER

Methodological considerations in developing local-scale health impact assessments: balancing national, regional, and local data

Bryan J. Hubbell · Neal Fann · Jonathan I. Levy

Abstract National-scale health impact assessments (HIAs) have been conducted for many years and have become reasonably systematized. Recently, there has been growing interest in utilizing HIA methods at local scales, in the context of Environmental Public Health Tracking and in other settings. This paper investigates the data and analytical challenges to estimating the incidence of health effects associated with changes in air pollution concentrations at the local scale, focusing on ozone and fine particulate matter. Although it could be argued that the local-scale HIA is simply a more geographically discrete version of the national-scale assessment and, therefore, has similar challenges, in practice, many key inputs in national-scale assessments are assumed to be spatially uniform or vary only at coarse geographic resolution. For a national-scale assessment, this assumption may not contribute appreciable bias, but the bias could be significant for any individual location. Thus, local-scale assessments require more geographically resolved air quality data, concentration–response (C-R) functions, and baseline incidence rates than are often used. However, comprehensive local data may not be available, may be incomplete, or may be time-intensive and resource-intensive to develop, especially for C-R functions for which small-scale epidemiological studies will often be underpowered. Given this context, this paper considers how best to develop credible local-scale HIAs, identifying factors that contribute to variability across geographic areas, study designs, and time periods. This paper also describes which key sources of analytical uncertainty change as the scope shifts from the national to the local scale. These challenges notwithstanding, the paper concludes that a well-designed local-scale HIA, following key principles and recommendations, can be both informative and defensible.

Keywords Health impact assessment · Baseline incidence rate · BenMAP · Concentration–response function · Health impact function · Air quality modeling

B. J. Hubbell · N. Fann (✉)
Office of Air Quality Planning and Standards,
U.S. Environmental Protection Agency,
Research Triangle Park, NC 27711, USA
e-mail: Fann.neal@epa.gov

J. I. Levy
Department of Environmental Health,
Harvard School of Public Health,
Boston, MA, USA

Introduction

In the United States (U.S.), regulatory agencies such as the U.S. Environmental Protection Agency (EPA) have an extensive history of performing national-scale health benefits assessments of major air pollution regulations. In one such analysis, the U.S. EPA estimated that implementation of the Clean Air Act amendments would result in approximately 23,000 premature deaths avoided in 2010, relative to a baseline of Clean Air Act implementation without the amendments (U.S. EPA 1999). More recently, EPA has systematized this approach by developing the Environmental Benefits Mapping and Analysis Program (BenMAP) for estimating national-scale health benefits of policies and regulations to reduce air pollution. BenMAP was used in recent Regulatory Impact Analyses (RIAs) of the Clean Air Interstate Rule and Nonroad Diesel Rule (U.S. EPA 2004, 2005), finding that, when fully implemented, those two rules combined would result in close to 30,000 premature deaths from air pollution avoided annually.

The general approach utilized within these analyses has been found to be reasonable and informative for policy decision-making in spite of inherent uncertainties (NRC 2002).

In recent years, interest has grown among both regulatory agencies and researchers in developing similar types of assessments at subnational or local scales, focusing on either the health impacts of pollutants or the health benefits of control strategies. For example, while the RIA of the Clean Air Interstate Rule (U.S. EPA 2005) estimated air quality changes at 36 km resolution and provided health benefits estimates only at the national scale, other studies have used atmospheric models with census tract resolution and provided health benefits estimates for individual tracts or as a function of distance from individual sources (Levy and Spengler 2002; Levy et al. 2002). In addition, BenMAP analyses have been conducted to address a variety of risk management questions in Philadelphia, Detroit, California, Georgia, and for the Lake Michigan Air Directors Consortium (Chang et al. 2007; Cohan et al. 2007; Clean Air Council 2004; Deck 2006; U.S. EPA 2007a).

A local-scale health impact assessment (HIA; considered henceforth to be urban-scale) applies essentially identical methods as a national-scale HIA, but is distinguished by the heightened importance of spatially refined input data. In principle, a national-scale assessment would rely on the same spatially varying input data as used in a local-scale assessment. However, in practice, many key inputs are assumed to be spatially uniform or to vary only across regions of the country. For a national-scale assessment focused on total national impacts, this may not contribute appreciable bias, but the bias may be larger for individual urban areas.

As a result, there are multiple tensions and tradeoffs that analysts must weigh in attempting to construct defensible local-scale HIAs. In general, as the spatial scale decreases, national or "generic" data may become less representative. At the same time, local data may not be available or may be more uncertain given smaller sample sizes. In addition, there is a tension between developing assessments that are locally meaningful and using methods that are consistent with other HIAs, so that results can be compared across settings. Finally, if HIAs are to be conducted on a regular basis, the HIA methodology should provide a mechanism for collection and use of regularly updated data and concentration–response (C-R) functions.

This paper explores each of these issues in the context of ozone and fine particulate matter ($PM_{2.5}$) HIAs in the U.S. Historically, these are the two pollutants that dominate the health benefits of air pollution control strategies. As a result, ozone and $PM_{2.5}$ were selected within the Environmental Public Health Tracking (EPHT) program framework as key pollutants to evaluate within a local-scale HIA context (CDC 2007). In the "Overview of health impact assessment methods" section, we provide a brief conceptual overview of HIA methods. In the "Analytical challenges to performing a local-scale HIA" section, we discuss critical data issues that must be addressed when conducting a local-scale HIA. In the "Consideration of time-varying factors in conducting local-scale health impact assessments over time" section, we discuss the need to consider the influence of time-varying factors in conducting local-scale HIAs over time. In the "Characterizing uncertainty in local-scale health impact assessments" section, we discuss characterization of uncertainty, including aspects that are amplified at the local scale. Finally, in the "Conclusions and recommendations" section, we conclude with some recommendations for the conduct of local-scale HIAs.

## Overview of health impact assessment methods

The key elements of health impact assessments (HIAs) are illustrated in Fig. 1, and include:

1. Estimate a change in or increment of ambient air quality, using monitoring data (from ground-based or satellite measurements), modeled air quality, or a combination of the two. These air quality changes may be actual or hypothetical.
2. Combine air quality changes with population information to determine changes in population exposure in a form that is relevant given the epidemiological evidence (e.g., the appropriate averaging time).
3. Combine changes in population exposure to ambient air pollution with impact functions[1] to generate distributions of changes in the incidence of health effects. The impact functions are constructed using population data, baseline health effect incidence and prevalence rates, and C-R functions.
4. Characterize the results of the HIA, through the use of summary statistics (e.g., mean, 95% confidence interval), graphs (e.g., cumulative distribution functions and box-plots), and maps.

Health impact functions estimate the change in a health endpoint of interest, such as hospital admissions, for a

---

[1] The term "impact function" as used here refers to the combination of (a) a C-R function obtained from the epidemiological literature, (b) the baseline incidence for the health effect of interest in the modeled population, and (c) the size of that modeled population. The impact function is distinct from the C-R function, which strictly refers to the estimated equation from the epidemiological study relating the relative risk of the health effect and ambient pollution. We refer to the specific value of the relative risk or estimated coefficients in the epidemiological study as the "C-R function." In referencing the functions used to generate changes in incidence of health effects for this paper, we use the term "impact function."

Fig. 1 HIA—analytical
framework
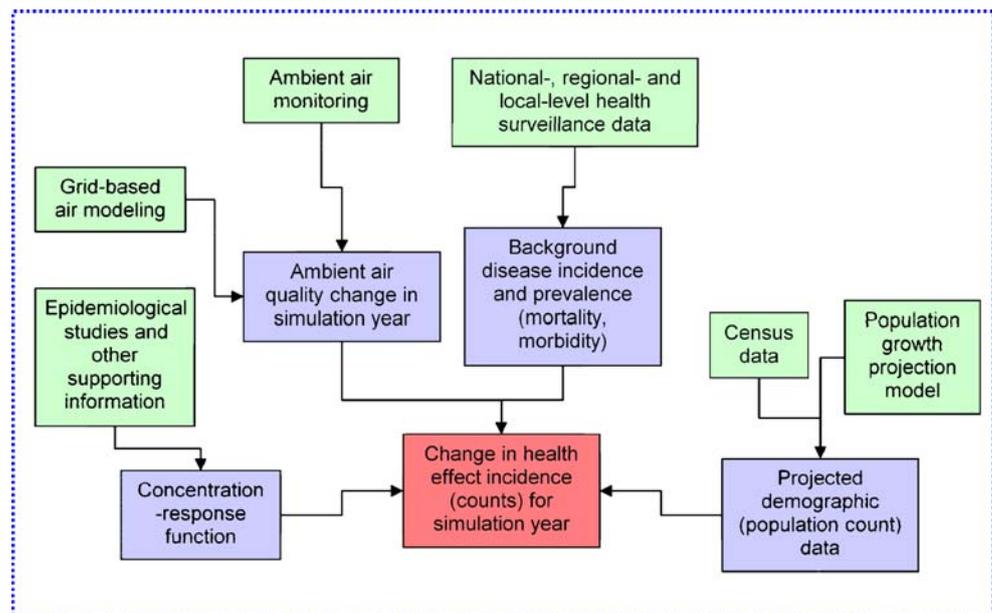
given change in air pollution concentrations (here, consid-
ered to be either ambient ozone or PM$_{2.5}$). A typical health
impact function might look like:

$$\Delta y = y_0 \times \left(e^{\beta \times \Delta x} - 1\right)$$

where y_0 is the baseline incidence (the product of the
baseline incidence rate times the potentially affected
population), β is the C-R function, and Δx is the estimated
change in ambient concentrations. There are other func-
tional forms, but the basic elements remain the same.

Identifying health outcomes of interest

An important initial step in an HIA (whether national-scale
or local-scale) is to consider which health outcomes to
include in the analysis. Several types of data can support
this determination, including toxicological studies (animal
and cellular studies), human clinical studies, and observa-
tional epidemiology studies. All of these data sources
provide important contributions to the weight of evidence
regarding the biological plausibility of a particular health
outcome; however, only epidemiological studies provide
direct C-R functions which can be used to evaluate
population-level impacts of reductions in ambient levels
of criteria air pollutants such as ozone and PM$_{2.5}$. The
selection of a health outcome, therefore, follows a weight of
evidence approach, based on the biological plausibility of
effects, availability of C-R functions from well-conducted
peer-reviewed epidemiological studies, cohesiveness of
results across studies, and a focus on endpoints reflecting
public health impacts (like hospital admissions) rather than
physiological responses (such as forced expiratory volume

in 1 s [FEV1]). Table 1 lists some of the health endpoints
included in recent HIAs for ozone and PM$_{2.5}$.

The quantitative aspect of HIA, therefore, relies on the
outputs from environmental epidemiology. A downside is
that standard epidemiological studies provide only a limited
representation of the uncertainty associated with a specific
C-R function, measuring only the statistical error in the
estimates, usually relating to the power of the underlying
study (driven largely by population size and the frequency
of the outcome measure). There are other sources of
uncertainty in the relationships between ambient pollution
and population-level health outcomes, including model
specification, potential confounding by factors that are both
correlated with the health outcome and each other, and
other factors such as exposure misclassification. Other
study designs may provide insight about these issues but

Table 1 Health endpoints included in recent U.S. EPA national-scale
HIAs (U.S. EPA 2005, 2006)

Health endpoint	PM$_{2.5}$	Ozone
Mortality	✓	✓
Chronic bronchitis	✓	
Nonfatal heart attacks	✓	
Hospital admissions	✓	✓
Asthma ER visits	✓	✓
Minor restricted activity days	✓	✓
Asthma attacks	✓	✓
Work loss days	✓	
Worker productivity		✓
School absence rates		✓

are difficult to capture quantitatively. For example, expert elicitation methods have been used to integrate across various sources of data in developing C-R functions for RIAs, a topic discussed in greater detail in the "Characterizing uncertainty in local-scale health impact assessments" section.

Selecting C-R functions

For health outcomes for which multiple studies are available, criteria need to be developed to determine which studies are likely to provide the best C-R functions given the context of the HIA, as well as how the estimates from these studies should be weighted and combined. These criteria may include consideration of whether the study was peer-reviewed, the match between the pollutant studied and the pollutant of interest, whether potential co-pollutant confounding was addressed, the study design and location, and characteristics of the study population, among other considerations. To account for the potential impacts of different health care systems or underlying health status of populations, it may make sense to give preference to studies conducted in the country of interest, although the relative importance of country-specific characteristics may vary by health outcome (with health care utilization measures likely depending more on country characteristics than disease development).

Table 2 is a summary table from the ozone National Ambient Air Quality Standards (NAAQS) RIA (U.S. EPA

2007b) which demonstrates some of the issues confronted in selecting and applying C-R functions in a national-scale analysis. First, EPA relied on multicity studies or meta-analyses of the mortality literature rather than selecting epidemiological studies conducted in individual cities, but did not attempt to formally combine the evidence across these studies. The individual row estimates for mortality, therefore, reflect the variability in the C-R functions for ozone mortality across these studies. Ranges within each column reflect the uncertainty in the ozone C-R functions as well as in the estimates of $PM_{2.5}$ premature mortality impacts across the available C-R functions for $PM_{2.5}$ mortality. Similar characterization could be conducted for morbidity outcomes as well, although the existence of fewer studies makes it more difficult to encapsulate uncertainty as fully.

Estimating baseline incidence

Given the C-R functions for various health outcomes, the remainder of the impact function depends on the baseline incidence rate and the size of the relevant exposed population. The way in which the health outcome is defined should be in agreement with the epidemiological studies underlying the C-R function. Some epidemiological studies examine the association between pollution levels and adverse health effects in a specific subpopulation, such as asthmatics or diabetics. In these cases, it is necessary to

Table 2 Summary of 2020 national health impacts associated with attainment strategies for alternative ozone NAAQS (U.S. EPA 2007a)

Standard alternative and model or assumption		Combined range of ozone benefits and $PM_{2.5}$ co-benefits			
		0.079ppm	0.075ppm	0.070ppm	0.065ppm
Combined estimate of mortality					
NMMAPS	Bell et al. (2004)	200 to 1,900	430 to 2,600	670 to 4,300	1,200 to 7,400
	Bell et al. (2005)	260 to 2,000	1,100 to 3,300	1,500 to 5,100	2,800 to 9,000
Meta-analysis	Ito et al. (2005)	270 to 2,000	1,200 to 3,300	1,600 to 5,200	3,000 to 9,200
	Levy et al. (2005)	260 to 2,000	1,300 3,500	1,800 to 5,400	3,000 to 9,200
No causality		180 to 1,900	230 to 2,400	390 to 4,000	660 to 6,900
Combined estimate of morbidity					
Acute myocardial infarction		1,100	1,400	2,300	4,000
Hospital and ER visits		1,300	5,600	7,600	13,000
Chronic bronchitis		370	470	780	1,300
Acute bronchitis		950	1,200	2,000	3,500
Asthma exacerbation		7,300	9,400	16,000	27,000
Lower respiratory symptoms		8,100	10,000	17,000	29,000
Upper respiratory symptoms		5,900	7,500	13,000	22,000
School loss days		50,000	610,000	780,000	1,300,000
Work loss days		51,000	65,000	110,000	190,000
Minor restricted activity days		430,000	2,000,000	2,700,000	4,700,000

develop not only baseline incidence rates, but also prevalence rates for the defining condition (e.g., asthma). For both baseline incidence and prevalence data, age-specific rates are preferred where available. Impact functions are applied to individual age groups and then summed over the relevant age range to provide an estimate of total population benefits.

Estimating health benefits

Health benefits are calculated by linking the impact function and the modeled changes in air pollution. The change in or increment of air quality is generally determined by either a policy scenario (e.g., implementation of SO_2 emission controls at power plants) or a specific standard or target (e.g., reduce $PM_{2.5}$ levels to the annual average standard of $15 \mu g/m^3$). For policy measures for which secondary pollutant formation (e.g., ozone, sulfate, nitrate) and regional-scale fate and transport are relevant, air quality changes are generally predicted using sophisticated air quality models such as EPA's Community Multiscale Air Quality model. For primary pollutants at smaller spatial scales, Gaussian plume models such as the Atmospheric Dispersion Modeling System or more simplified monitor rollback methods (Bell et al. 2005) can be used. The appropriate methodology, scale, and resolution of air quality assessment will depend on the problem context.

Analytical challenges to performing a local-scale HIA

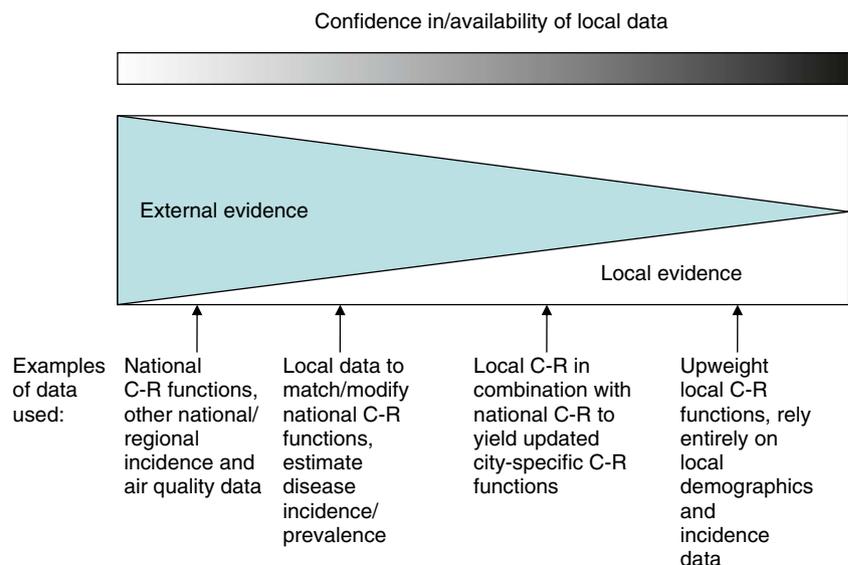Although national-scale HIAs in principle include local-scale impacts, in practice, these assessments are oriented around aggregate national impacts and are relatively less concerned with robust characterization of subnational impacts. Use of nationally representative impact functions could result in large errors for any given location, which would tend to cancel out in an aggregate assessment. Thus, the sources of analytical uncertainty change as the scope shifts from national-scale impacts to local-scale impacts. The discussion below focuses on each facet of the HIA, describes the tension between national-scale and local-scale data, and discusses how the role of local data in the analysis is affected by the shift in geographic scale.

In general, the tension is captured in the continuum proposed within Fig. 2—in some cases, either limited local data will be available or the data will be less reliable than similar data developed at the national level (either because of statistical power issues or because of weakness in the study design). In those cases, the evidence used in national-scale HIA may need to be directly applied at the local scale, with minor modifications and an acknowledgement of how the uncertainties change. Increasing availability of local data may provide refined incidence/prevalence data or insight about factors that could influence the C-R functions. At the far end of the continuum, sufficient local data may be available to provide site-specific epidemiological and other local data. More detail is provided in the discussion below for each input.

Development of appropriate C-R functions

In most cases, local-scale HIAs will need to rely on "off-the-shelf" information available from the epidemiological literature, given either a lack of local epidemiological studies or the likelihood that such studies would not have

Fig. 2 Continuum for use of national versus local data in local-scale HIAs. Note that a separate continuum would be applicable for each health outcome

Confidence in/availability of local data

External evidence

Local evidence

| Examples of data used: | National C-R functions, other national/ regional incidence and air quality data | Local data to match/modify national C-R functions, estimate disease incidence/ prevalence | Local C-R in combination with national C-R to yield updated city-specific C-R functions | Upweight local C-R functions, rely entirely on local demographics and incidence data |

adequate statistical power to allow the global literature to be ignored or downweighted and preference given to the findings of local studies. The critical question is, therefore, whether C-R functions are transferable from the contexts in which they were generated to a specific local-scale HIA, given the various factors that could lead these functions to differ by location or over time.

Although the nature of these variations will differ by outcome and study design, the key sources of potential heterogeneity in C-R functions between cities can be divided into two general categories: those attributable to differences in exposure and those attributable to differences in potential susceptibility. Because most epidemiological studies relate health outcomes to central-site monitored concentrations rather than personal exposure, factors that affect the relationship between personal exposure and ambient concentrations can potentially affect the C-R function. Some of these exposure-related factors include air conditioning prevalence and utilization, availability and effectiveness of air quality alerts, and amounts of time spent outdoors or in traffic. Monitor siting characteristics can also affect the magnitude and statistical significance of estimated C-R functions as well as their interpretation and transferability. For example, elemental carbon displays greater spatial heterogeneity within a city than sulfates, with potential differential effects on the degree of measurement error resident in central-site monitors used to represent community exposures to ambient $PM_{2.5}$. As a result, cities with relatively greater contributions from secondary pollutants such as sulfate may display different C-R functions strictly due to monitor placement. More generally, the pollutant mixture could influence the underlying C-R function, related to the relative levels and temporal profiles of ambient pollution, the chemical composition of $PM_{2.5}$, or the size distribution of $PM_{2.5}$.

Susceptibility, as measured by differences in demographic factors and baseline health across cities or across time, can affect estimated C-R functions. The literature identifies key factors such as age, prevalence of heart and lung disease, education, income/poverty, access to health care, the nature of the health care systems, and asthma prevalence. Some recent evidence also suggests that it is important to account for public health interventions that might modify the impacts of air pollution on health. For example, a recent study by Schwartz et al. (2005) found that the introduction of statin drugs reduces the relative risk from $PM_{2.5}$. The widespread and increasing use of statin drugs in the population may then affect the observed C-R functions for $PM_{2.5}$ over time. Careful comparisons of those exposure or susceptibility factors that might change the relationship between ambient pollutant concentrations and health outcomes should be conducted prior to selecting C-R functions for a particular location, as this will help determine whether evidence is applicable directly, applicable with modification, or inapplicable.

In general, when conducting a local-scale HIA, there may be data from a sufficient number of cities or studies to be able to identify factors that explain variability in C-R functions across cities, or the literature may be inadequate to do so. Assuming a sufficient number of candidate C-R functions, meta-analysis and pooling techniques can be used to develop estimates that reflect potential heterogeneity across cities. Standard meta-analysis or pooling approaches involve weighting candidate studies by (for example) the inverse of their reported variance, providing a central estimate across the literature. An alternative to this approach uses random rather than fixed effects, which allows the possibility that the estimates from the different studies may in fact be estimates of different parameters, rather than just different estimates of a single underlying parameter. While these simple pooling approaches can provide better mean C-R functions for use in national-scale HIAs, they can lead to biased results when applied to local areas because some of the variability in C-R functions between locations may be due to systematic factors influencing exposure, susceptibility, or other factors. In this case, more sophisticated meta-analysis approaches may be required.

In recent years, several articles have been published that involve either meta-analyses or new multicity studies for ozone (Stieb et al. 2002; Bell et al. 2004, 2005; Ito et al. 2005; Levy et al. 2005) and $PM_{2.5}$ mortality (Levy et al. 2000; Stieb et al. 2002; Dominici et al. 2003; Franklin et al. 2007). In addition to developing an overall mean C-R function, some of these studies attempted to determine whether the C-R functions vary as a function of co-pollutant concentrations, temperature, air conditioning prevalence, and other factors. Among other findings, locations with higher air conditioning prevalence appear to have a smaller effect from ozone (Levy et al. 2005) and $PM_{2.5}$ (Franklin et al. 2007). The implication of this for developing C-R functions for specific locations is that national mean estimates may need to be adjusted to account for local factors, although it should be recognized that the covariates in metaregressions may not necessarily be the causal factors driving the C-R functions.

In general, multicity analyses have some analytical advantages over multistudy meta-analyses, as they impose a consistent model specification, use the same time period for each city included, and can be more inclusive of locations, resulting in less chance of any publication bias. However, if the C-R functions are strongly influenced by model specification, single studies may be more likely to be biased than multistudy meta-analyses that may draw on multiple model specifications. Regardless, either type of study can provide city-specific estimates using hierarchical

Bayes models in which individual C-R functions for each city represent priors for those cities, but the posterior estimates represent a weighted average between those observations and the results from a pooling process or metaregression. If a city-specific estimate is highly uncertain and there is either little heterogeneity in the city-specific estimates or such heterogeneity can be explained by defined characteristics of the cities, then less weight is given to the city-specific observation. In contrast, if there is significant unexplained heterogeneity and the city-specific estimates have good statistical power, then the city-specific estimates are given more weight. This approach recognizes that C-R functions from cities other than the one being evaluated within the local-scale HIA can provide insight about the appropriate estimate for the city of interest.

For many exposures and outcomes, the literature will not be sufficient to conduct formal metaregressions. Given the likelihood that neither extreme on the continuum in Fig. 2 (no relevant local data or substantial and well-powered local epidemiology) will occur, the process for determining appropriate C-R functions for local-scale HIA in this context requires careful development of profiles of characteristics of the city of interest and study locations to find the closest match along a range of attributes that can impact C-R functions. Profiles can be generated using available databases on air quality composition (obtained from the EPA Air Quality System—http://www.epa.gov/ttn/airs/airsaqs/ or the HEI Air Quality Database—http://hei.sf.aer.com/login.php), baseline health status (using numerous sources from CDC), demographics (using databases from the U.S. Census Bureau), and other factors such as climate and meteorological variability. Formal matching analyses can be conducted (e.g., clustering of cities based on health and air quality attributes), or less formal approaches based on expert analytical judgment can be used. There is no single rule of thumb on how close areas need to be in attributes space— to some extent, this will depend on how much uncertainty is acceptable in the analysis. It will also depend on the attribute. If an attribute has been shown to have a large impact on C-R functions, then the focus should be on matching that attribute as closely as possible. In cases where the information base is more limited, matching may be inapplicable; in which case, the analyst should broadly consider the degree of uncertainty and/or bias associated with the application of the available C-R functions.

It should be noted that the aforementioned approach is most relevant to time-series studies, which focus on day-to-day variations in pollution within cities. However, another form of study is often used to examine health outcomes associated with chronic exposures. These studies use the variation in long-term pollution concentrations between cities to estimate the C-R function. The best of these studies use prospective cohort designs, which track the survival rates (or disease-free status rates) for individuals over time, and calculate relative risks from pollution controlling for differences in individual-level factors such as smoking status, diet, etc. Because these studies use between-city variability to generate the C-R function, city-specific C-R functions are not available. Thus, when applying the C-R functions from cohort studies, the same estimate should be used in each location, unless there is evidence of non-linearities in the function. While there is some recent evidence (e.g., Jerrett et al. 2005) that C-R functions for cohort mortality are larger when using within-city pollution gradients, the overall evidence is mixed, and the most robust findings are currently based on between-city comparisons of $PM_{2.5}$ and standardized mortality risks.

However, care should be taken to identify any systematic differences in exposure or susceptibility between the populations used in the cohort study and the populations in the local-scale HIA. For example, it has been recognized that the population studied in one of the most widely used cohort studies, the American Cancer Society (ACS) Study (Pope et al. 2002), is not representative of the demographic mix in the general population. The ACS cohort is almost entirely white and has higher income and education levels relative to the general population. In EPA's recent expert elicitation study, many of the experts suggested that these sample characteristics led to a downward bias in the estimated C-R functions relative to a C-R function that would be representative of the general U.S. population (Roman et al. 2008). In spite of this concern, most previous HIA in the U.S. (U.S. EPA 2005, 2006; Levy and Spengler 2002; Levy et al. 2002) have used estimates from the Harvard Six Cities (Laden et al. 2006) and ACS (Pope et al. 2002) studies of $PM_{2.5}$-related mortality, given concerns that other published cohort studies are less representative of national or local populations. For example, the Washington University-EPRI Veterans Cohort Study (Lipfert et al. 2006) involved male hypertensive veterans receiving treatment at VA clinics, with 57% current smokers (versus 24% in the general population). Another cohort study (McDonnell et al. 2000) focused only on nonsmoking Seventh-Day Adventists in California. In either case, the populations differed in demographics, risk factors, and disease status in ways that could significantly impact the C-R functions. As it is unlikely that any general population HIA will exclusively capture such targeted populations, the Six Cities and ACS study estimates will be more relevant for local-scale HIA.

Baseline incidence/prevalence data

Unlike with C-R functions, it is likely that local baseline incidence/prevalence data could be available for at least some health outcomes, putting this step of the local-scale

HIA further along the continuum in Fig. 2. In addition, baseline incidence/prevalence may vary to a greater extent across settings than C-R functions, implying that utilizing baseline incidence data that are not specific to a given location or that are not adequately geographically resolved can introduce important uncertainties to the analysis. For example, ZIP code-level asthma hospitalization rates vary substantially within Detroit, ranging from a minimum of 10 to a maximum of 129 per 10,000 (Fig. 3). In contrast, EPA has used a single estimate of 28 in 10,000 in its national-scale RIAs (U.S. EPA 2005, 2006).

The implications of using spatially resolved local baseline incidence rates become clear when estimating the changes in asthma-related hospital admissions resulting from a hypothetical 20% reduction in $PM_{2.5}$ levels in Detroit (Fig. 3). Using the EPA default baseline hospitalization rate generates a total reduction in asthma-related hospitalizations of 36 cases (90%CI=17, 54). In contrast, using the local-scale rates produces a reduction in asthma-related hospitalizations of 53 cases (90%CI=26, 81). Clearly, using default baseline incidence rates would underestimate the total change in this particular health endpoint in Detroit and would not capture the spatial and demographic variability in that endpoint.

The chief impediment to using high-resolution baseline incidence data is that it is very resource-intensive to produce or may not be available for the outcome of interest with sufficient coverage within a given urban area. For the Detroit example, while the Michigan Department of Environmental Quality (Wasilevich 2005) furnished U.S. EPA with estimates of baseline incidence rates for many key $PM_{2.5}$-related health endpoints, such as nonfatal heart attacks and chronic bronchitis, small ZIP code-level population sizes triggered data suppression rules, resulting in omitted incidence estimates for a number of locations within the city. Moreover, while the department maintains a comprehensive asthma epidemiology database which provides good spatial coverage of Detroit, it was still necessary for an epidemiologist to reformat these data to generate tables in a format suitable for use in an HIA. The EPHT program may serve a valuable role by providing technical guidance to localities as they consider collecting such data.

Given these constraints, local-scale baseline incidence rates for each health endpoint of concern will not always be available. However, rather than relying solely on national or broad regional estimates, it may be possible to apply interpolation or other estimation techniques to infer the baseline rates for the city—or perhaps for certain locations
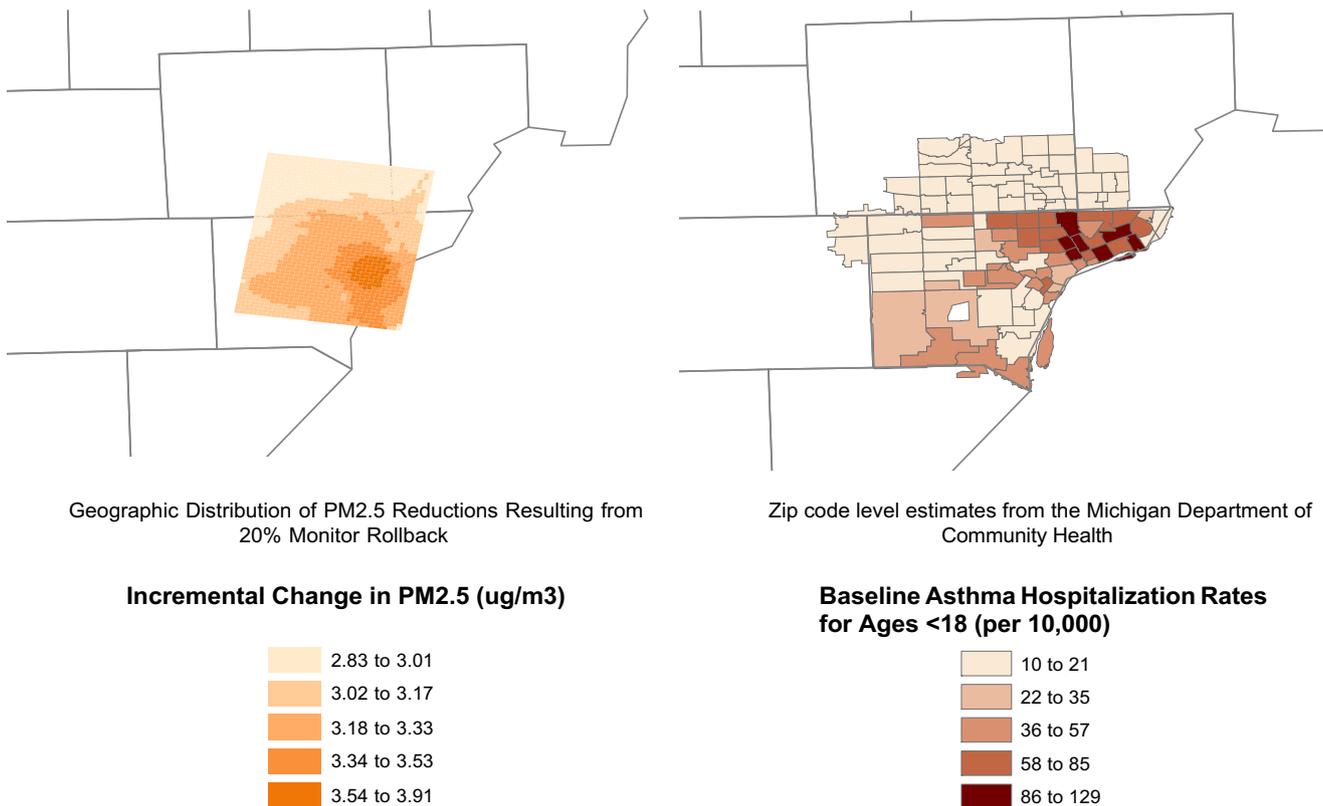

Geographic Distribution of PM2.5 Reductions Resulting from 20% Monitor Rollback

Incremental Change in PM2.5 (ug/m3)

- 2.83 to 3.01
- 3.02 to 3.17
- 3.18 to 3.33
- 3.34 to 3.53
- 3.54 to 3.91

Zip code level estimates from the Michigan Department of Community Health

Baseline Asthma Hospitalization Rates for Ages <18 (per 10,000)

- 10 to 21
- 22 to 35
- 36 to 57
- 58 to 85
- 86 to 129

Fig. 3 Comparison of the geographic distribution of ZIP code-level asthma hospitalization rates and a hypothetical 20% reduction in monitored $PM_{2.5}$ in the Detroit metropolitan area

within a city—of interest. For example, if incidence rates correlate well with some number of easily observed independent variables—perhaps age, race, and geographic region—then one could estimate the baseline incidence rate. This approach has been followed previously, as baseline incidence data were simulated at high resolution as a function of demographic factors (Levy et al. 2002). A limitation is the fact that this relies heavily on the assumption that factors such as race and education are universally applicable causal factors rather than covariates reflecting complex contextual relationships within the underlying studies from which the correlations were developed. Geostatistical interpolation techniques, such as kriging and co-kriging, may also be of some use in creating a spatial surface of interpolated baseline incidence rates. The uncertainty these methods introduce, though significant, is likely outweighed by the improved representation of geographic heterogeneity inherent in baseline incidence rates.

Air quality data

Changes in air quality—monitored or modeled—ultimately drive estimates of health impacts. A key analytical challenge is to represent both the spatial distribution and scale of these air quality changes, an issue of particular significance for local-scale HIA, for which more finely resolved estimates of air quality changes may take on added importance. For pollutants such as directly emitted $PM_{2.5}$, we would expect a high degree of variability in the geographic distribution of air quality changes across urban areas, given certain source controls. For example, evidence suggests that directly emitted particles from motor vehicles (Roorda-Knape et al. 1998; Zhu et al. 2002) or point sources such as metal processors and glass manufacturers (U.S. EPA 2006) have a pronounced spatial gradient as a function of distance from the source. However, for HIA, the impact of these emissions on total population exposure is more relevant than the concentration profile itself, and some studies have shown a much larger spatial extent of impact even for primary $PM_{2.5}$ when considered from a total population exposure perspective (Greco et al. 2007). Local-scale HIA for $PM_{2.5}$ and ozone are further complicated by the importance of secondary formation, which implies that sources outside of the geographic area of interest will influence local air quality and that sources inside the area of interest will influence public health at a regional or national scale.

Without knowing the policy context for the local-scale HIA, it is difficult to know the extent to which monitors or models can be used or the necessary scale and resolution. One general statement that can be made is that the estimated changes in air quality must be reasonably in agreement with the way in which exposure was characterized within the epidemiological studies used to derive C-R functions. If the C-R function is derived from a single central-site monitor, highly spatially resolved exposure characterization would be more difficult to interpret within an HIA. That being said, finer-scale air quality data will clearly take on added importance for a local-scale HIA in which it may be of interest to align inputs such as the baseline incidence rates and populations with the spatial air quality gradient.

Comparing locally developed health impact estimates with literature-based estimates

In many cases, even if analysts conduct a local-scale HIA using locally generated baseline incidence and prevalence rates, C-R functions, and air quality estimates, it may be useful to generate health impacts using literature-based estimates as well as a way of putting the local data-based estimates into context. In comparing these estimates, it will be important to understand what differences might be expected versus differences that cannot be expected or explained. In addition to evaluating differences between attributes of locations, analysts should also be aware of the timeframe in which a national analysis was conducted. Some population or air quality characteristics may have changed significantly between the time when a study was conducted and the present, as discussed in more detail below.

In addition to local area attributes, analysts should also compare the statistical methodologies used in generating the local area estimates relative to those used in generating the literature-based estimates. For example, if different functional forms are used for a local epidemiological study and the published literature (e.g., 2-day moving average versus distributed lags), then differences in the results would be expected, solely from the statistical methods used. Similarly, data collection and aggregation methods may differ for baseline incidence/prevalence data.

Consideration of time-varying factors in conducting local-scale health impact assessments over time

As noted above, population and air quality attributes will likely change over time, making interpretation of the outputs of local-scale HIAs over time challenging. From an accountability perspective, if these changes are not accounted for, then the "signal" from programs intended to reduced air pollution-related health risks can be masked or overstated. For example, if the age composition of a population is becoming older over time, but air quality and other factors are unchanged, the population health risks

will appear to increase over time simply because the "at risk" susceptible proportion of the population is increasing. While these demographic changes should clearly be incorporated, the resulting HIAs could be misleading if not interpreted carefully.

Some time-varying factors that should be considered when designing a multiple year assessment program include demographics, exposure modifiers, air pollution sources, pollutant composition, and meteorology/climate. Demographic factors may include age composition, race, educational levels, income and income disparity, and population health characteristics such as rates of obesity, asthma, and heart disease. This includes covariates that could explain between-city variability in C-R functions or baseline incidence rates, so the process of determining the appropriate impact functions for a local-scale HIA can help to elucidate the most significant time-varying factors to consider.

Changes over time in the composition of air pollution could influence local atmospheric chemistry, leading to increased or decreased formation of secondary $PM_{2.5}$ or ozone for a given amount of emissions controls and could potentially influence the toxicity of the pollutant mixture. These changes may occur both due to pollution control programs and due to natural economic factors such as plant closures. Finally, as climate changes over time, both the susceptibility of populations to air pollution and the nature of air pollution events may change. Higher temperatures may increase susceptibility to air pollution-related health effects across the U.S., changing C-R functions (Roberts 2004). Several recent studies have found an increased risk of air pollution stagnation events under projected changes in the global and regional climate, which are expected to decrease cyclone frequencies throughout the U.S. (Leung and Gustafson 2005; Mickley et al. 2005; Wu et al. 2007).

Forward-looking design of HIA protocols can help to avoid comparability problems by including explicit consideration of time-varying factors. With the proper study design, as conditions change, the changes in effects expected from air pollution reductions can be isolated from changes in effects due to other time-varying factors.

Characterizing uncertainty in local-scale health impact assessments

An important component of any HIA is a characterization of the uncertainties associated with estimates of health impacts. While techniques exist to provide probabilistic estimates of impacts, those techniques are limited by a lack of input data on uncertainty in individual impact function elements. For some elements, such as the C-R function, there is at least some limited information on uncertainty, in

the form of standard errors from the statistical estimation, while for others, e.g., baseline incidence rates, there is no information available. Bayesian approaches, such as hierarchical Bayes analyses used in multicity epidemiological studies, can provide additional characterization of uncertainty because they partially account for heterogeneity between cities, thus accounting for some of the uncertainty about transferability. However, even these approaches cannot address overall model uncertainty or uncertainty about causality.

As alluded to earlier, one approach that has been utilized in a limited number of settings involves formal expert elicitation. The U.S. EPA recently conducted an expert elicitation to try to provide a more complete assessment of the uncertainties associated with the C-R function for $PM_{2.5}$-related mortality (Roman et al. 2008). Expert elicitation is useful in integrating the many sources of information about uncertainty in the C-R function because it allows experts to synthesize these data sources using their own mental models and provide a probabilistic representation of their synthesis of the data in the form of a probability distribution of the C-R function. The goal of the study was to elicit from a sample of health experts probabilistic distributions describing uncertainty in estimates of the reduction in mortality among the adult U.S. population resulting from reductions in ambient annual average $PM_{2.5}$ levels. Expert elicitation methods are still somewhat controversial, and care should be taken to follow formal expert elicitation protocols. It is also unlikely that expert elicitations at this scale would be conducted for an individual local-scale HIA, or even for many of the key elements, given the cost, time, and lack of empirical data from which experts could derive informed opinions. However, results from available expert elicitations can provide a more robust understanding of uncertainties in different analytical components and may eventually provide insight about some of the transferability concerns central in local-scale HIA.

There are some types of uncertainties that are especially important when conducting a local-scale HIA using national-scale data or transferring data from a different location. These uncertainties relate to the factors identified in the sections above on choosing C-R functions. These uncertainties can be minimized (but not eliminated) by careful selection of C-R functions based on cities that are similar to the city under analysis. Other sources of uncertainty at the local scale include uncertainty in the application of regional-average or national-average baseline incidence rates and uncertainty in changes in air quality at fine spatial resolution or for specific populations. These uncertainties can be minimized by using as spatially refined estimates as available. While many of these uncertainties may be difficult to quantify, sensitivity analyses can be

conducted to examine individual assumptions or combinations of assumptions. When communicating the results of sensitivity analyses, care should be taken to indicate the potential likelihood of the set of assumptions to avoid giving an artificially distorted impression of the likely range of impacts.

Conclusions and recommendations

Properly conducted local-scale HIAs can be informative and defensible. However, there are many opportunities for biases and uncertainties to be introduced into the analytical process and the findings. Careful attention to the inputs to the analysis can help to minimize uncertainties and reduce the potential for biased results. In addition, comparison with results from other locations or with national-scale results can provide context for the results and a check on the reasonableness of estimates.

Some key recommendations for conducting interpretable local-scale HIAs include:

- Clearly identify characteristics of study locations that may influence health impact estimates, including major emission sources, air quality composition, population demographics (age, race, income, etc.), and climate/meteorology. When applying previously published epidemiological studies, pay attention to statistical designs, including functional forms, controls for weather and other confounders, lag structures, treatment of missing values, and timeframes for each study from which a C-R function is derived.

- Where available, use city-specific estimates derived from multicity studies, which ideally use Bayesian methods to leverage findings from all cities. In these cases, use meta-analytic techniques to formally evaluate factors contributing to between-city variability in C-R functions and to determine appropriate local-scale estimates. If there are no multicity studies or sufficient studies for a meta-analysis, compare characteristics in the local area to the characteristics of the source study locations and select C-R functions from studies or cities with characteristics closely matching those of the local area to the extent possible.

- Choose a spatial resolution for estimation of changes in air quality appropriate given the exposure assessment within the epidemiological studies of interest, the spatial gradient in air pollution given the pollutant and sources being controlled within the assessment, and potential for correlation with population demographics. In cases where there is a sharp gradient in air pollution that is highly correlated with spatial gradients in population density or susceptibility characteristics, high-resolution data should

be used in spite of the attendant uncertainties and potential issues in combining these data with epidemiological evidence based on central-site monitors.

- Baseline health outcome data should be as spatially and demographically refined as possible. When local baseline health data are not available, analysts should consider using prediction or interpolation methods to derive local baseline rates using national or regional estimates coupled with locally available data shown to be correlated with the health outcomes of interest.

- Design multiyear HIA studies with accountability and comparability in mind. Designs should incorporate time-varying factors, including baseline population health (which may be influenced by the availability of certain types of medical interventions like statin drug use), socioeconomic factors, climate (e.g., mean summer temperatures), sources of emissions (which can affect the composition of local air pollution), and availability of detailed air quality forecasts, that can influence the relationship between air pollution and health.

- Uncertainty should be discussed and characterized quantitatively where possible. Probabilistic approaches can be used to characterize some uncertainties, but should be accompanied by single-attribute and multi-attribute sensitivity analyses to address uncertainties that do not have good probabilistic data available.

References

Bell ML, McDermott A, Zegler SL et al (2004) Ozone and short-term mortality in 95 US urban communities, 1987–2000. JAMA 292:2372–2378. doi:10.1001/jama.292.19.2372

Bell ML, Dominici F, Samet JM et al (2005) A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. Epidemiology 16:436–445. doi:10.1097/01.ede.0000165817.40152.85

Centers for Disease Control and Prevention (2007). National Environmental Public Health Tracking Network Technical Network Implementation Plan (TNIP) V 1.0. In: CDC. Available at http://www.cdc.gov/nceh/tracking/pdfs/TNIP_V1.pdf. Accessed 17 February 2009

Chang E, Tsiopulos LT, Whynot J (2007) Final Socioeconomic Report for the 2007 AQMP. In: South Coast Air Quality Management District. Available at http://www.aqmd.gov/aqmp/07aqmp/07AQMP_socio.html Accessed 17 February 2009

Clean Air Council (2004) Philly Particulate Matter Analysis. In: Philadelphia Diesel Difference. Available at http://www.cleanair.org/dieseldifference/resources/reports/pmphillyreport.html Accessed 17 February 2009

Cohan D, Boylan JW, Marmur A et al (2007) An integrated framework for multipollutant air quality management and its application in Georgia. Environ Manage 40:545–554. doi:10.1007/s00267-006-0228-4

Deck L (2006) Benefit Study of MRPO Candidate Control Options for Electricity Generation. Stratus Consulting Inc. In: Lake Michigan Air Directors Consortium. Available at http://64.27.125.175/reports/rpo/Regional%20Air%20Quality/EGU%20Benefit-Cost%20analysis%20Stratus.pdf. Accessed 17 February 2009

Dominici F, McDermott A, Zegler SL et al (2003) National maps of the effects of particulate matter on mortality: exploring geographic variation. Environ Health Perspect 111:39–43

Franklin M, Zeka A, Schwartz J (2007) Association between $PM_{2.5}$ and all-cause and specific-cause mortality in 27 US communities. J Expo Sci Environ Epidemiol 17:279–287. doi:10.1038/sj.jes.7500530

Greco SL, Wilson AM, Spengler JD et al (2007) Spatial patterns of mobile source particulate matter emissions-to-exposure relationships across the United States. Atmos Environ 41:1011–1025. doi:10.1016/j.atmosenv.2006.09.025

Ito K, De Leon SF, Lippmann M (2005) Associations between ozone and daily mortality: analysis and meta-analysis. Epidemiology 16:446–457. doi:10.1097/01.ede.0000165821.90114.7f

Jerrett M, Burnett RT, Ma R et al (2005) Spatial analysis of air pollution and mortality in 35 Los Angeles. Epidemiology 16:727–736. doi:10.1097/01.ede.0000181630.15826.7d

Laden F, Schwartz J, Speizer FE et al (2006) Reduction in fine particulate air pollution and mortality. Am J Respir Crit Care Med 173:667–672. doi:10.1164/rccm.200503-443OC

Leung R, Gustafson W (2005) Potential regional climate change and implications to U.S. air quality. Geophys Res Lett 32:L16711. doi:10.1029/2005GL022911

Levy JI, Spengler JD (2002) Modeling the benefits of power plant emission controls in Massachusetts. J Air Waste Manage Assoc 52:5–18

Levy JI, Chemerynski SM, Sarnat JA (2005) Ozone exposure and mortality: an empiric Bayes metaregression analysis. Epidemiology 16:458–468. doi:10.1097/01.ede.0000165820.08301.b3

Levy JI, Hammitt JK, Spengler JD (2000) Estimating the mortality impacts of particulate matter: what can be learned from between-study variability? Environ Health Perspect 108:109–117. doi:10.2307/3454508

Levy JI, Greco SL, Spengler JD (2002) The importance of population susceptibility for air pollution risk assessment: a case study of power plants near Washington, DC. Environ Health Perspect 110:1253–1260

Lipfert FW, Wyzga RE, Baty JD et al (2006) Traffic density as a surrogate measure of environmental exposures in studies of air pollution health effects: long-term mortality in a cohort of US Veterans. Atmos Environ 40:154–169. doi:10.1016/j.atmosenv.2005.09.027

McDonnell WF, Nishino-Ishikawa N, Petersen FF et al (2000) Relationships of mortality with the fine and coarse fractions of long-term ambient PM10 concentrations in nonsmokers. J Expo Anal Environ Epidemiol 10:427–436. doi:10.1038/sj.jea.7500095

Mickley LJ, Jacob DJ, Field BD et al (2005) Impact of climate change on air pollution episodes in the United States. U.S. Climate Change Science Program, Climate Science in Support of Decisionmaking Workshop, 14–16 November, Arlington, VA

National Research Council (NRC) (2002) Estimating the public health benefits of proposed Air Pollution Regulations. National Academies Press, Washington, DC

Pope CAIII, Burnett RT, Thun MJ et al (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287:1132–1141. doi:10.1001/jama.287.9.1132

Roberts S (2004) Interactions between particulate air pollution and temperature in air pollution mortality time series studies. Environ Res 96:328–337. doi:10.1016/j.envres.2004.01.015

Roman H, Walker KD, Walsh TL et al (2008) An Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S. Environ Sci Technol 42:2268–2274

Roorda-Knape MC, Janssen NAH, de Hartog JJ et al (1998) Air pollution from traffic in city districts near major motorways. Atmos Environ 32:1921–1930. doi:10.1016/S1352-2310(97)00496-2

Schwartz J, Park S, O'Neill MS, Vokonas PS et al (2005) Glutathione-S-transferase M1, obesity, statins, and autonomic effects of particles gene-by-drug-by-environment interaction. Am J Respir Crit Care Med 172:1529–1533. doi:10.1164/rccm.200412-1698OC

Stieb DM, Judek S, Burnett RT (2002) Meta-analysis of time-series studies of air pollution and mortality: effects of gases and particles and the influence of cause of death, age, and season. J Air Waste Manage Assoc 52:470–484

U.S. Environmental Protection Agency (U.S. EPA) (1999) The benefits and costs of the Clean Air Act, 1990–2010. Prepared for U.S. Congress by U.S. EPA, Office of Air and Radiation/Office of Policy Analysis and Review, Washington, DC, November; EPA report no. EPA-410-R-99-001. Available at http://www.epa.gov/air/sect812/

U.S. Environmental Protection Agency (U.S. EPA) (2004) Final regulatory analysis: control of emissions from nonroad diesel engines. EPA420-R-04-007. Prepared by Office of Air and Radiation. Available at http://www.epa.gov/nonroad-diesel/2004fr/420r04007.pdf

U.S. Environmental Protection Agency (U.S. EPA) (2005) Regulatory impact analysis for the Clean Air Interstate Rule. EPA 452/-03-001. Prepared by Office of Air and Radiation. Available at http://www.epa.gov/cair/pdfs/finaltech08.pdf

U.S. Environmental Protection Agency (U.S. EPA) (2006) Regulatory impact analysis of the 2006 National Ambient Air Quality Standards for Particle Pollution. Office of Air Quality Planning and Standards, October

U.S. Environmental Protection Agency (U.S. EPA) (2007) Regulatory impact analysis of the proposed revisions to the National Ambient Air Quality Standards for Ground-Level Ozone. Office of Air Quality Planning and Standards. EPA-452/R-07-008

U.S. Environmental Protection Agency (U.S. EPA) (2007) Detroit MP Study. In: Clean Air Info. Available at http://www.cleanairinfo.com/regionalstatelocalmodelingworkshop/old/2007/presentations/Thursday%20-%20May%2017%202007/Air_Toxics_Update_Detroit_Study.pdf Accessed 17 February 2009

Wasilevich W (2005) Epidemiology of asthma in Wayne County. In: Asthma Institute of Michigan, Michigan Department of Community Health. Available at http://www.getasthmahelp.org/Wayne%20Fact%20Sheet%202005.pdf Accessed 15 December 2007

Wu S, Mickley LJ, Leibensperger EM, Jacob DJ et al (2007) Effects of 2000–2050 global change on ozone air quality in the United States. J Geophys Res 113:D06302 Available online at: http://www-as.harvard.edu/chemistry/trop/publications/wu2007/GCAP_AQ.pdf

Zhu YF, Hinds WC, Kim S et al (2002) Study of ultrafine particles near a major highway with heavy-duty diesel traffic. Atmos Environ 36:4323–4335. doi:10.1016/S1352-2310(02)00354-0