Appendix H: Additional Details on Benefits Methodologies

H.1 Methodology Used to Develop Threshold Adjusted Concentration-Response Functions

For mortality and morbidity outcomes associated with short-term exposure to $PM_{2.5}$, log-linear C-R functions are developed based on a continuous function from the epidemiological studies. Generally, the lowest measured concentrations in the short-term exposure studies were relatively near or below the estimated background levels such that little or no extrapolation of the C-R function is required beyond the range of data in the studies. Among the studies of mortality associated with long-term exposure to $PM_{2.5}$ that have been included in the benefits analysis, the lowest measured long-term levels were in the range 7.5 to 11 $\mu g/m^3$. For the base cases and sensitivity analyses we applied various alternative "cutpoint" models. While there are likely biological thresholds in individuals for specific health responses, the available epidemiological studies do not support or refute the existence of thresholds at the population level for either long-term or short-term PM exposures within the range of air quality observed in the studies. It may therefore be appropriate to consider health risks estimated not only with the reported linear or log-linear C-R functions, but also with modified functions that approximate non-linear, sigmoidal-shaped functions that would better reflect possible population thresholds. We approximated such sigmoidal functions by "hockeystick" functions based on the reported linear or log-linear functions. This approximation consisted of (1) imposing a cutpoint (i.e., an assumed threshold) on the original C-R function, that is intended to reflect an inflection point in a typical sigmoidal shaped function, below which there is little or no population response, and (2) adjusting the slope of the original C-R function above the cutpoint. This approach mirrors the approach used in the Particulate Matter Health Risk Assessment (Post et al., 2005).

If the researchers in the original study fit a log-linear or a linear model through data that actually better support a sigmoidal or "hockeystick" form, the slope of the fitted curve would be smaller than the slope of the upward-sloping portion of the "true" hockeystick relationship, as shown in Figure H-1a. The horizontal portion of the data below the cutpoint would essentially cause the estimated slope to be biased downward relative to the "true" slope of the upwardsloping portion of the hockeystick. The slope of the upward-sloping portion of the hockeystick model should therefore be adjusted upward (from the slope of the reported C-R function), as shown in Figure H-1a. This rationale applies equally in the case of mortality associated with long- and short-term exposure to PM. In each case, under the threshold hypothesis a log-linear curve has been fit to data that are better characterized by a hockeystick model. In the case of a short-term exposure mortality or morbidity study, the curve represents the relationship between daily PM and daily mortality or morbidity; in the case of a long-term exposure mortality study, the curve represents the relationship between annual average PM and annual mortality. In both cases, however, if the "true" relationship looks like a hockeystick, then the log-linear curve fitted to the data would understate the impact of increases in PM (either daily, in the case of a short-term study, or annual average, in the case of a long-term study) on mortality or morbidity at PM levels above the cutpoint.

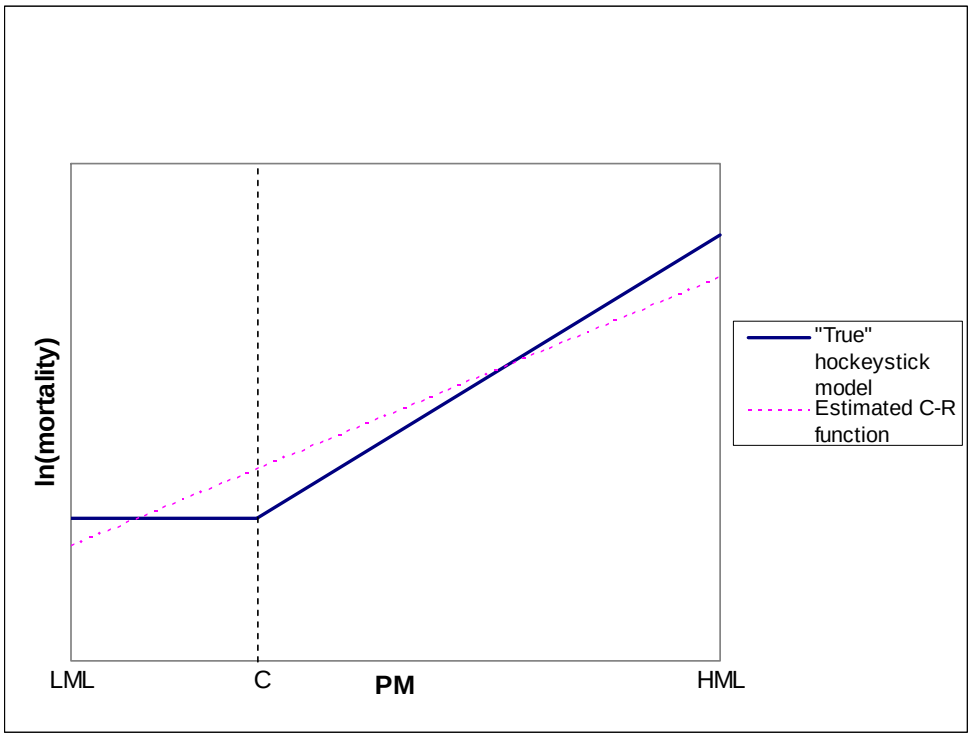


Figure H-1a. Relationship Between Estimated Log-Linear Concentration-Response Function and Hockeystick Model

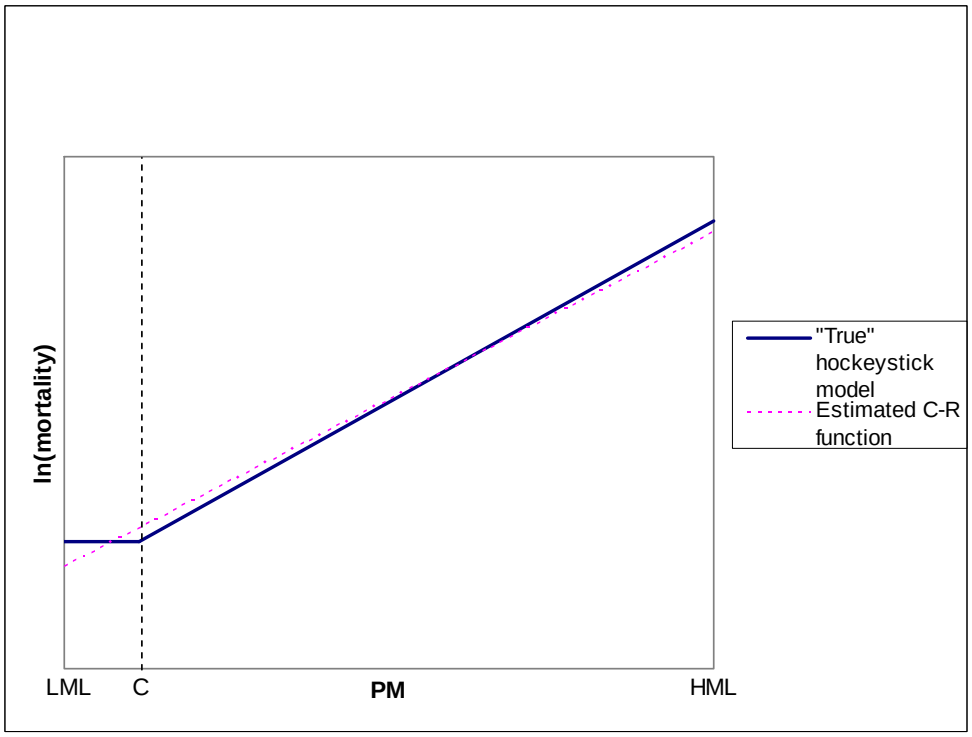


Figure H-1b. Relationship Between Estimated Log-Linear Concentration-Response Function and Hockeystick Model

If the data used in a study do not extend down below the cutpoint or extend only slightly below it, then the extent of the downward bias of the reported PM coefficient will be minimal. This is the case, for example, when the cutpoint is 10 μg/m^3 or 12 μg/m^3 for long-term exposure mortality, given that the lowest measured $PM_{2.5}$ levels in the long-term exposure mortality studies were 7.5, 10, or 11 μg/m^3. In this case, the data in the study provided hardly any information about the relationship between $PM_{2.5}$ and mortality at levels below the cutpoints and would have biased an estimate of the slope of the upward-sloping portion of a hockeystick only minimally if at all, as illustrated in Figure 2.1b.

We used a simple slope adjustment method based on the idea discussed above—that, if the data in the study were best described by a hockeystick model with a cutpoint at c, then the slope estimated in the study using a log-linear model would be approximately a weighted average of the two slopes of the hockeystick—namely, zero and the slope of the upward-sloping portion of the hockeystick. If we let

LML denote the lowest measured PM level in the study,

c denote the cutpoint,

HML denote the highest measured PM level in the study,

β^{est} denote the slope (the PM coefficient) estimated in the study (using a loglinear model), and

β^{T} denote the "true" slope of the upward-sloping portion of the hockeystick,

then, assuming the estimated coefficient reported by the study is (approximately) a weighted average of the slope below the cutpoint (0) and the slope above the cutpoint,

$$\beta^{est} = 0 * \frac{(c - LML)}{(HML - LML)} + \beta^{T} * \frac{(HML - c)}{(HML - LML)}$$

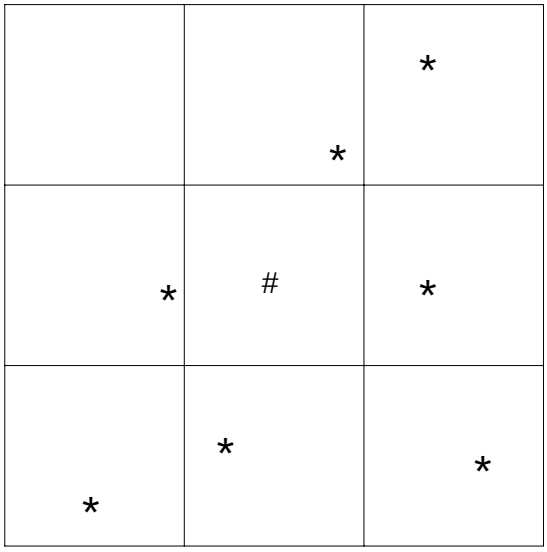
and, solving for β T,

$$\beta^{T} = \beta^{est} * \frac{(HML - LML)}{(HML - c)}$$

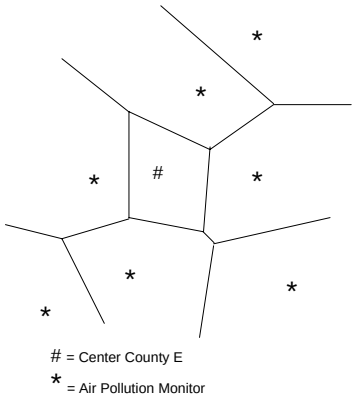
That is, the "true" slope of the upward-sloping portion of the hockeystick would be the slope estimated in the study (using a log-linear model rather than a hockeystick model) adjusted by the inverse of the proportion of the range of PM levels observed in the study that was above the cutpoint. Note that if the LML was below the estimated background level (or if it was not available for the study), the estimated background level was substituted for LML in the above equation. We believe that this slope adjustment method is a reasonable approach to estimating health effects under various assumed cutpoint models. A more definitive evaluation of the impact of alternative cutpoints and non-linear models is a subject that should be explored in further research.

H.2 Spatial Interpolation Method: Voronoi Neighbor Averaging

The first step in VNA is to identify the set of neighboring monitors for each grid cell in the Continental United States. The figure below presents nine grid cells and seven monitors, with the focus on identifying the set of neighboring monitors for grid cell E.

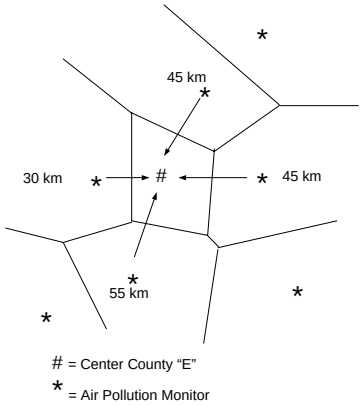


In particular, BenMAP identifies the nearest monitors, or "neighbors," by drawing a polygon, or Voronoi cell, around the center of each grid cell. The polygons have the special property that the boundaries are the same distance from the two closest points.



We then choose those monitors that share a boundary with the center of grid cell E. These are the nearest neighbors, and we use these monitors to estimate the air pollution level for this grid cell.

To estimate the air pollution level in each county, BenMAP calculates the PM metrics for each of the neighboring monitors, and then calculates an inverse-distance weighted average of the metrics. The further the monitor is from the grid cell center, the smaller the weight.



The weight for the monitor 30 kilometers from the center of grid cell E is calculated as follows:

$$weight_1 = \frac{\frac{1}{30}}{\left(\frac{1}{30}+\frac{1}{45}+\frac{1}{45}+\frac{1}{55}\right)} = 0.35$$

The weights for the other monitors would be calculated in a similar fashion. BenMAP would then calculate an inverse-distance weighted average of the nearest neighbors for grid cell E as follows:

Forecast = 0.35*80 μg + 0.23*90 μg+ 0.23*60 μg + 0.19*100 μg = 81.5 μg.

H.3 The Random/Fixed Effects Pooling Procedure

A common method for weighting estimates involves using their variances. Variance takes into account both the consistency of data and the sample size used to obtain the estimate, two key factors that influence the reliability of results. The exact way in which variances are used to weight the estimates from different studies in a pooled estimate depends on the underlying model assumed.

The fixed effects model assumes that there is a single true concentration-response relationship and therefore a single true value for the parameter β. Differences among β's reported by different studies are therefore simply the result of sampling error. That is, each reported β is an estimate of the *same underlying parameter.* The certainty of an estimate is reflected in its variance (the larger the variance, the less certain the estimate). Pooling that assumes a fixed effects model therefore weights each estimate under consideration in proportion to the *inverse* of its variance.

Suppose there are n studies, with the ith study providing an estimate β_i with variance v_i (I = 1, ..., n). Let

$$S = \sum_i \frac{1}{v_i},$$

denote the sum of the inverse variances. Then the weight, w_i, given to the ith estimate, β_i, is

$$w_i = \frac{1/v_i}{S}$$

This means that estimates with small variances (i.e., estimates with relatively little uncertainty surrounding them) receive large weights, and those with large variances receive small weights.

The estimate produced by pooling based on a fixed effects model, then, is just a weighted average of the estimates from the studies being considered, with the weights as defined above. That is,

$$\beta_{fe} = \sum_i w_i \beta_i$$

The variance associated with this pooled estimate is the inverse of the sum of the inverse variances:

$$v_{fe} = \frac{1}{\sum_i 1/v_i}$$

An alternative to the fixed effects model is the random effects model, which allows the possibility that the estimates β_i from the different studies may in fact be estimates of *different* parameters, rather than just different estimates of a single underlying parameter. In studies of the effects of ozone on mortality, for example, if the level of air conditioning use varies among study locations the underlying relationship between mortality and ozone may be different from one study location to another. If air conditioning use causes individuals to stay inside more on days with high ozone, then the mortality risk may be lower in areas with high prevalence of air conditioning. As such, one would expect the true value of β in cities with low air conditioning prevalence to be greater than the true value of β in cities with high air conditioning prevalence. This would violate the assumption of the fixed effects model.

The following procedure can test whether it is appropriate to base the pooling on the random effects model (vs. the fixed effects model):

A test statistic, Q_w, the weighted sum of squared differences of the separate study estimates from the pooled estimate based on the fixed effects model, is calculated as:

$$Q_w = \sum_i \frac{1}{v_i} \left(\beta_{fe} - \beta_i\right)^2$$

Under the null hypothesis that there is a single underlying parameter, β, of which all the β_i 's are estimates, Q_w has a chi-squared distribution with n-1 degrees of freedom. (Recall that n is the number of studies in the meta-analysis.) If Q_w is greater than the critical value corresponding to the desired confidence level, the null hypothesis is rejected. That is, in this case the evidence does not support the fixed effects model, and the random effects model is assumed, allowing the possibility that each study is estimating a different β.

The weights used in a pooling based on the random effects model must take into account not only the within-study variances (used in a meta-analysis based on the fixed effects model) but the between-study variance as well. These weights are calculated as follows:

Using Q_w, the between-study variance, η^2, is:

$$\eta^2 = \frac{Q_w - (n-1)}{\sum 1/v_i - \frac{\sum 1/v_i^2}{\sum 1/v_i}}$$

It can be shown that the denominator is always positive. Therefore, if the numerator is negative (i.e., if $Q_w < n-1$), then η^2 is a negative number, and it is not possible to calculate a random effects estimate. In this case, however, the small value of Q_w would presumably have led to accepting the null hypothesis described above, and the meta-analysis would be based on the fixed effects model. The remaining discussion therefore assumes that η^2 is positive.

Given a value for η^2, the random effects estimate is calculated in almost the same way as the fixed effects estimate. However, the weights now incorporate both the within-study variance (v_i) and the between-study variance (η^2). Whereas the weights implied by the fixed effects model used only v_i, the within-study variance, the weights implied by the random effects model use $v_i + \eta^2$.

Let $v_i^* = v_i + \eta^2$. Then

$$S^* = \sum_i \frac{1}{v_i^*}$$

and

$$w_i^* = \frac{1/v_i^*}{S^*}$$

The estimate produced by pooling based on the random effects model, then, is just a weighted average of the estimates from the studies being considered, with the weights as defined above. That is,

$$\beta_{rand} = \sum_i w_i^* \times \beta_i$$

The variance associated with this random effects pooled estimate is, as it was for the fixed effects pooled estimate, the inverse of the sum of the inverse variances:

$$v_{rand} = \frac{1}{\sum_i 1/v_i^*}$$

The weighting scheme used in a pooling based on the random effects model is basically the same as that used if a fixed effects model is assumed, but the variances used in the calculations are different. This is because a fixed effects model assumes that the variability among the estimates from different studies is due only to sampling error (i.e., each study is thought of as representing just another sample from the same underlying population), while the random effects model assumes that there is not only sampling error associated with each study, but that there is also *between-study* variability—each study is estimating a different underlying β. Therefore, the sum of the within-study variance and the between-study variance yields an overall variance estimate.

Weights can be derived for pooling incidence changes predicted by different studies, using either the fixed effects or the random effects model, in a way that is analogous to the derivation of weights for pooling the β's in the C-R functions. For a given change in pollutant level and a given baseline incidence rate, corresponding to every possible value of β, there is an incidence change. Corresponding to β_i, with variance v_i (calculated from the reported standard error of β_i,) from the ith study, there is therefore an estimate of incidence change, I_i, with variance $v(I)_i$. In practice, we generate a sample mean and a sample variance of incidence changes by calculating an incidence change for each of many β's pulled from the distribution of β's for the study.

This can be done either using Monte Carlo methods (making many random pulls) or by a Latin Hypercube approach, in which we pull the nth percentile β from the distribution of β's, for, e.g., n = 2.5, 7.5, ..., 97.5. Either way, the result is a corresponding sample distribution of incidence changes that would be predicted by the study, from which we calculate the sample mean and the sample variance. The sample means of incidence change from the studies to be pooled are used in exactly the same way as the reported β's are used in the discussion of fixed effects and random effects models above. The sample variances of incidence change are used in the same way as the variances of the β's. The formulas above for calculating fixed effects weights, for testing the fixed effects hypothesis, and for calculating random effects weights can all be used by substituting the sample mean incidence change for the ith study for β_i and the sample variance of incidence change for the ith study for v_i.