An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure


http://dx.doi.org/10.1289/ehp.1307049

Received: 6 May 2013
Accepted: 7 February 2014
Advance Publication: 11 February 2014
An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure

Richard T. Burnett,1 C. Arden Pope III,2 Majid Ezzati,3 Casey Olives,4 Stephen S. Lim,5 Sumi Mehta,6 Hwashin H. Shin,1 Gitanjali Singh,7 Bryan Hubbell,8 Michael Brauer,9 H. Ross Anderson,10 Kirk R. Smith,11 John R. Balmes,11-12 Nigel G. Bruce,13 Haidong Kan,14 Francine Laden,15 Annette Prüss-Ustün,16 Michelle C. Turner,17 Susan M. Gapstur,18 W. Ryan Diver,18 and Aaron Cohen19*

1Health Canada, Ottawa, Ontario, Canada; 2Brigham Young University, Provo, Utah, USA; 3MRC-PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK; 4School of Public Health, University of Washington, Seattle, Washington, USA; 5Institute for Health Metrics and Evaluation, Seattle, Washington, USA; 6Global Alliance for Clean Cookstoves, Washington, DC, USA; 7Harvard School of Public Health, Harvard University, Cambridge, Massachusetts, USA; 8United States Environmental Protection Agency, Research Triangle Park, North Carolina, USA; 9School of Population and Public Health, University of British Colombia, Vancouver, Canada; 10MRC-PHE Centre for Environment and Health, King’s College London, London, UK; 11University of California, Berkeley, Berkeley, California, USA; 12School of Medicine, University of California, San Francisco and Berkeley, California, USA; 13Department of Public Health and Policy, University of Liverpool, Liverpool, UK; 14School of Public Health, Fudan University, Shanghai, China; 15Exposure, Epidemiology, and Risk Program, Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA; 16World Health Organization, Geneva, Switzerland; 17Institute of Population Health, University of Ottawa, Ottawa, Ontario, Canada; 18American Cancer Society, Atlanta, Georgia, USA; 19Health Effects Institute, Boston, Massachusetts, USA; *Senior Author
Address correspondence to Richard T. Burnett, Population Studies Division, Environmental Health Sciences and Research Bureau, Environmental and Radiation Health Sciences Directorate, Healthy Environment and Consumer Safety Branch, Health Canada, Room 134, Environmental Health Center, 50 Columbine Driveway, Ottawa, ON, Canada K1A 0K9. Telephone: 613-952-1364. Fax: 613-941-3883. E-mail: rick.burnett@hc-sc.gc.ca

Short running title: Fine Particulate Mass Based Risk Model

Acknowledgments: We would like to thank the members of the Ambient Air Pollution Expert Group of the Global Burden of Diseases, Injuries, and Risk Factors Study (the GBD 2010 Study) for their helpful comments and discussion of this material and to Michael Lipsett for the relative risk estimate for COPD mortality from the California Teachers Study. Aaron Cohen wishes to acknowledge the support provided for his work by the William and Flora Hewlett Foundation.

Disclaimer: Aaron Cohen is employed by HEI. HEI receives about half of its core funds from the U.S.EPA and half from the worldwide motor vehicle industry, although other public and private organizations periodically support special projects or certain research programs. The views expressed in this article do not represent those of HEI or its sponsors. Annette Prüss-Ustün is a staff member of the World Health Organization. She alone is responsible for the views expressed in this publication and they do not necessarily represent the views, decisions or policies of the World Health Organization. This article should not be reproduced for use in association with the promotion of commercial products, services or any legal entity. WHO does not endorse any specific organization or products. Any reproduction of this article cannot include the use of the WHO logo.

Competing Financial Interests: We declare that we have no competing financial interests.
Abstract

Background: Estimating the burden of disease attributable to long-term exposure to fine particulate matter (PM$_{2.5}$) in ambient air requires knowledge of both the shape and magnitude of the relative risk function (RR). However, there is inadequate direct evidence to identify the shape of the mortality RR functions at high ambient concentrations observed in many places in the world.

Objective: Develop relative risk (RR) functions over entire global exposure range for causes of mortality in adults: ischemic heart disease (IHD), cerebrovascular disease (stroke), chronic obstructive pulmonary disease (COPD), and lung cancer (LC). In addition, develop RR functions for the incidence of acute lower respiratory infection (ALRI) that can be used to estimate mortality and lost-years of healthy life in children less than 5 years old.

Methods: An Integrated Exposure-Response (IER) model was fit by integrating available RR information from studies of ambient air pollution (AAP), second hand tobacco smoke (SHS), household solid cooking fuel (HAP), and active smoking (AS). AS exposures were converted to estimated annual PM$_{2.5}$ exposure equivalents using inhaled doses of particle mass. Population attributable fractions (PAF) were derived for every country based on estimated world-wide ambient PM$_{2.5}$ concentrations.

Results: The IER model was a superior predictor of RR compared to seven other forms previously used in burden assessments. The PAF (%) attributable to AAP exposure varied among countries from: 2-41 for IHD, 1-43 for stroke, < 1-21 for COPD, < 1-25 for LC, and < 1-38 for ALRI.

Conclusions: We developed a fine particulate mass-based RR model that covered the global range of exposure by integrating RR information from different combustion types that generate emissions of particulate matter. The model can be updated as new RR information becomes available.
Introduction

It is now well established that long-term exposure to ambient PM$_{2.5}$ is associated with increased mortality from non-accidental and cause-specific diseases (Brook et al. 2010; COMEAP 2009; Cooke et al. 2007; Krewski et al. 2009). Epidemiologic cohort studies, conducted largely in the US, have reported this association for annual ambient average concentrations from approximately 5 µg/m$^3$ to 30 µg/m$^3$, though definitive knowledge of which specific sources or characteristics of PM$_{2.5}$ are responsible for these associations is currently lacking (US EPA 2009; WHO 2006, 2007). No epidemiologic study, however, has estimated the association of long-term exposure to direct measurements of PM$_{2.5}$ with mortality from chronic cardiovascular and respiratory disease at higher ambient exposures common in cities and other areas in Asia and other developing countries where annual average exposures can exceed 100 µg/m$^3$ (Brauer et al. 2012; Health Effects Institute 2010). As a result, estimates of disease burden attributable to ambient air pollution in these locations have had to extrapolate the results of epidemiologic studies from locations with lower ambient PM$_{2.5}$ exposures (Anenberg et al. 2010; Cohen et al. 2004; Evans et al. 2013).

Previous efforts to estimate global burden from exposure to ambient air pollution (AAP) in the form of PM$_{2.5}$ postulated risk functions for cardiopulmonary mortality as linearly increasing in RR from 7.5 to 50 µg/m$^3$ with no further change in RR at higher concentrations (Cohen et al. 2004). Sensitivity analyses included a model in which RR varied as the logarithm of concentration, producing a more gradual diminution of the marginal increase in RR than the base-case model. The logarithmic model was subsequently recommended by the World Health Organization (WHO) for use in air pollution burden of disease estimates at the national level (Ostro 2004). The coefficients of these models were based on information from a single US cohort study - American Cancer Society Cancer Prevention II (CPS-II) (Krewski et al. 2009; Pope et al. 2002;) with exposure
assignments below 22 µg/m³. The form of the models used for global burden assessment was motivated largely by the concern that linear extrapolation using these coefficients would produce unrealistically large estimates of RR compared to other known PM₂.₅ related mortality risks, such as active smoking (AS) and exposure to second hand tobacco smoke (SHS) (Cohen et al. 2004; Ostro 2004). These RR models were also employed in more recent estimates of global mortality associated with ambient PM₂.₅ concentrations (Anenberg et al. 2012; Evans et al. 2013).

Absent empirical epidemiologic evidence on the magnitude of the association with mortality at high exposures of PM₂.₅ in ambient environments, Pope et al. (2011) suggested that the integration of epidemiologic evidence on cardiovascular (CV) and lung cancer (LC) mortality RR from disparate types of PM₂.₅ exposure such as AAP, SHS, and AS, may provide insight into the shape of the exposure-response relation over a much wider range of exposures.

In this paper we present the methodology used to estimate the population attributable fraction (PAF) from exposure to ambient PM₂.₅ in the Global Burden of Disease 2010 project (Lim et al. 2012). We selected a mathematical form of the RR function with PM₂.₅ concentration that could describe the observed relationships between RR and exposure for the five outcomes examined. We fit this model for cause-specific adult mortality for four causes of death: IHD, stroke, COPD, and LC, using relative risk information from epidemiologic studies of long-term exposure to particulate matter not only from AAP, SHS, and AS, but also from studies of household air pollution from solid cookfuel (HAP). We use these models to estimate the PAF associated with exposure to ambient PM₂.₅ for each of the 187 countries included in the GBD 2010. We identify a specific model form that best predicts source-specific RR for all four causes of death. In addition, we examine the relationship between PM₂.₅ exposure and the incidence of ALRI in infants, another health outcome considered
Methods

Underlying assumptions

The model we propose in this paper is based on the following underlying assumptions:

(1) Exposure to PM$_{2.5}$ from diverse combustion sources is associated with increased mortality from IHD, stroke, COPD, and LC, and increased incidence of ALRI. This assumption is based on systematic review of the available epidemiologic literature conducted by the GBD 2010 Ambient Air Pollution Expert Group as part of the GBD 2010 project (Lim et al. 2012).

(2) The observed RRs from AAS, SHS, HAPs, and AS are a function of PM$_{2.5}$ mass inhaled concentration across all combustion particle sources (Smith 1987). The toxicity of PM$_{2.5}$ is assumed to differ only with regard to inhaled mass (exposure) and not with PM$_{2.5}$ composition. The toxicity of emissions from different combustion sources may well differ, but current knowledge does not allow definitive and quantifiable conclusions regarding their relative toxicity and little is known about international variation in source contributions around the world (Stanek et al. 2011; USEPA 2009; WHO 2006).

(3) The relation between PM$_{2.5}$ exposure and excess mortality RR is not necessarily restricted to a linear function over the range of human exposure to PM$_{2.5}$ from diverse sources (Pope et al. 2009; 2011).
(4) The RR of mortality from chronic disease experienced by people exposed to AAP, SHS, HAP, and AS is a function of long-term, cumulative, exposure quantified in terms of daily average exposure concentration and does not depend on the temporal pattern of exposure (Pope et al. 2010; 2011). This assumption is required since the temporal nature of PM$_{2.5}$ exposure is different for AAS, SHS, HAPs, and AS.

(5) The RR associated with each type of exposure does not depend on the other types of exposure. That is we are assuming no interaction among the different exposure types for any cause of mortality. We are aware of no empirical epidemiologic evidence that tests that assumption, however the direct epidemiologic evidence from the cohort studies we use to estimate the burden attributable to ambient PM$_{2.5}$ shows that active cigarette smokers are also affected adversely by exposure to ambient PM$_{2.5}$, and these studies do not provide support for significant heterogeneity of the relative excess AAP relative risk across smoking categories.

Model Form

We selected a mathematical form of the IER that could describe several patterns in RR thought to be a priori applicable to exposure-response models. We wanted the IER to be able to take shapes similar to models previously used for burden assessment such as linear and log-linear (Cohen et al. 2004) and a power function (Pope et al. 2009, 2011). In addition to these shapes we also required the IER to have a property that it flattens out at high exposures consistent with evidence of the relationship between IHD mortality and smoking intensity (Pope et al. 2009).

The form must equal 1 when PM$_{2.5}$ values are below some concentration that represents a counterfactual low exposure where below this level there is no excess risk. We also desire a model
that increases monotonically with increasing PM$_{2.5}$ exposure concentration and can take the variety of shapes such as near linear, sub-linear, and supra-linear. Our Integrated Exposure-Response (IER) model has the following form:

for $z < z_{cf}$, $RR_{IER}(z) = 1$

for $z \geq z_{cf}$, $RR_{IER}(z) = 1 + \alpha \{1 - \exp[-\gamma (z - z_{cf})^\delta]\}$ \hspace{1cm} [1]

where $z$ is the exposure to PM$_{2.5}$ in $\mu g/m^3$ and $z_{cf}$ is the counterfactual concentration below which we assume there is no additional risk. For very large $z$, $RR_{IER} \sim 1 + \alpha$. We include a power of PM$_{2.5}$, $\delta$, to predict risk over a very large range of concentrations. We note that $RR_{IER} (z_{cf} + 1) \sim 1 + \alpha \gamma$. Thus, $\gamma = (RR_{IER} (z_{cf} + 1) - 1)/(RR_{IER} (\infty) - 1)$ can be interpreted as the ratio of the RR at low to high exposures. We term our model the Integrated-Exposure Response (IER) model since its development requires the integration of exposures to PM$_{2.5}$ from different combustion types (AAP, SHS, HAP, and AS).

In formulating our RR model we rely on information on the RR of mortality at specified PM$_{2.5}$ exposure concentrations from the available literature. Suppose we have a set of RR estimates $\{\hat{\gamma}^{(s)}_1, \ldots, \hat{\gamma}^{(s)}_{K_s}\}_{s = 1, \ldots, S}$ and corresponding confidence intervals based on PM$_{2.5}$ concentrations $\{z^{(s)}_1, \ldots, z^{(s)}_{K_s}\}_{s = 1, \ldots, S}$, for $S$ different types of PM$_{2.5}$ sources, where $K_s$ is the number of RR estimates available from for source type $S$. The unknown parameters ($\alpha$, $\gamma$, $\delta$) are estimated by non-linear regression methods. We weight the RR estimates by the inverse of the variance estimate of the logarithm of the RR in order to reflect the uncertainties in each estimate.

We compared the IER model to seven other models that have been previously suggested for burden assessment. These include a RR model which is linear in exposure throughout the global
concentration range (Lin), a model that is linear up to 30 µg/m$^3$ and constant above 30 µg/m$^3$ (Lin30) a model that is linear up to 50 µg/m$^3$ and constant above 50 µg/m$^3$ (Lin50) and a model which is a function of the logarithm of exposure (Log). These models were used in a previous assessment of global burden of disease due to AAP exposure (Cohen et al. 2004). We also postulated a model in which we added an unknown parameter to concentration in the Log model to allow more flexibility in fitting the type-specific RRs (Log2). The sixth model examined related RR to a power of exposure as proposed by Pope et al. (2009, 2011) with the seventh model equivalent to the IER with $\delta=1$ (Exp). The mathematical forms of the models are given in the Supplemental Material (Sensitivity of RRs and PAFs to Model Form). We then calculated both the Akaike and Bayesian Information Criteria for each of the eight models examined and five health outcomes as measures of goodness-of-fit.

The method of constructing uncertainty bounds on model predictions is described in detail in the Supplemental Material (Characterizing Uncertainty). In summary, we simulated 1000 sets of source type-specific RRs based on their point estimates and confidence intervals and fit the IER model to these simulated values obtaining 1000 sets of parameter estimates of $(\alpha, \gamma, \delta)$. Using these parameter estimates we then generated 1000 IER functions over the global concentration range. Estimates of uncertainty were also generated for the PM$_{2.5}$ concentrations. Uncertainty in the PAFs is a function of the uncertainty in the IER model predictions and the exposure estimates and determined by simulation methods as described in the Supplementary Materials (Characterizing Uncertainty).

Specifics of the selection of source type-specific RR and PM$_{2.5}$ exposure for each type are described below for the four mortality outcomes. The logarithm of the RR per µg/m$^3$, its standard error, and associated PM$_{2.5}$ concentration for the five outcomes is given in the Supplemental Material, Table S1 by type of PM$_{2.5}$.
**Ambient Air Pollution**

To fit the risk models we used cause-specific mortality AAP RR estimates from available published cohort studies. We evaluated each RR estimate at its study-specific PM$_{2.5}$ mean concentration minus a less-polluted counterfactual level (Lim et al. 2012). Most RRs were obtained from published reports but in some cases new analyses were conducted for this project. These estimates are identified in Supplemental Material, Table S1. We had eight studies reporting RR estimates for IHD mortality, five for stroke mortality, three for COPD mortality, and four for LC mortality.

**Second Hand Smoke**

We selected RRs for both IHD (8 studies reporting separate estimates for males and females) and LC (46 studies) mortality from studies included in the US Surgeon General’s Report (SGR) (2006). We associated the relative risk of death due to SHS exposure with an equivalent ambient PM$_{2.5}$ concentration of 20 $\mu$g/m$^3$ for low-moderate SHS exposure and 50 $\mu$g/m$^3$ for moderate-high exposure based on the analysis of Pope et al. (2009) for IHD mortality since RRs were reported by the SGR (2006) for these two descriptive exposure categories. We assigned a concentration of 35 $\mu$g/m$^3$ based on the mid-point of the range 20-50 $\mu$g/m$^3$ for LC mortality since no specific description of the level of SHS exposure was provided by SGR (2006). We selected 29 RRs from studies examined by Oono et al. (2011) for stroke mortality based on prospective cohort studies with an associated PM$_{2.5}$ concentration of 35 $\mu$g/m$^3$. There was insufficient evidence to estimate a RR due to SHS exposure for COPD mortality. We assume that the SHS RRs are associated with a change in PM$_{2.5}$ exposure based on non-smoking subjects living with a smoker compared to those not living with a smoker. We have not incorporated other potential sources of PM$_{2.5}$ exposure for these subjects, such as from indoor sources, near-roadway conditions, or occupational exposures by subject.
Active Smoking

Following Pope et al. (2009; 2011) we estimated the RR of each of the four causes of death for current cigarettes smoked per day compared to never smokers from the CPS-II. We estimated the RR and 95% confidence intervals associated with 10 cigarette/day groupings of 1-3, 4-7, 8-12, 13-17, 18-22, 23-27, 28-32, 33-37, 38-42, and > 42 cigarettes/day. We estimated that smoking a single cigarette was equivalent to breathing a daily ambient concentration of PM$_{2.5}$ of 667 $\mu$g/m$^3$ assuming an average breathing rate of 18m$^3$/day and inhaled dose of 12,000 $\mu$g PM$_{2.5}$ mass per cigarette (Pope et al. 2009). We then estimated the equivalent ambient concentration of PM$_{2.5}$ by multiplying the average cigarettes/day smoked in each interval by 667 $\mu$g/m$^3$. The shape of the curve fitted by Pope et al. (2009; 2011) was not sensitive to the estimate of equivalent ambient PM$_{2.5}$ concentrations for AS.

Household Air Pollution

Smith et al. (2014) conducted a meta-analysis of studies examining COPD and LC incidence rates among men and women exposed to air pollution from burning coal or biomass for cooking. There were no studies relating IHD or stroke mortality or incidence to HAPs at the time of the GBD 2010 analyses, and thus this PM$_{2.5}$ type cannot contribute to the fit of our RR function. The equivalent long-term PM$_{2.5}$ exposure from HAPs was estimated for subjects using coal or biomass for cooking and those using cleaner fuels for the purposes of integrating this information into our IER risk model. PM$_{2.5}$ exposure estimates for women (300 $\mu$g/m$^3$) were higher than for men (200 $\mu$g/m$^3$). For COPD meta-analysis the relevant female control group was assumed to be using a mixture of gas and chimney stoves (100 $\mu$g/m$^3$). The PM$_{2.5}$ exposure for males was estimated to be 65% of that for females (65 $\mu$g/m$^3$). For LC, the female control group was assumed to be using only gas stoves with an estimated PM$_{2.5}$ exposure of 70 $\mu$g/m$^3$. For males, the exposure was again assumed
to be 65% of females resulting in an equivalent exposure of 45.5 µg/m³. The meta-analytic summary risk estimate for male COPD incidence in association with HAP PM$_{2.5}$ was 1.90 (95% CI: 1.56, 2.32) and for females was 2.70 (95% CI: 1.95, 3.75). For LC incidence among males, the summary risk estimate was 1.26 (95% CI: 1.04, 1.52) and among females was 1.81 (95% CI: 1.07, 3.06).

The lower exposure estimates in the HAPs studies are substantially higher than counterfactual exposure due to the nearby use of less clean fuels and therefore these RRs are not directly comparable to those obtained from AAP, SHS, or AS types that compare to either the counterfactual (AAP) or a 0 µg/m³ exposure (SHS, AS). This information was included in the curve fitting process by equating the observed RRs to the ratio of the IER model evaluated at the respective two PM$_{2.5}$ concentrations.

The HAP studies estimated effects on incidence rather than mortality. For the purposes of building the IER, we assume that the RRs of mortality and incidence are equal.

**Age-Modification Risk Models for Ischemic Heart Disease and Stroke Mortality**

Epidemiologic studies of risk factors for both IHD and stroke indicate that the RR declines with the logarithm of age, reaching 1 between 100 and 120 (Singh et al. 2013). We thus modified the type-specific RR for both IHD and stroke mortality using a linear regression model of the logarithm of the median age at death for each study with intercept equal to 1 at age 110. The slope of the regression line was estimated from a meta-analysis of several risk factors (Singh et al. 2013). We applied this age-modification to the RRs and fit the IER model for each age group separately.

**Selecting the Counterfactual Exposure**

For each risk factor examined by GBD 2010 (Lim et al. 2012), the distribution of exposure was compared to an alternative (counterfactual) distribution termed the theoretical-minimum-risk
exposure distribution (TMRED). For AAP, zero exposure is not a practical counterfactual level as it is impossible to achieve even in pristine environments (Brauer et al. 2012). Furthermore, the lowest level of exposure to PM$_{2.5}$ that is deemed beneficial has not been clearly identified. Defining the TMRED was based on two criteria (Lim et al. 2012): the availability of convincing evidence from epidemiological studies that support a continuous reduction in risk of disease to the chosen distribution; and a distribution that is theoretically possible at the population level.

Lim et al. (2012) suggest that a positive counterfactual concentration be used. Their counterfactual concentration is bounded by the minimum concentrations observed in the studies used to estimate risk and some low percentile of the PM$_{2.5}$ distribution. There is clearly no evidence of an association below observed levels and it is impractical to estimate the shape of the curve at the extremes of the exposure distribution. Lim et al. (2012) suggest that the 5$^{th}$ percentile be used and that the lower and upper bounds on the counterfactual concentration be determined by the corresponding minimum and 5$^{th}$ percentiles, respectively, of the AAP PM$_{2.5}$ exposure distribution for the American Cancer Society Cancer Prevention cohort (Krewski et al. 2009), the largest cohort study of air pollution. The minimum was 5.8 $\mu$g/m$^3$ and the 5$^{th}$ percentile was 8.8 $\mu$g/m$^3$. Uncertainty in the counterfactual concentration was modelled as a uniform distribution between the minimum and 5$^{th}$ percentile.

**Estimation of PAF**

We estimated the PAF associated with ambient PM$_{2.5}$ exposure for all 187 countries separately for 2005. We first estimated surface PM$_{2.5}$ concentrations on a 0.1° by 0.1° grid for the globe using a combination of remote sensing and atmospheric models calibrated to ground monitoring data (Brauer et al. 2012). For each grid cell within a country we estimated the RR based on the IER model at the estimated PM$_{2.5}$ concentration. We then constructed a population-weighted average RR for each country using the corresponding population count 0.1° by 0.1° grid cell (Brauer et al. 2012
Both the gridded PM$_{2.5}$ and population values can be obtained from Brauer et al. 2012. The country-specific PAF = 1 – 1/$WRRIER$, where $WRRIER$ is the population weighted average of the $RRRIER$ values at each PM$_{2.5}$ grid cell within the country.

**Integrated Exposure-Response Model for Acute Lower Respiratory Infection**

Mehta et al. (2011) reviewed the evidence for an association between exposure to ambient PM$_{2.5}$ and ALRI. Four cohort studies were deemed appropriate to include in an IER model (Mehta et al. 2011). We included 23 studies of parental SHS and ALRI reported by the SGR (2006) with each study-specific odds ratio (OR) assigned a PM$_{2.5}$ equivalent ambient exposure of 50 µg/m$^3$, assuming a moderate to high level of exposure. Smith et al. (2011) examined the relationship between exposure to carbon monoxide (CO) from the burning of solid biomass for heating and cooking and the incidence of ALRI in Guatemala and reported incidence rates by decile average of CO personal exposures. These decile CO averages were converted to PM$_{2.5}$ concentrations using the equation: $\text{PM}_{2.5}$ (mg m$^{-3}$) = 0.10 (0.093, 0.12) × CO (mg m$^{-3}$) + 0.067 (0.0069, 0.13) with 95% confidence intervals displayed in parenthesis (Northcross et al. 2010). This equation had good predictive power ($R^2 = 0.76$).

Incidence rates, $I(z_i)$, corresponding to the 10 decile values of PM$_{2.5}$, denoted by $z_i$ for $i=1,…,10$, can be compared to the risk model by taking the ratio of incidence rates for all unique pairs of PM$_{2.5}$ deciles, a total of 45 pairs, and equating them to the ratio of the corresponding risk model evaluated at the appropriate decile average. That is

$$RR_{ALRI}(z_i, z_j) = \frac{I(z_i)}{I(z_j)} = \frac{\left[ 1 + \alpha \{ 1 - \exp[-\gamma (z_i - z_{cf})^\delta] \} \right]}{\left[ 1 + \alpha \{ 1 - \exp[-\gamma (z_j - z_{cf})^\delta] \} \right]} \quad [2]$$
for all 45 unique pairs of concentrations \((z_i, z_j)\), \(\forall i > j = 1, \ldots, 10\). The 45 incidence rate ratios were combined with the four AAP cohort study ORs and the 23 SHS ORs in order to fit the IER model for ALRI. We assumed the same counterfactual uncertainty distribution as with the mortality IER models.

**Results**

The average of the \(RR_{IER}\) predictions among the simulations are displayed for the four causes of death in Figure 1 in addition to the 95% confidence intervals (CI) and the type-specific RR estimates and corresponding 95% confidence intervals used to fit the curves. The HAPs RRs for COPD and LC are presented in this figure as a pink shaded box with height of the box representing the uncertainty in the RR estimates and the width representing the exposure contrast at which the RRs was assumed to pertain to. Two shaped boxes are displayed for both COPD and LC corresponding to the two RR estimates for males and females. The box is centered at the RR estimate and the midpoint of the two exposure values. This alternate depiction of the HAPs information was necessary since the lowest exposure levels were substantially higher than the counterfactual exposure and thus not directly comparable to the RRs from the other sources. The pooled estimate of RR and its corresponding CI for SHS is displayed in placed of the study-specific SHS RRs for each unique PM\(_{2.5}\) value since the study-specific RRs and CI could not be visually distinguished. Similar results are presented for ALRI in Figure 2. In addition to the RR the incidence of ALRI is also displayed on the right hand y-axis.

The \(RR_{IER}\) function fits well the RRs for all types of PM\(_{2.5}\) and causes of mortality except for COPD and HAP in which the IER model underestimates the observed RRs (Figure 1). This may be due to the use of the ratio of incidence rates rather than RR based on mortality data for this outcome.
However, the IER curve fits the LC incidence data reasonably well. The time between diagnosis of COPD and mortality is much longer than that for LC and thus the LC incidence data may better reflect mortality patterns than the COPD incidence data.

We compared the country specific estimated PAFs using the age-modified models to those models using age independent data. Age-modified $RR_{IER}$ curves are displayed for IHD and stroke mortality in Supplemental Material, Figure S15 (top panels) with generally decreasing risk with increasing age. The country-specific PAFs based on risk models not modified by age and those in which age-modification models were used for both IHD and stroke mortality are presented in Supplemental Material, Figure S15 (bottom panels). Incorporation of age-modification risk models tends to slightly decrease the PAF estimates.

The distribution of population-weighted country-average PM$_{2.5}$ concentrations and PAFs are displayed in Figure 3. The country average PM$_{2.5}$ concentrations ranged from 2 $\mu$g/m$^3$ to 70 $\mu$g/m$^3$ for 2005 (Figure 3A) while the country level PAFs were less than 0.4 for ALRI, IHD, and stroke and less than 0.25 for LC and 0.2 for COPD (Figure 3B).

Plots similar to Figures 1 and 2 are displayed for the other seven model forms examined in Supplemental Material, Figures S1-S14 for both the four causes of death (Supplemental Material, Figures S1-S7) and ALRI (Supplemental Material, Figures S8-S14). In addition, both the AICs and BICs are given in Supplemental Material, Table S2 for all eight models and five outcomes. The IER model was a better predictor of the type-specific RRs than the other seven model examined for ALRI and three of the four causes of death. For COPD mortality, the Power model provided a better fit than the IER model based on lower AIC and BIC values (Supplemental Material, Table S2). This was likely due to the better prediction of the HAPs RR for which the IER model clearly
underestimated the RR. Graphical comparisons of the predicted values (solid lines) to the type-specific RRs (dots) in Supplemental Material, Figures S1-S14 verify the conclusions drawn from the AIC/BIC results.

**Discussion**

Exposure to PM$_{2.5}$ in ambient air has been linked to increased risk of death from chronic cardiovascular and respiratory disease and lung cancer in cohort studies in the United States and Europe (Chen et al. 2008; USEPA 2009). Unfortunately there are few long-term cohort studies for these diseases in other regions, such as East and South Asia and the Middle East, where ambient exposures are much higher and where the relative contribution of specific sources of air pollution differ from the North America and Europe (Brauer et al. 2012; Heath Effects Institute 2010).

To derive the shape of the exposure-response curve at higher ambient concentrations we incorporated information on risk due to exposure to SHS, HAP, and AS to extend the risk estimates to higher exposures. The Integrated Exposure-Response (IER) model combines information on mortality RR from separate types of combustion, unified by equating delivered dose from all types in terms of equivalent ambient PM$_{2.5}$ exposures. Although we assume that the toxicity of PM$_{2.5}$ exposure, as characterized by RR, changes with the magnitude of exposure, we also assume that at any fixed exposure level, toxicity is roughly equivalent among all types and temporal patterns of PM$_{2.5}$ exposure. These are important assumptions since estimated PM$_{2.5}$ exposure throughout the world, whether from ambient origin or household indoor combustion, has not been differentiated by the components or sources of fine particulate matter.

Only evidence from multiple epidemiologic studies of long-term exposure to PM$_{2.5}$ in highly polluted settings can provide definitive estimates of the shape of the exposure-response function for
mortality from chronic cardiovascular and respiratory diseases. These are starting to appear however. For example, Cao and colleagues (2011) report increased risk of mortality from cardiovascular and respiratory disease and lung cancer associated with long-term exposure to total suspended particulates (TSP) in 71,000 residents of 31 Chinese cities. This study offers an opportunity to assess the ability of our \( RR_{IER} \) model to estimate the observed RRs in situations with very high levels of outdoor air pollution. In order to estimate PM\(_{2.5}\) RRs in the cohort, the authors used a 3/1 ratio to convert TSP to PM\(_{2.5}\) based on current and historical Chinese data (Cao et al. 2011). Estimated PM\(_{2.5}\) (converted from TSP) concentrations ranged among cities from 38 to 166 \( \mu g/m^3 \). Increases of 2.1% (-0.3%, 4.6%), 3.3% (0.9%, 5.4%), and 3.3% (-0.3%, 6.9%) in IHD, stroke, and LC mortality, respectively, were associated with a 10 \( \mu g/m^3 \) change in estimated equivalent PM\(_{2.5}\) exposures in this cohort (Haidong Kan, personal communication).

Because the cohort members did not experience exposures near the lowest concentrations applicable to our RR model (i.e., the counterfactual concentration), we cannot determine RRs estimated from the cohort and directly compare them to our RR model, which is relative to a much lower counterfactual concentration. However, we can determine RR between concentrations observed in the cohort itself. We first determined the mean of the four quartiles of PM\(_{2.5}\) concentrations as 40, 91, 106, and 127 \( \mu g/m^3 \), respectively (Haidong Kan, personal communication) and calculated the RR between consecutive quartile averages assuming the exponential risk model form as was used by the study authors. The geometric average of these three RRs was then determined as a summary measure of change in risk over the PM\(_{2.5}\) exposure distribution. A similar calculation was undertaken for the \( RR_{IER} \) model. The RRs observed in the Chinese cohort and those predicted by \( RR_{IER} \) were similar for the three causes of death examined (IHD: China RR = 1.06; 95% CI: 0.99, 1.14 and IER RR = 1.05; 95% CI: 1.03, 1.1; stroke: China RR = 1.10; 95% CI: 1.03, 1.17 and IER 18
RR = 1.08; 95% CI: 1.01, 1.14; LC: China RR = 1.10; 95% CI: 0.99, 1.22 and IER RR = 1.09; 95% CI: 1.06, 1.12), suggesting that our IER model yielded reasonable predictions in the change in risk over a range of concentrations that prevail in China and other highly polluted settings that were not observed in cohort studies conducted in North America and Western Europe.

We note, however, some limitations in this comparison. First, TSP was a poorer predictor of cardiovascular mortality than PM$_{2.5}$ in US-based cohort studies (Pope et al. 2002). Secondly, we note that uncertainty about the temporal and spatial consistency of the TSP/PM$_{2.5}$ conversion ratio of 3/1 adds uncertainty to our interpretation of the results from the Chinese cohort.

There are additional uncertainties due to lack of information on actual exposure to PM$_{2.5}$ for some source-specific RRs used to fit the model, notably (a) scarce information on actual exposure from SHS in the relevant epidemiological studies (Pope et al. 2009; 2011) which required the estimation of PM$_{2.5}$ concentrations from other studies, (b) potential misclassification of exposure for SHS estimates due to possible co-exposure from AAP of the exposed group, and (c) the duration of exposure, which differs when it comes to exposures from AAP, SHS, HAP, and active smoking: the lifetime duration of exposure in active smoking may be much shorter than in the other exposures and the received doses may therefore not be proportional to concentrations according to type of exposure. Uncertainties may be reduced by improving precision in the actual exposure estimates of the relative risks from the epidemiologic literature used for developing the proposed model.

Multiple studies were used to estimate RRs associated with exposure to AAP, SHS, and HAPs. For AS, we estimated RRs for active cigarette smokers from a single cohort, the American Cancer Society Cancer Prevention II. This cohort was also used by GBD 2010 to estimate risk specifically for AS (Lim et al. 2012). However, the pattern of the association between the number of cigarettes
smoked per day and cause-specific mortality observed in the CPS II cohort may not reflect the patterns observed in other cohort studies of active smoking (see Pirie et al. 2012 as one example). Similarly, the IER for ALRI is fit through RR from studies of AAP and SHS conducted in a limited number of mostly high-income countries, and a single developing country RR estimate for HAP PM$_{2.5}$ exposure and ALRI (RESPIRE, Smith et al. 2011). We thus recommend that future work on the IER function include additional sensitivity analyses of the type-specific RRs to which the curve is fit. Future work could also include the uncertainty in the estimate of PM$_{2.5}$ from CO and new information in this relationship (McCracken et al. 2013).

The key assumptions that underlie the IER, discussed above, largely serve to justify the integration of risk estimates for different types of PM exposure. These assumptions, and their tenability, have been addressed elsewhere (Pope et al. 2009; 2010; 2011). Unfortunately, for several of the most critical assumptions, those concerning the relative toxicity per unit mass of PM$_{2.5}$ of different types (e.g., AAP and AS), not accounting for the temporal pattern of exposure, and the absence of interaction among types of combustion, there is little empirical evidence against which to evaluate those assumptions or to evaluate in detail specific implications of their violation. Each warrants additional research.

Although we set the counterfactual concentration to be drawn from a uniform distribution with lower bound of 5.8 $\mu$g/m$^3$ and an upper bound of 8.8 $\mu$g/m$^3$, we are not suggesting that there is convincing evidence that PM$_{2.5}$ mortality and ALRI risk is zero below any specific concentration based on biological considerations (Brook et al. 2010). Absence of such evidence from epidemiological studies does not necessarily imply evidence of the absence of such a counterfactual concentration. We thus take the conservative approach and set a positive counterfactual concentration. We note however, our approach can be adapted to a different counterfactual if new
evidence supporting a positive association at lower concentrations becomes available. One such piece of evidence was observed in Canada where positive associations down to 2 µg/m³ were noted (Crouse et al. 2012).

The Lin50 and Log models proposed by Cohen et al. (2004) were used for the previous GBD estimates and the Log model is currently recommended by WHO (Ostro 2004). However the unknown parameters in these models were estimated from a single cohort study of AAP, the CPS-II, which required analysis of the original data. The IER model uses RR estimates available in the open literature, allowing periodic updating of risk functions based on systematic review of the literature, and does not require analyses of primary data not in the public domain. As new epidemiologic studies and evidence on type-specific PM$_{2.5}$ exposure appear the models can be re-estimated by any interested member of the scientific community using publically available information

**Conclusion**

Fine particulate mass-based RR models can be developed that cover the entire global range of ambient exposure to PM$_{2.5}$ by integrating RR information from different combustion sources that generate emissions of particulate matter. A specific RR model form was identified that provided superior predictive power for leading global causes of mortality for air pollution compared to a range of alternative model forms.
References


Estimates of global mortality attributable to particulate air pollution using satellite imagery. Environ. Res. 120:33–42.


Figure Legends

Figure 1. Predicted values of $I_{ER}$ model (solid line) and 95% confidence interval (dashed line) for ischemic heart disease (IHD), stroke, chronic obstructive pulmonary disease (COPD), and lung cancer (LC) mortality. Type-specific relative risks (RR) (points) and 95% confidence intervals (error bars) also presented. Green dots represent AAP cohort study RRs, blue triangles represent pooled SHS RRs, and black dots represent AS RRs. Shaded boxes for COPD and LC mortality represent uncertainty (height) and exposure contrast (width) of RR HAP estimates for males and females separately. PM$_{2.5}$ concentrations are on the logarithmic (base 10) scale on x-axis.

Figure 2. Predicted values of $I_{ER}$ model (solid line) and 95% confidence interval (dashed line) for Acute Lower Respiratory Infection infants (ALRI). Type-specific relative risks (RR) (points) and 95% confidence intervals (error bars) also presented. Green dots represent AAP cohort study RRs, blue triangle represents pooled SHS RRs, and red dots represent household air pollution (HAP) RRs. Right hand y-axis represents HAP incidence rates. PM$_{2.5}$ concentrations are on the logarithmic (base 10) scale on x-axis.

Figure 3. Density plots of country-specific PM$_{2.5}$ concentrations ($\mu$g/m$^3$) (3A) and Population Attributable Fractions (3B) by risk model and health outcome. Dashed line represents smooth fit of density function.
Figure 1.
Figure 2.
Figure 3.