Supporting Information for

# Blue Skies Bluer?

Julian D. Marshall[*,1], Joshua S. Apte[2], Jay S. Coggins[3], Andrew L. Goodkind[3]

[1] Dept. of Civil, Environmental, and Geo-Engineering, University of Minnesota, Minneapolis, MN 55455-0233;
[2] Dept. of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX 78712;
[3] Dept. of Applied Economics, University of Minnesota, St. Paul, MN 55108.
[*] Corresponding author: E-mail: julian@umn.edu, phone: +1 (612) 625-2397

**Number of pages:** 10

**Number of tables:** 2

**Number of figures**: 1

**Contents:**

## S1. Derivation of C-R functions

*Log-linear C-R ("linear")*

The log-linear Cox proportional hazards model is presented by Krewski et al. (2009) in equation (3), on Page 16:

$$\lambda_{ijk}^{(s)}(t) = \lambda_0(t) u_{ij} \exp(\beta^T x_{ijk}^{(s)}) \quad [S1]$$

Here, $\lambda_{ijk}^{(s)}(t)$ is the hazard function for individual $i$ in cluster $j$ and sub-cluster $k$ from stratum $s$ at time $t$; and $x_{ijk}^{(s)}$ is the corresponding exposure matrix.[1] Rearranging terms, the logarithm of the hazard ratio $HR$ can be demonstrated to be linearly proportional to the exposure $x$ with coefficient $\beta$. Neglecting the model terms and subscripts not relevant for this proportionality yields the following relationship:

$$\ln(HR) = \ln\left(\frac{\lambda^{(s)}(t)}{\lambda_0(t)}\right) \sim (\beta x) \quad [S2]$$

Thus, the logarithm of the hazard ratio is linearly proportional to the matrix of exposures $x$ and vector of coefficients $\beta$. Simplifying the notation, the continuous hazard ratio (equivalent to relative risk) function for exposure to an ambient concentration $C = x$ has the following concentration dependence:

$$HR = \exp\left(\beta(C - C_{\text{min}})\right). \quad [S3]$$

In equation [S3], $HR$ is the hazard ratio, $C_{\text{min}}$ is the referent condition for which $HR = 1$, and $\beta$ is the empirical coefficient to be determined. Here, we use $C_{\text{min}} = 5.8$ μg m$^{-3}$ $PM_{2.5}$, consistent with the minimum annual-average concentration assigned to subjects in the extended analysis of the ACS Cohort Study.[1] A key feature of this C-R function is that the relative risk for any fixed increment in concentration is constant. For cardiopulmonary mortality, Krewski et al. (2009)

report HR = 1.128 for any 10 μg $m^{-3}$ increment in concentration. Rearranging terms, we compute $\beta$ as follows:

$$HR = 1.128 = \frac{\exp\left(\beta(C + 10 - C_{\min})\right)}{\exp\left(\beta(C - C_{\min})\right)} = \exp\left(10\beta\right) \quad [S4]$$

$$\beta = \frac{\ln(1.128)}{10} = \boxed{0.012045}.$$

*Log-log C-R ("supralinear")*

In the log-log Cox proportional hazards model of Krewski et al. (2009), the $PM_{2.5}$ concentration is entered into the survival model as a logarithmically transformed variable (see unnumbered equations on Page 27 of Krewski et al. 2009):

$$\lambda^e(t) = \lambda_0(t) U^{r(e)} \exp(\{\ln(\mathrm{PM}_{2.5})\gamma + X^e\beta\}). \quad [S5]$$

As above, $\lambda$ represents the hazard function, while $PM_{2.5}$ is entered into the model as a distinct, log-transformed variable with its own coefficient γ that is separate from the other covariates *X*. Rearranging terms, the hazard ratio can be demonstrated to have the following proportionality for $PM_{2.5}$ after omitting other covariates:

$$\ln(HR) = \ln\left(\frac{\lambda^e(t)}{\lambda_0(t)}\right) \sim \gamma \ln(\mathrm{PM}_{2.5}). \quad [S6]$$

Therefore, the resulting C-R function for this model conforms to the following power-law functional form for concentration *C:*

$$HR = \left(C / C_{\min}\right)^{\gamma}. \quad [S7]$$

In equation [S7], γ is an empirical regression coefficient that can be determined from a point estimate of *HR* for any two concentrations. In Table 11, Krewski et al. (2009) report a

cardiopulmonary mortality *HR* = 1.208 for a change in concentration from 5 to 15 μg $m^{-3}$. Mathematically, the following relationship holds:

$$1.208 = \frac{\left(15/C_{\min}\right)^{\gamma}}{\left(5/C_{\min}\right)^{\gamma}}.$$

We rearrange to solve for γ:

$$\gamma = \frac{\ln(1.208)}{\ln(15/5)} = 0.17200.$$

To check this result, a second cardiopulmonary mortality point estimate from Table 11 is employed. For a change in concentration between 10 and 20 μg $m^{-3}$, *HR* = 1.127, which corresponds to the following value of γ:

$$\gamma = \frac{\ln(1.127)}{\ln(20/10)} = 0.17249$$

This second value differs from the first by 0.3%, or well within the range of rounding error for amounts given in the table. We employed the average of these two point estimates above:

$\boxed{\gamma = 0.17225}$.

## S2. Derivation of change in attributable burden for C-R functions

For the family of C-R functions employed by Krewski et al. (2009), the reduction in attributable burden Δ*AM* is independent of assumptions about the theoretical minimum risk-concentration $C_{min}$ for any initial and final concentrations $C_1$ and $C_2$ that satisfy the conditions that ($C_1 \geq C_{min}$) and ($C_2 \geq C_{min}$). We will demonstrate this property of these C-R functions by first manipulating equation 2, and then substituting in the particular functional forms of the C-R relationships of Krewski et al.:

$$\Delta AM = AM(C_1) - AM(C_2) = M_{\min} \times \left[ RR(C_1) - RR(C_2) \right] = M_{obs} \times \frac{RR(C_1) - RR(C_2)}{RR(C_{obs})} \quad [S8]$$

Equation S8 embodies a conventional attributable-fraction type calculation. For the sign convention employed here, a reduction in attributable mortality (i.e, by moving to a lower concentration $C_2 < C_1$) results in a positive value of $\Delta AM$: i.e., a reduction in attributable mortality.

We will now substitute the functional form for Krewski's linear C-R relationship (equation S3) into equation S8 in order to demonstrate that $\Delta AM$ is independent of the theoretical minimum-risk concentration:

$$\Delta AM = M_{obs} \times \frac{RR(C_1) - RR(C_2)}{RR(C_{obs})} = M_{obs} \times \left[ \frac{\exp\left(\beta \times (C_1 - C_{\min})\right) - \exp\left(\beta \times (C_2 - C_{\min})\right)}{\exp\left(\beta \times (C_{obs} - C_{\min})\right)} \right]$$

$$\boxed{\Delta AM = M_{obs} \times \left[ \exp(-\beta C_{obs}) \times \left( \exp(\beta C_1) - \exp(\beta C_2) \right) \right].} \qquad [S9]$$

Thus, for a linear C-R function, $\Delta AM$ is entirely independent of the minimum-risk concentration $C_{min}$.

We now turn to the supralinear C-R of Krewski et al. Analogously, we substitute that C-R function (equation S7) into equation S8 to demonstrate the independence of $\Delta AB$ on $C_{min}$.

$$\Delta AM = M_{\text{obs}} \times \frac{RR(C_1) - RR(C_2)}{RR(C_{\text{obs}})} = M_{\text{obs}} \times \left[ \frac{\left( \frac{C_1}{C_{\min}} \right)^{\gamma} - \left( \frac{C_2}{C_{\min}} \right)^{\gamma}}{\left( \frac{C_{\text{obs}}}{C_{\min}} \right)^{\gamma}} \right] =$$

$$\boxed{\Delta AM = M_{obs} \times \left[ \left( \frac{C_1}{C_{obs}} \right)^{\gamma} - \left( \frac{C_2}{C_{obs}} \right)^{\gamma} \right].} \qquad [S10]$$

Again, for the supralinear function of Krewski et al, the change in risk $\Delta AM$ between any two concentrations does not depend on the assumed minimum-risk concentration $C_{min.}$

## S3. Data for mortality calculations (Figure 1)

*Mortality data*

We obtained a value of $M_{obs}$ equivalent to ~345 adult cardiopulmonary deaths (age > 30) per 100,000 all-age population. To derive this value, we employed cause-specific mortality data from the 2010 Global Burden of Disease (GBD) assessment, focusing on regional data for high-income North America, which is dominated by US adults. Krewski et al. (2009) classified as cardiopulmonary mortality all subject deaths with ICD-9 codes in the ranges 409-440 and 460-519. Examples of major causes of death within these codes include ischemic and hypertensive heart diseases, cerebrovascular disease, acute respiratory infections, chronic obstructive pulmonary disease (COPD), pneumonia, and influenza. To reconstruct the cardiopulmonary disease category from GBD data, we used the GBD causes of death indicated in Table S1, below.

**Table S1** – Cardiopulmonary causes of death in GBD 2010 [2]

| *Disease Grouping* | *Cardiovascular diseases* | *Pulmonary diseases* | *Lower respiratory infections* |
|---|---|---|---|
| *Specific Causes* | Ischemic heart disease<br>Ischemic stroke<br>Hemorrhagic and other non-ischemic stroke<br>Hypertensive heart disease<br>Atrial fibrillation and flutter<br>Cardiomyopathy and myocarditis<br>Endocarditis | COPD<br>Asthma<br>Pneumoconiosis<br>Other chronic respiratory diseases | Influenza<br>Pneumococcal pneumonia<br>H influenzae type B pneumonia<br>Respiratory syncytial virus pneumonia<br>Other lower respiratory infections |
| *Mortality rate* [a] | 265 | 53 | 27 |

[a] Year 2010 over-30 adult deaths per 100,000 all-age population, in high-income North America. Corresponding all-cause mortality rate is ~846 adult deaths per 100,000 all-age population.[2]

## S4. Sensitivity case: integrated-exposure response function

To examine the sensitivity of our results to the choice of supralinear C-R function, we consider here the supralinear integrated exposure-response (IER) functions of employed in the Global Burden of Disease 2010 study[3]. Burnett et al. (2014)[3] provide IER functions for five

major disease endpoints: acute lower respiratory infections (ALRI) in children under 5 y of age, and ischemic heart disease (IHD), cerebrovascular disease (stroke), chronic obstructive pulmonary disease (COPD) and lung cancer (LC) in adults over 30 y of age. Of these functions, the IHD, stroke, and COPD endpoints correspond most closely to the cardiopulmonary disease endpoints that are considered by Krewski et al. Apte et al. (2015) developed a set of "synthetic" IER functions that report a single C-R function for multiple disease endpoints by weighting the relative risk of each cause by its baseline incidence rate.[4] Figure S1 reproduces two synthetic IER functions here: a function that incorporates all five IER outcomes ("IER-5"), and a function that includes only the three cardiopulmonary outcomes (IHD, stroke, and COPD, "IER-3"). Full details of the derivation are provided in the supplementary information of Apte et al. (2015).[4]

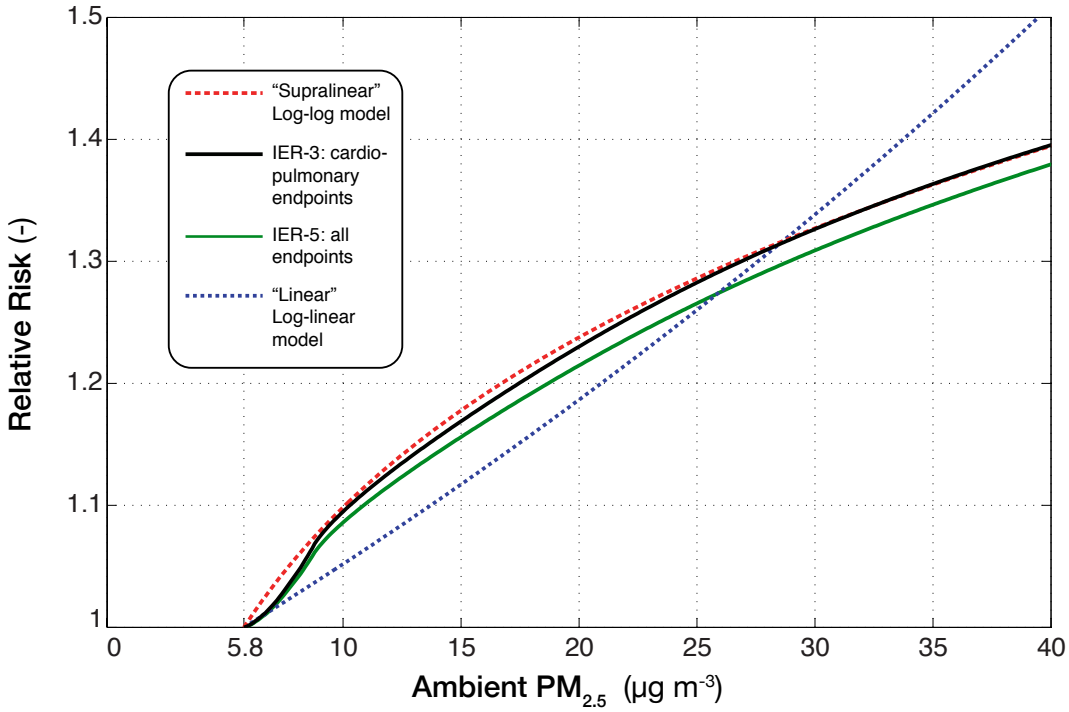


**Figure S1** – Comparison of the supralinear IER functions of Burnett et al. (solid lines, marked "IER-3" and "IER-5") with the Krewski et al. C-R functions used in the core analysis (dashed lines, marked "supralinear" and "linear". Note the similarity between the IER-3 function (black line) and the Krewski supralinear function (dashed red line).

In our sensitivity analyses below, we employ the IER-3 function, as it pertains to the cardiopulmonary endpoints we consider in base-case analyses. Note that the IER-3 function very closely tracks the Krewski et al. supralinear function down to ~ 8 μg $m^{-3}$. The inflection in the IER-3 function below ~8 μg $m^{-3}$ reflects averaging over alternative possible realizations of $C_{min}$ (sampling interval: 5.8 – 8 μg $m^{-3}$), as described by Apte et al. (2015).

Table S2 presents a reassessment of our core results from Table 2 using the IER-3 functions. The total year-2010 attributable cardiopulmonary disease mortality for the IER-3 function (116,000 deaths $y^{-1}$) is similar to that assessed by the core supralinear C-R of Krewski et al. For scenarios where $PM_{2.5}$ levels are limited to a specific maximum concentration (e.g., 15, 12, 10 or 8 μg $m^{-3}$), the reduction in attributable mortality predicted by the IER-3 function generally results in similar values to those predicted by the Krewski et al supralinear function. This finding should not be altogether surprising, since the IER functions are informed in part by data from the ACS cohort studies.

**Table S2 –** Potential reductions in $PM_{2.5}$-attributable cardiopulmonary mortality by limiting maximum $PM_{2.5}$ levels in the U.S.

| | **Linear C-R** | | **Supralinear C-R** | | **IER-3**[e] | |
|---|---|---|---|---|---|---|
| | *deaths $y^{-1}$* | *%*[a] | *deaths $y^{-1}$* | *%* | *deaths $y^{-1}$* | % |
| **Year-2010 attributable mortality** | 80,100 | | 122,000 | | 116,000 | |
| **Mortality reduction for achieving $C \leq C_{max}$ throughout the U.S.** | | | | | | |
| $C_{max}$ = 15 μg $m^{-3}$ [b] | 7,600 | 9% | 6,400 | 5% | 6,500 | 5% |
| $C_{max}$ = 12 μg $m^{-3}$ [c] | 20,400 | 25% | 20,400 | 17% | 19,700 | 17% |
| $C_{max}$ = 10 μg $m^{-3}$ [d] | 35,600 | 44% | 40,400 | 33% | 38,400 | 33% |
| $C_{max}$ = 8 μg $m^{-3}$ | 55,300 | 69% | 71,800 | 59% | 77,900 | 68% |

[a] Percentage reduction in annual $PM_{2.5}$ attributable mortality relative to year-2010 levels for a hypothetical standard that immediately limited annual $PM_{2.5}$ levels to the target concentration $C_{max}$.
[b] Previous U.S. EPA $PM_{2.5}$ National Ambient Air Quality Standard (NAAQS) was 15 μg $m^{-3}$ annual average.
[c] Current U.S. EPA $PM_{2.5}$ NAAQS is 12 μg $m^{-3}$ annual average.
[d] Current World Health Organization $PM_{2.5}$ air quality guideline is 10 μg $m^{-3}$ annual average
[e] Synthetic integrated exposure-response (IER) function for three cardiopulmonary endpoints, as described in text above. IER functions derived from data presented by Burnett et al. (2014).[3]

## S5. Concentration targets to reach a fixed level of attributable mortality

In the case of both cardiopulmonary and all-cause mortality, and for the majority of the concentration range considered here (< ~27 μg $m^{-3}$), a lower concentration is required to reach a given level of attributable mortality for a supralinear log-log C-R relationship than for a log-linear C-R relationship (e.g., see Figure 1(a)). For example, the attributable cardiopulmonary mortality for the log-linear relationship at 12 μg $m^{-3}$ is comparable to the attributable cardiopulmonary mortality in the log-log relationship at ~9.1 μg $m^{-3}$. These calculations were performed by rearranging equation 1 to solve for the concentration consistent with a fixed level of per-capita attributable mortality. Let $RR^{L}(C)$ denote the C-R function for the linear model, and let $RR^{S\text{-}L}(C)$ denote the C-R function for the alternative supralinear model. For the supralinear model to reach a level of attributable mortality at some concentration $C'$ that is equal to the attributable mortality of the log-linear model at a different concentration $C^*$, the following equality must hold:

$$\frac{RR^{L}(C^{*})-1}{RR^{L}\left(C_{\text{obs}}\right)}=\frac{RR^{S-L}(C')-1}{RR^{S-L}\left(C_{\text{obs}}\right)}. \qquad \text{[S11]}$$

For the example provided above, a given level of attributable cardiopulmonary mortality for linear C-R is computed at $C^*$ = 12 μg $m^{-3}$. Solving equation S11 for C′ demonstrates that a comparable level of attributable cardiopulmonary mortality for the supralinear model is reached at ~ 9.1 μg $m^{-3}$.

## References


1. Krewski, D.; Jerrett, M.; Burnett, R. T.; Ma, R.; Hughes, E.; Shi, Y.; Turner, M. C.; Pope, C. A.; Thurston, G.; Calle, E. E., et al. *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality*; Research report 140; Boston, MA, Jun, 2009; http://pubs.healtheffects.org/view.php?id=315.
2. Lozano, R.; Naghavi, M.; Foreman, K.; Lim, S.; Shibuya, K.; Aboyans, V.; Abraham, J.; Adair, T.; Aggarwal, R.; Ahn, S. Y., et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: A systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* **2012**, *380*, 2095-2128.
3. Burnett, R. T.; Pope, C. A.; Ezzati, M.; Olives, C.; Lim, S. S.; Mehta, S.; Shin, H. H.; Singh, G.; Hubbell, B.; Brauer, M., et al. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. *Environ. Health Persp.* **2014**, *122*, 397-403.
4. Apte, J. S.; Marshall, J. D.; Brauer, M.; Cohen, A. J. Addressing global mortality from ambient $PM_{2.5}$. *Environ. Sci. Technol.* **2015**, *49*, 8057–8066.