User's Manual

Appendices

September 2008

Prepared for
Office of Air Quality Planning and
Standards
U.S. Environmental Protection Agency
Research Triangle Park, NC
Neal Fann, Project Manager

Prepared by
Abt Associates Inc.

Table of Contents

Appendix B: Monitor Rollback Algorithms   191

Appendix C: Air Pollution Exposure Estimation Algorithms   207

Appendix D: Deriving Health Impact Functions   227

Contents

Appendix J: Uncertainty & Pooling   362

Appendix K: Command Line BenMAP   374

Appendix L: Function Editor   387

References 392

Appendix A: Training Courses

BenMAP is intended primarily as a tool for estimating the health impacts, and associated economic values, associated with changes in ambient air pollution. It accomplishes this by running health impact functions, which relate a change in the concentration of a pollutant to a change in the incidence of a health endpoint. Inputs to health impact functions typically include:

(a) the change in ambient air pollution level,

(b) a health effect estimate,

(c) the background incidence rate of the health endpoint, and

(d) the exposed population.

For example, in the case of a premature mortality health impact function, we might have the following:

Mortality Change = Air Pollution Change × Mortality Effect Estimate

× Mortality Incidence × Exposed Population

where each term on the right is defined as follows:

Air Pollution Change. The air quality change is calculated as the difference between the starting air pollution level, also called the baseline, and the air pollution level after some change, such as that caused by a regulation. In the case of particulate matter, this is typically estimated in micrograms per cubic meter (µg/m3).

Mortality Effect Estimate. The mortality effect estimate is an estimate of the percentage change in mortality due to a one unit change in ambient air pollution. Epidemiological studies provide a good source for effect estimates.

Mortality Incidence. The mortality incidence rate is an estimate of the average number of people that die in a given population over a given period of time. For example, the mortality incidence rate might be the probability that a person will die in a given year. Mortality incidence rates and other health data are typically collected by each country's government. In addition, the World Health Organization is a good source for data.

Exposed Population. The exposed population is the number of people affected by the air pollution reduction. The government census office is a good source for this information. In addition, private companies may collect this information and offer it for sale.

BenMAP also calculates the economic value of health impacts. For example, after the calculation of the mortality change, you can value the premature deaths avoided by multiplying the change in mortality by an estimate of the value of a statistical life:

Value Mortality = Mortality Change × Value of Statistical Life

where the rightmost term is defined as follows:

Value of Statistical Life. The value of a statistical life is the economic value placed on eliminating the risk of one premature death.

Figure 1-1 provides the overall schematic of BenMAP, and the various major steps involved in using it. This figure also highlights that BenMAP does not have air quality modeling capabilities, and instead relies on modeling and monitoring inputs.
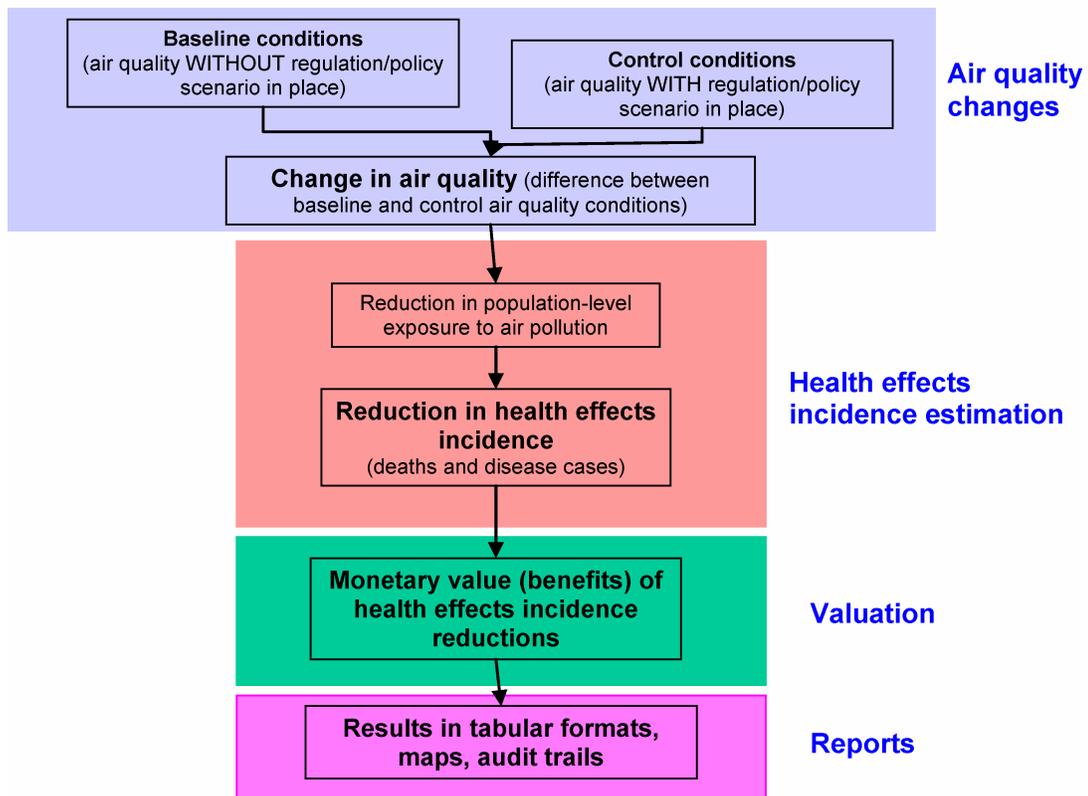
Figure 1-1. Overall schematic of BenMAP.

BenMAP also serves as a Geographic Information System (GIS), allowing users to create, utilize, and visualize maps of air pollution, population, incidence rates, incidence rate changes, economic valuations, and other types of data. BenMAP can thus be used for a variety of purposes, including:

Generating population/community level ambient pollution exposure maps;

Comparing benefits associated with regulatory programs;

Estimating health impacts and costs of existing air pollution concentrations;

Estimating health benefits of alternative ambient air quality standards; and

Performing sensitivity analyses of health or valuation functions, or of other inputs.

A wide range of people can use BenMAP, including scientists, policy analysts, and decision makers. Advanced users can explore a broad array of options, such as using the map querying features and exploring the impacts of different health impact and valuation functions.

A.1   United States

This training is intended for a beginning user of BenMAP that wants to use the U.S. setup that comes with BenMAP. The training manual is organized into seven sections.

Section 1. Data Files Needed. This provides a list of the files that you will need for this training and where you can get them.

Section 2. Mapping. This module is designed to help you become comfortable using BenMAP by analyzing the various types of data that we will use throughout this course.

Section 3. One-step Analyses. You will create health incidence and valuation results. This one-step analysis uses pre-defined health incidence and economic valuation configurations, which you can combine with your own baseline and control air quality grids.

Section 4. Creating Grids. In this section, you will create air quality grids (aqg) from both air quality (AQ) model data and from monitor data. The overall goal will be to produce baseline and control air quality grids for later estimations of health incidence and valuation.

Section 5. Health Incidence Estimation. In this section, you will modify an existing health incidence configuration and use it to create new health incidence results. You will create two separate sets of health incidence results based on the same configuration and two similar, but distinct, control strategies.

Section 6. Aggregation, Valuation, and Pooling. In this section, you will create an aggregation, pooling, and valuation (APV) configuration and use it to produce new valuation results. You will create two separate sets of valuation results based on the same configuration and two similar, but distinct, control strategies.

Section 7. Adding New Datasets and an Independent Case Study. In this section, you will add new datasets to BenMAP and run a local-scale benefit analysis in Detroit, Michigan.

We have broken down each of the above sections into various modules, and each module provides a context section, a module goal, and one or more example BenMAP applications. In addition, there is a section with answers to embedded exercise questions for Sections 2 through 7.

A.1.1   Section 1. Data Files Needed for Training

A range of data files are needed for this training and can be accessed at the BenMAP website: http://www.epa.gov/air/benmap/.

A.1.2   Section 2. Mapping Introduction

This module is designed to help you become comfortable using BenMAP by analyzing the various types of data that you will use throughout this course. In this lesson, you will focus on using BenMAP's GIS tool to view maps of air quality data (file names ending in .aqg),

health incidence results (ending in .cfgr), and economic valuation results (ending in .apvr).

A.1.2.1 Raw Monitor Data

Several years of recent monitor data for Ozone, PM10 and PM2.5 have been pre-loaded into the BenMAP database. These data are derived from the extensive network of monitoring sites throughout the U.S. For the PM (particulate matter) data, most of the monitor data is recorded on a daily basis. The ozone measurements are recorded on an hourly basis. When displaying the data for a pollutant, you can display different averaging techniques, typically referred to as air quality metrics.

A.1.2.1.1   Example PM2.5 monitor data for 2000

Goal: To start learning about BenMAP and the GIS tool, and to explore PM2.5 monitor data for the year 2000. You will be working with monitor values preloaded into BenMAP's underlying database

(a) Open BenMAP by clicking on the desktop icon or by choosing "Launch BenMAP 3" from the Window's Start menu.   This will bring up the main BenMAP window (Figure 2-1).
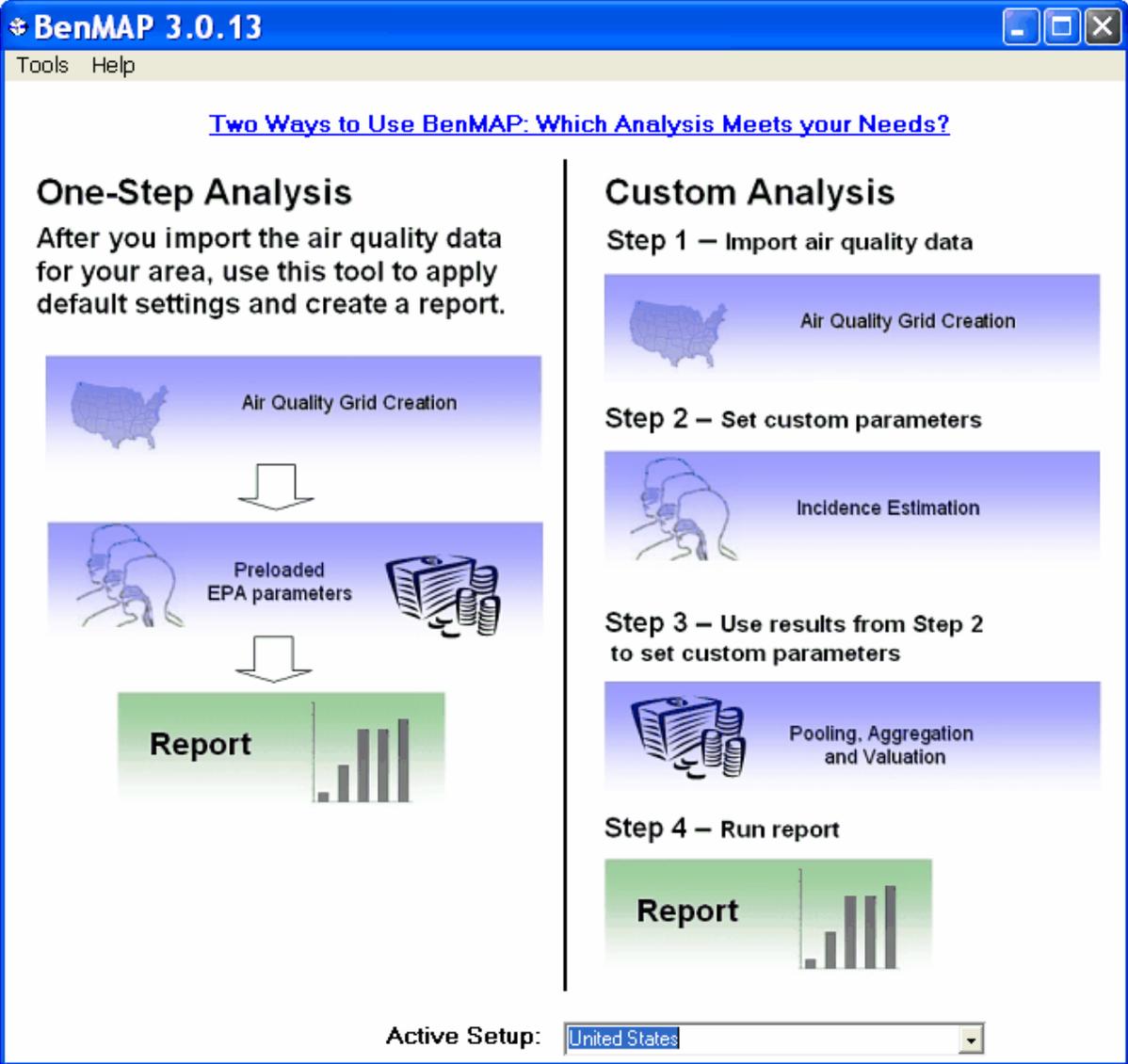
Figure 2-1. Main BenMAP window.

(b) From the Tools menu, choose "GIS/Mapping" (Figure 2-2).
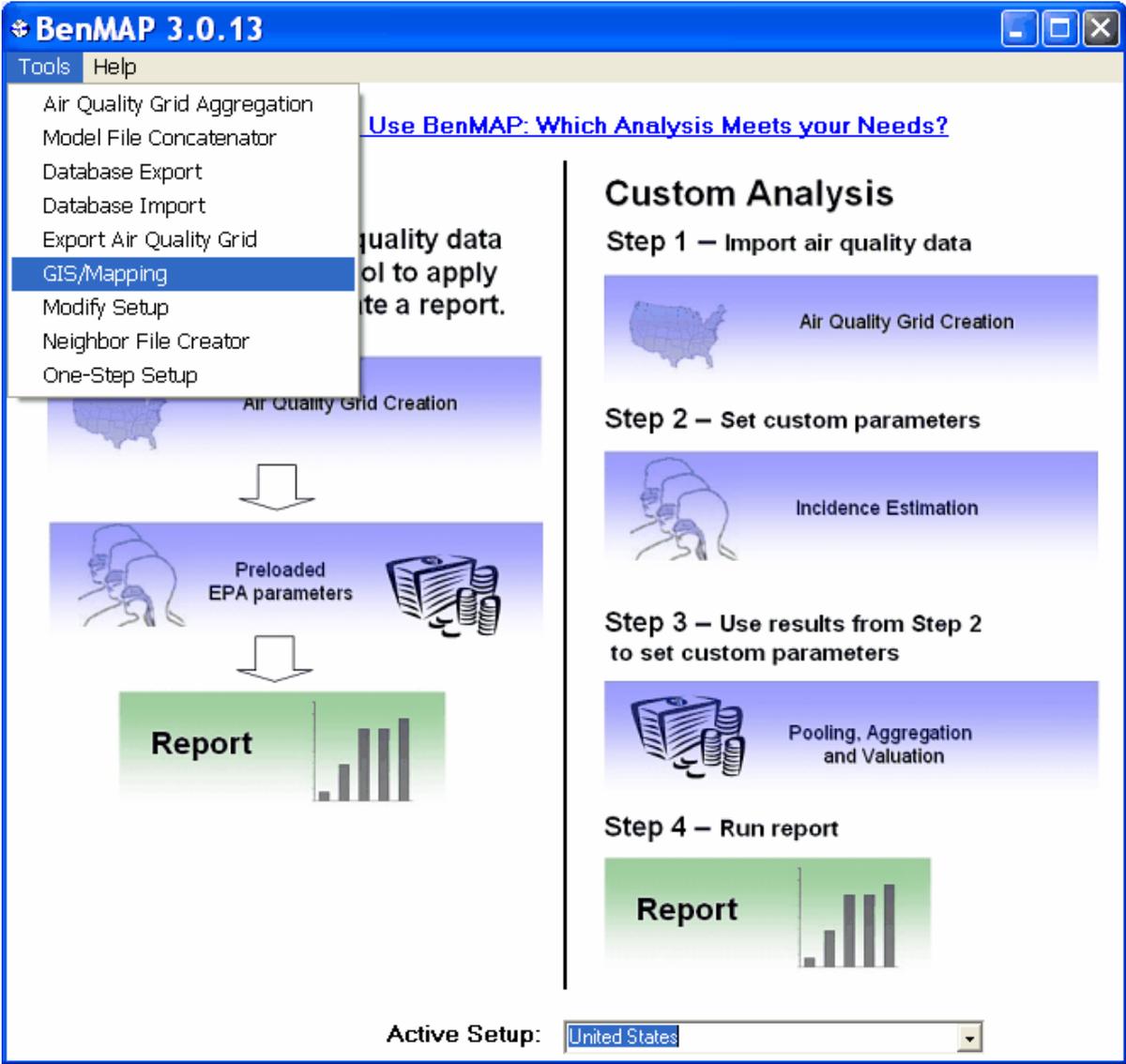
Figure 2-2. Opening the GIS tool.

This will open the BenMAP GIS window (an example is shown later in Figure 2-4). At the top of the GIS window, you will see a series of buttons, described below. Note that in this terminology a layer is a map, and "active layer" means the topmost map in the GIS window. You can have multiple maps layered on top of each other in the GIS tool.

Open a File

Save active layer

Zoom to full extent

Zoom in, Zoom out

Select a region to zoom into

Drag map

Display info - Displays information for all variables in active layer of that cell

Build queries to limit data displayed (e.g. values > 80 ppb)

Calculate layer statistics of the active layer

(c) Click on the Open a File button and choose "Monitors". In the Select Monitors window (Figure 2-3), set the pollutant to "PM2.5", and under the library tab set the dataset to "EPA Standard Monitors" and year to "2000".

Although you will not use the Advanced button in this module, we could use it to further limit the monitors that we would display based on location, state, or other monitor criteria.
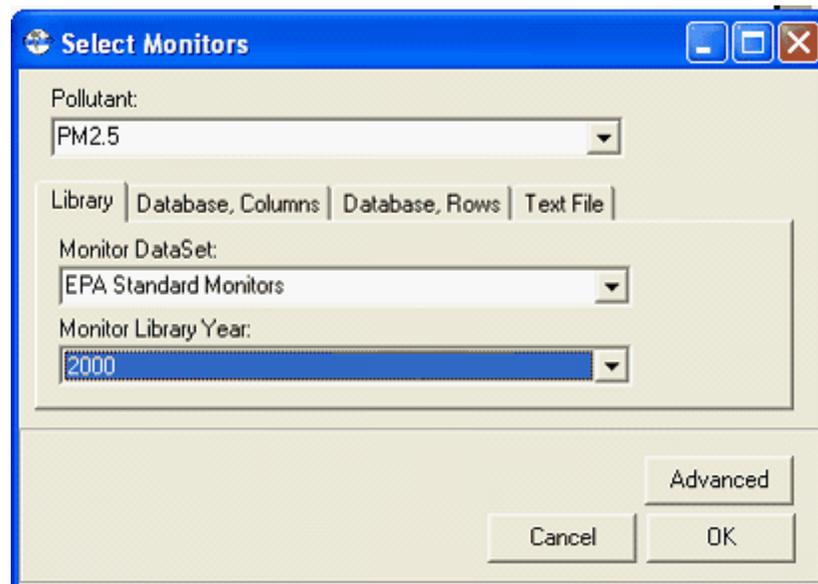
Figure 2-3. *Select Monitors* **window.**

(d) Click the **OK** button. This will bring up the monitor locations on the *BenMAP GIS* window (Figure 2-4).
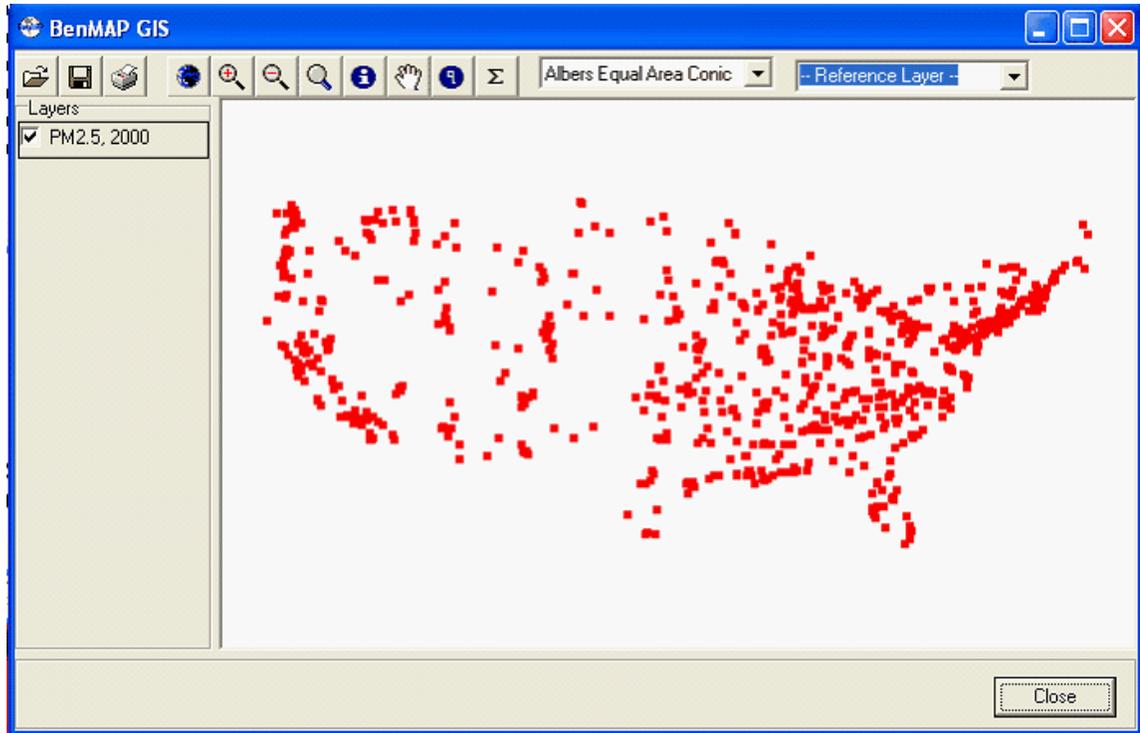
Figure 2-4. PM$_{2.5}$ monitors locations.

(e) Double-click on the "PM2.5, 2000" layer. This will open the Display Options window (Figure 2-5).

Set the variable to "QuarterlyMean", which represents the average of the four quarterly means (i.e., the annual average).

Set start and end sizes to "100"; this is the size of the monitor points.

Do not change the other seven fields in the window. The min and max values define your range. You can edit this to narrow in on a specific range. The start and end color allow you to pick the colors of the monitor points in your range. The default size and default color are for areas that are outside your range. The decimal digits are the number of decimal digits displayed.
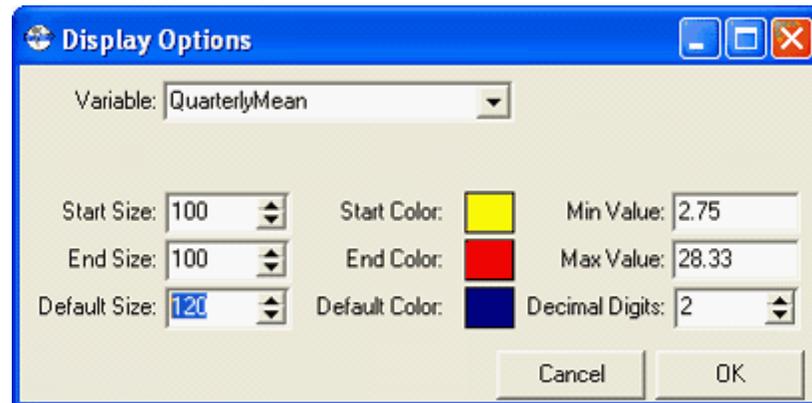
Figure 2-5. *Display Options* **window.**

> *Background*: The main PM metrics are D24HourMean (the daily value) and QuarterlyMean (the mean of all days within an individual yearly quarter). The GIS tool does not display the monitor values for particular day or quarter; rather, the GIS tool will show seasonal averages of these variables. For PM, the default "season" is typically the full year. This will likely cause some confusion. When we display D24HourMean, we are actually displaying the average of the 365 days of data, and when we display QuarterlyMean we are showing the average of the 4 quarterly means. In short, both these variables should be thought of as more akin to an annual average. Note: the definition of metrics and seasons can be changed (discussed in Lab 7).

(f) After clicking **OK**, you will get a map of the annually averaged $PM_{2.5}$ values at all the monitor locations. To gain a better sense of the monitor locations, use the "-- Reference Layer --" drop down menu in the top right corner of the *BenMAP GIS* window to select the "County" overlay (Figure 2-6). The reference layer overlays a specific grid type (e.g. county, state, CMAQ) on top of your data layer. It provides geographic context to your data layer.
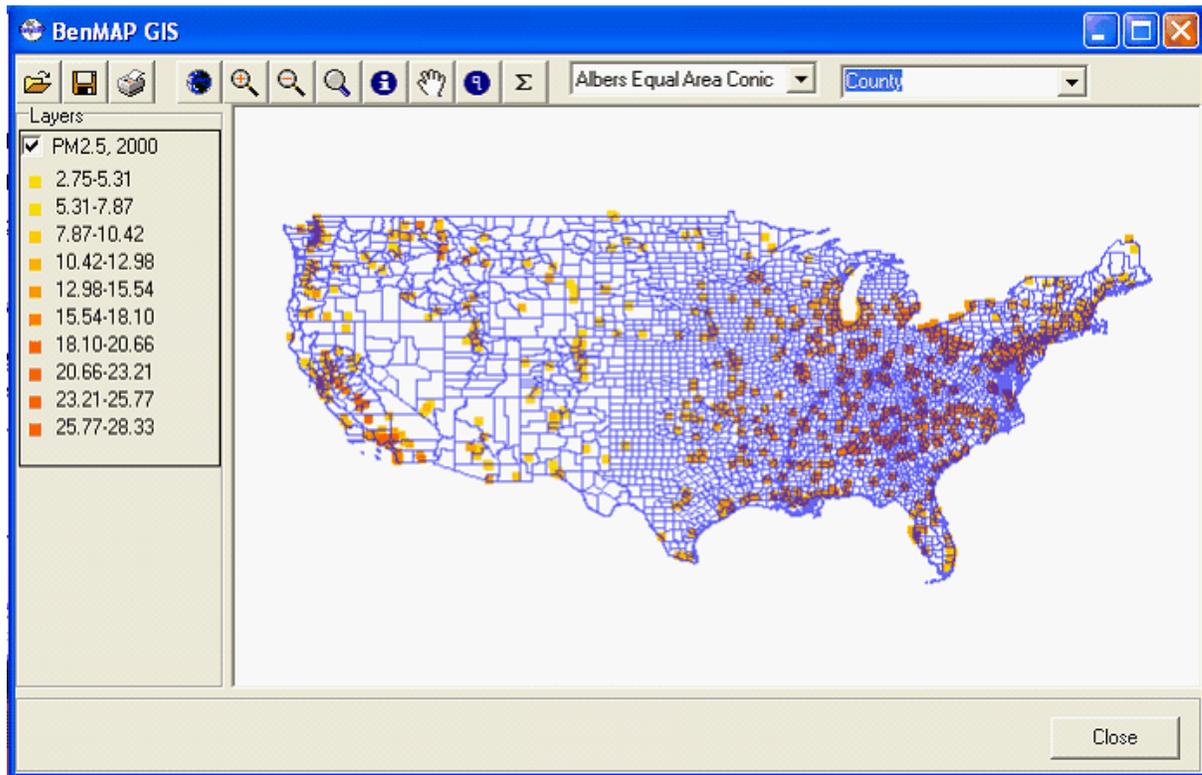
Figure 2-6. PM$_{2.5}$ monitor locations overlaid with county boundaries.

(g) Now overlay a State reference layer. Experiment with zooming in and out of the map (using the toolbar). Try out some of the other buttons, including **Display info** and **Create layer statistics**. *Note*: the values are in micrograms per cubic meter (μg/m³).

(h) **Exercise (2.1):** What are the D24HourMean, QuarterlyMean, and lat/lon of the monitor at the northern tip of Maine? *Hint:* Use the **Display info** and **Zoom** buttons.

Answer:

(i) Use the **Build queries** button to bring up the *Build Query* window (Figure 2-7). Use the fields list and mathematical operator buttons or simply type in that window to create a query that limits the monitors to those with QuarterlyMean less than 10 micrograms per cubic meter (μg/m³). Click **OK** or **Execute.**
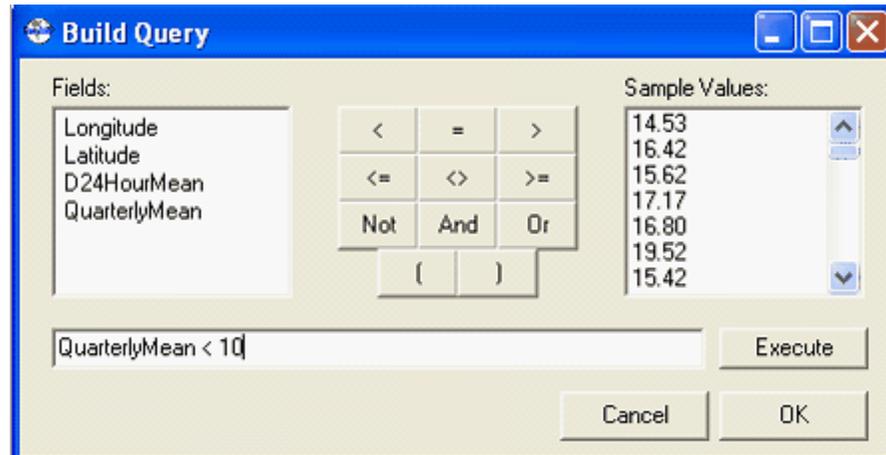
Figure 2-7. *Build Query* window.

Tip: You construct a query in the query text field (indicated by red arrow above). You can type in the field name (or double-click it, e.g. "QuarterlyMean"), the values, and the operators (e.g. ">" or "=") in the query text field. To remove the query (i.e. see all your monitors), delete the query text field and click **OK** (or **Execute**). **Execute** is the same as **OK** except that it keeps the *Build Query* window open.

(j) **Exercise (2.2):** What states have monitors with QuarterlyMean values >20 µg/m³?

Answer:

(k) Remove the query so that you can see all the monitors. **Do not close** the *BenMAP GIS* window, because it will be used again in the next example.

A.1.2.1.2   Example O3 monitor data for 2000

The goal of this task is to learn about layers in the GIS tool and to explore the O3 monitor data for 2000.

(a) Using the same GIS window, open a second dataset by clicking on Open a File and selecting "Monitors". This time, set the pollutant to "Ozone". As before, set the year to "2000". Click OK.

(b) Uncheck "PM2.5" in the Layers panel by clicking in the checkbox. The PM2.5 monitors should disappear.

(c) Double-click on the "Ozone, 2000" layer to open up the Display Options window for that layer (Figure 2-8). For the variable, select "D8HourMax", which is the average of the 8 hr maximum

window for each day in the ozone season.

Change the "Start Size" and "End Size" to "100" and colors from dark blue (Start Color) to light blue (End Color). In other words, we are differentiating the ozone monitors from the PM2.5 monitors by having them range from dark to light blue. Click OK. Note: these values are the concentrations of ozone in parts per billion (ppb).
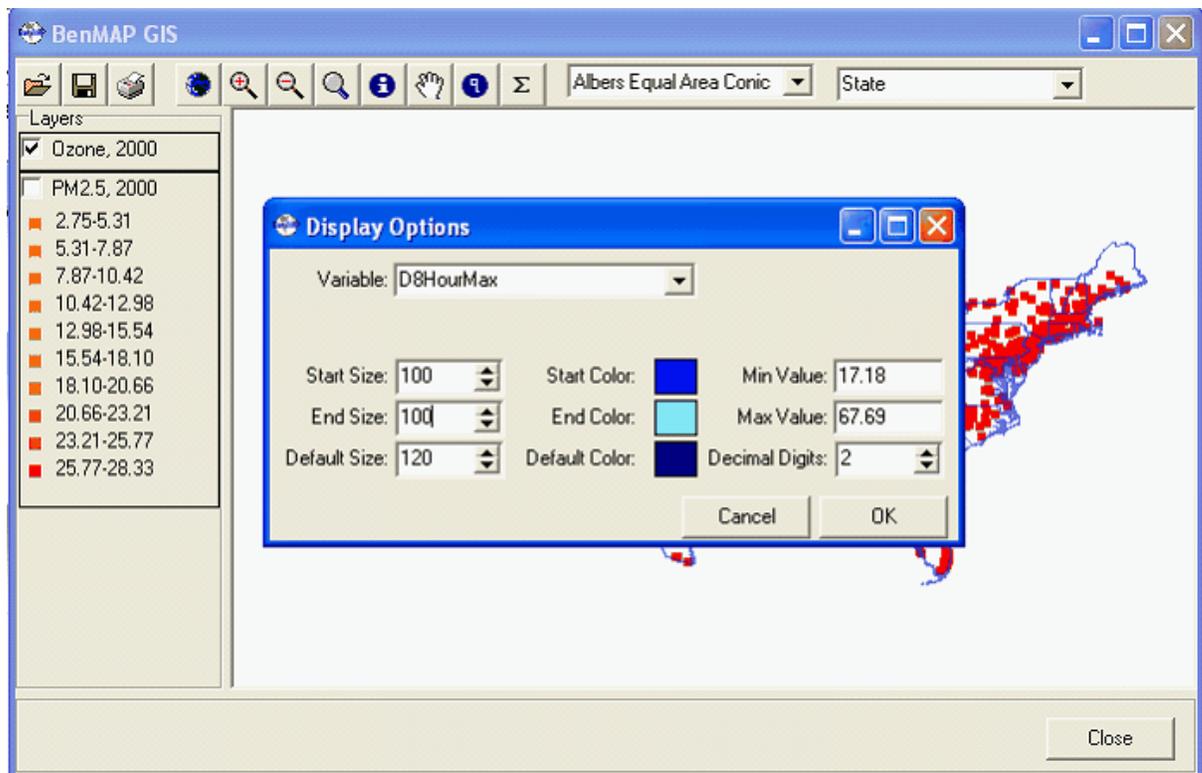
Figure 2-8. *Display Options* **window values described in step (c) above.**

Background: There are a series of ozone metrics: D1HourMax (the maximum 1 hour value in a day), D24HourMean (daily mean), D5HourMean (daily mean of hours 10am - 2pm), D8HourMax (the greatest mean for any 8 hour window in a day), and D8HourMean (daily mean of hours 9am - 4pm). Again, the GIS tool does not display the monitor values for any particular day. It calculates and displays a seasonal average for each of the above metrics. The default ozone season is from May 1st through September 30th. For example, the D1HourMax is calculated by adding up the maximum value for that monitor for each day in the season and then dividing by the number of days in the ozone season. Note: the definition of metrics and seasons can be changed (discussed in Lab 7).

(d) Recheck "PM2.5, 2000" in the Layers panel. You should now see both the O3 and PM2.5 data.

In the Layers panel, switch the order of the layers by right-clicking on the "PM2.5, 2000" layer and select "Move up" in the pop-up window. The active layer is always the topmost layer in the Layers panel. Note: Only the active layer is used in getting information or performing queries.

(e) Exercise (2.3): What are the maximum and minimum QuarterlyMean values for the PM2.5 monitors? Hint: use the Layer Statistics button and PM2.5 should be the active layer.

> *Answer:*

(f) **Exercise (2.4):** What are the maximum and minimum D8HourMax values for the ozone monitors? Which states have a D8HourMax greater than 60 parts per billion (ppb)? *Hint*: ozone should be the active layer. When performing the query, you might want to uncheck $PM_{2.5}$ so that it is easier to see the ozone monitors.

> *Answer:*

(g) After you are done, click **Close** at the bottom right of the GIS window. This completes the "Raw Monitor Data" module for this lab.

A.1.2.2 Model Data

You will use CMAQ (an AQ model that simulates the chemistry and movement of various pollutants) outputs from the PM2.5 Regulatory Impact Analysis (RIA) for this training module. Unlike the monitor data visualized in the previous module, the model data has values that are on a regular grid and cover the whole area of the map. The RIA model data is a forecast of the air quality (AQ) for the year 2020. We will focus on a baseline scenario (think of this as "business-as-usual") and two control scenarios (in these cases additional regulations have been applied to emission source, resulting in generally lower pollution levels).

We recommend that you are detailed and consistent in naming your BenMAP files. In this lab, the file name includes references to the annual PM2.5 NAAQS (National Ambient Air Quality Standards) and a daily PM2.5 NAAQS. For example, the file "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg" refers to the baseline model run on the CMAQ grid type in which most annual PM2.5 values are below 15 µg/m3 and most daily values are below 65 µg/m3. For the rest of the modules, we will use the shorthand "Baseline PM2.5 RIA 2020 15/65" to refer to the baseline scenario with 15 µg/m3 annual NAAQS and 65 µg/m3 daily NAAQS. An equivalent shorthand will be used

for the control scenarios.

The goal of this example is to look at model data using the BenMAP GIS tool and to learn about the differences between political-type grids and CMAQ-type grids.

(a)From the main BenMAP window, open a new GIS window by choosing "GIS/Mapping" in the Tools menu.

(b)Click on the Open a File button, then choose "Air Quality Grid (*.aqg)". Under the "Air Quality Grids" folder in the Open an Air Quality Grid window, navigate to the folder "PM25_RIA", select the file named "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg", and click Open.

(c)In the GIS window that appears, double-click on the "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg" layer to open the Display Options window.

In that window, set the variable to "QuarterlyMean". This will cause the annual average data to be displayed.

Also uncheck the grid outline; this is usually preferred, because the window often looks messy with both the data and the grid outlines displayed. Click OK in the Display Options window.

Finally, use the drop-down menu near the upper right corner of the GIS window to overlay a "State" reference layer (Figure 2-9).
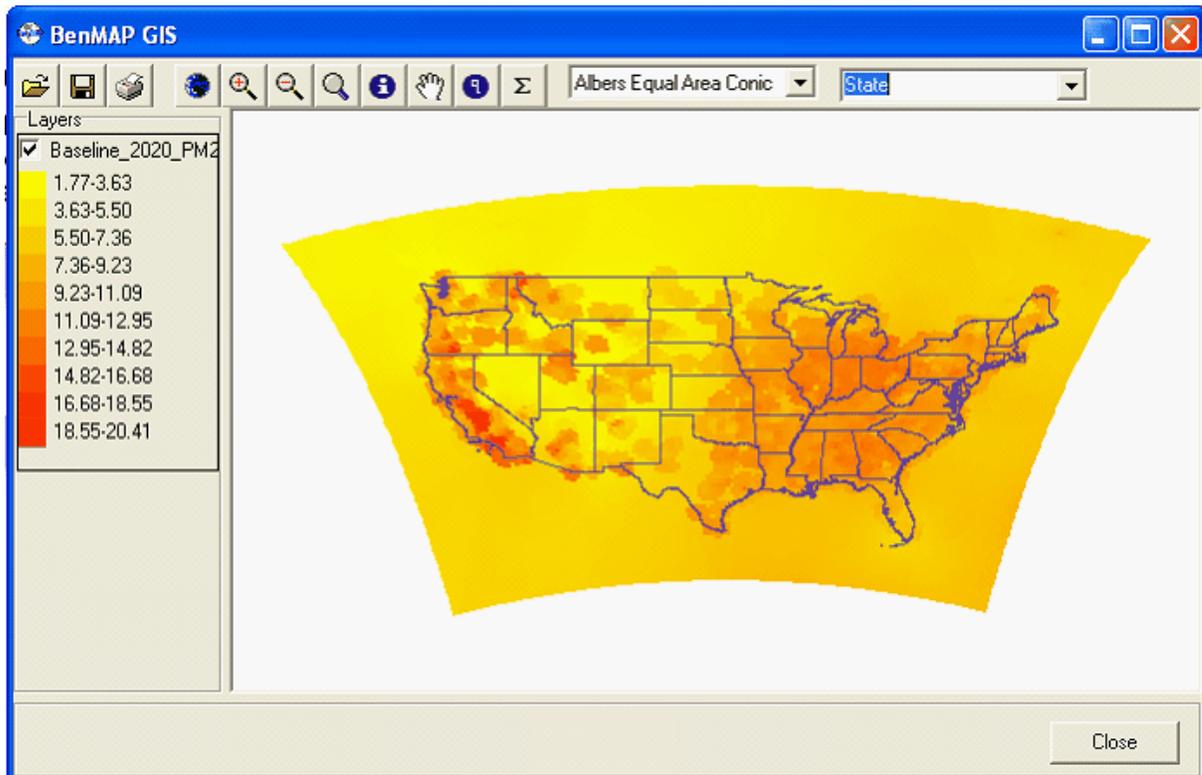
Figure 2-9. PM$_{2.5}$ model output overlaid with state boundaries.

(d)**Exercise (2.5):** Use the **Zoom in** button to zoom into a state border region until you can see the model grid cells. Do the grid cells align with the state boundaries? In other words, do the model grid cells perfectly fit within the political boundaries?

> *Answer:*

(e)**Exercise (2.6):** What states are out of attainment in this baseline model scenario? States that are out of attainment are those that have at least one QuarterlyMean grid cell value above 15 µg/m^3.

> *Answer:*

(f) Try overlaying the "CMAQ 36km Nation Overlap" overlay reference layer. Zoom into a small region. Note how the model values line-up with the reference layer.

(g) When you are done exploring, close the GIS window by clicking **Close**. This completes the "Model Data" module.

A.1.2.3 Health Incidence

To produce health incidence results, the first step is to calculate the change in pollution concentrations that would be produced by the application of a given set of emissions controls. The concentration change in a pollutant (say, PM2.5) is the difference (the "delta") between the modeling results from a control scenario and the modeling results from the baseline scenario. These deltas and a gridded population dataset are then used in concentration-response (C-R) functions to calculate the change in health incidence that would result from this change in pollution. These C-R functions are based on epidemiological studies and can be selected by the user (see Lab 5). Typically, these health incidence results show the number of avoided health incidence (e.g. the decrease in asthma, bronchitis, mortality, etc) due to a decrease in pollution.

In the rest of this module, we refer to the health incidence results via the shorthand versions of their control scenario names (as explained in Section 2.2.1). Also recall that the abbreviation "cfgr" refers to health incidence results. Note: we don't actually go through the procedure of creating these health incidence results in this lab (see Labs 3 and 5); rather, we are just looking at the pre-computed results.

A.1.2.3.1   Example Configuration Results File: Control PM2.5

The goal of this exercise is to use the GIS tool to explore the reductions in health incidence that would be due to the reductions in PM2.5 caused by the RIA control scenario. In particular, we will look at reductions in mortality, acute respiratory symptoms, chronic bronchitis, and emergency room (ER) visits.

(a) From the main BenMAP window, open a new GIS window.

(b) Click on the Open a File button, then select "Configuration Results (*.cfgr)". Under the "Configuration Results" folder, select the file named "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily.cfgr". Click Open.

(c) This will bring up an additional window Edit GIS Field Names. In the last column of this window, change the names of the fields to more meaningful names, by highlighting the contents of each of the four cells and typing in "Mortality", "ChronBronc", "ER", and "AcutResp" (Figure 2-10). Note: the GIS field names cannot exceed 10 characters in length.

Before clicking OK, do the next exercise.

Figure 2-10. *Edit GIS Field Names* **window.**

(d) **Exercise (2.7):** Look at the various columns in the Edit GIS Field Names window (information and variables used in the C-R functions, e.g. "Pollutant" and "Author"). For what age range are we calculating the change in mortality? When finished with this exercise, click OK.

> Answer:

(e) In the BenMAP GIS window, display mortality data, uncheck the CMAQ grid outline, and overlay a state reference layer. Note: these values are number of deaths prevented by the control scenario.

(f) **Exercise (2.8):** Under this control strategy, which states had more than 25 avoided deaths?

> Answer:
>
>
>
>

(g) Display acute respiratory symptoms, "AcutResp".

(h) **Exercise (2.9)**: How do the number of acute respiratory incidences avoided compare to the number of premature deaths avoided? Hint: use the Display info or Calculate layer statistics buttons.

> Answer:
>
>
>
>

(i) Display and explore the GIS fields DELTAX and POPX. Here, "X" refers to a number.

> Background: For each health incidence result, there is a corresponding population and delta variable. For example, DELTA0 and POP0 are the PM2.5 delta and the population that mortality was calculated from. The number comes from the initial result names.   For example, Result0 (mortality) matches DELTA0 and POP0; Result1 (ChronBronc) matches DELTA1 and POP1; etc. Note, the population of interest is determined by the C-R function and is not necessarily equivalent between the various health incidence results (e.g. different age ranges, gender, or ethnic groups).

(j) **Exercise (2.10):** Compare the delta and mortality values. Look at the spatial pattern of these two variables. Why do the delta and the mortality values not exactly correlate?

> Answer:
>
>
>
>

(k) Close the GIS window. This completes the "Health Incidence" module.

A.1.2.4 Valuation Map

For the purposes of this course, "valuation results" are the economic benefits of avoiding the previously calculated changes in health incidence. These are the monetized benefits of the avoided premature mortalities and morbidities (i.e., non-fatal health effects), which are the two measures we discuss in this module. These valuation results are calculated by taking the estimates of health incidence avoided and applying valuation functions to express those incidences in dollar terms. For example, the economic value of premature mortality avoided has a monetary value that is expressed as a distribution. If you calculate a mean value of that distribution, you can then multiply that mean value by the number of mortalities avoided (due to your control strategy), resulting in an estimate of the economic benefit of avoiding those premature mortalities.

The valuation results (and the underlying health incidence) are typically aggregated from the CMAQ or county grid to State or National totals. The results (health incidence and valuation) are often pooled. Pooling are methods for combining similar health incidence or valuation results. For example, if you have two different studies (valuation functions) for calculating the monetized benefit of avoided ER visits, you would "pool" together these two results to create a single valuation for ER (see Lab 6).

As in the "Health Incidence" module just completed, here we refer to the valuation results via the shorthand versions of their control scenario names. (Recall that the abbreviation "apvr" refers to economic valuation results.)

A.1.2.4.1   Example Valuation Results File: Control PM2.5

The goal of this exercise is to use the GIS tool to explore the economic benefits of the above reductions in health incidence due to this control scenario. In particular, we will look at the cost savings due to reductions in mortality and in morbidity. We will also compare the economic benefits to the pooled and aggregated health incidence results.

(a) Open a new GIS window, click on the Open a File button, select "APV Configuration Results (*.apvr)", then select "Pooled Valuation Results" (Figure 2-11).
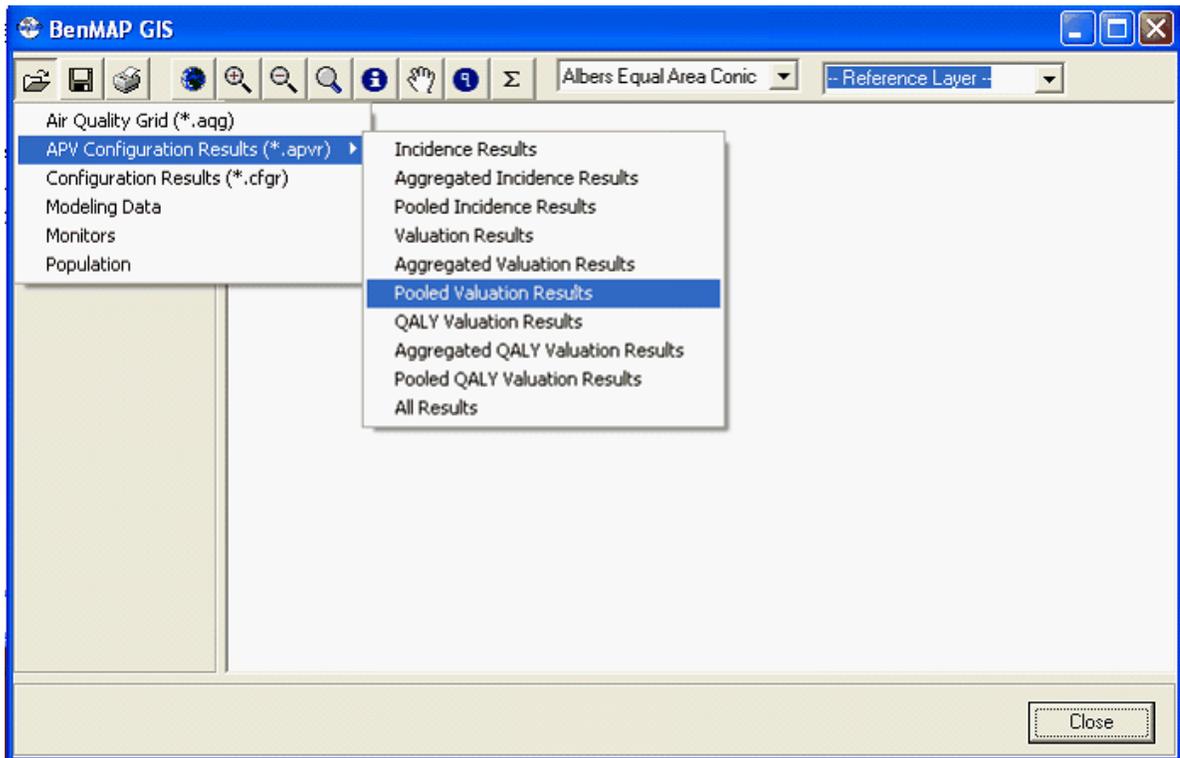
Figure 2-11. Selecting the "Pooled Valuation Results" file.

Tip: There are a lot of options for what type of data to map with apvr files. Typically we use "Pooled Incidence Results" (aggregated and pooled health incidence) or "Pooled Valuation Results" (aggregated and pooled valuations).

(b) In the *Open an APV Configuration Results File* window, under the folder "Configuration Results", select the file named "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr". Click Open.

(c) In the Edit GIS Field Names window that appears, edit the GIS field names. Use the same four field names as in the previous module ("Mortality", "ChronBronc", "ER", and "AcutResp").

(d) After clicking OK, a Valuation Sums Layer window (Figure 2-12) will appear. This window is the starting point for adding together various valuation results to get total values for mortality and morbidity. We will view these summed results in the GIS window.
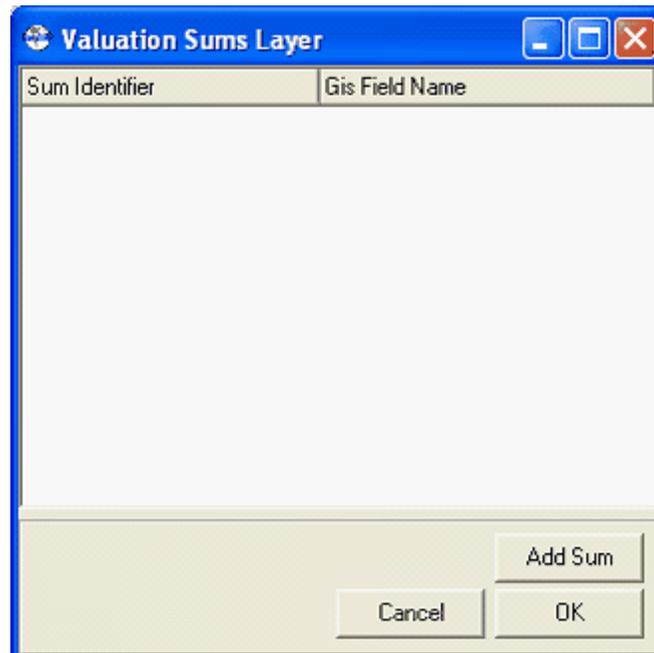
Figure 2-12. *Valuation Sums Layer* **window.**

(1) To create a morbidity layer, click on the Add Sum button. An Add Valuation Sum window will appear (Figure 2-13).

In the Include in Total column at the right, check all the health incidences that will be summed together to make   morbidity (chronic bronchitis, ER visits, and acute respiratory symptoms).

At the bottom left corner of this window, type "Morbidity" into the Valuation Sum Identifier field. Leave the   Summation Type field as "Dependent".
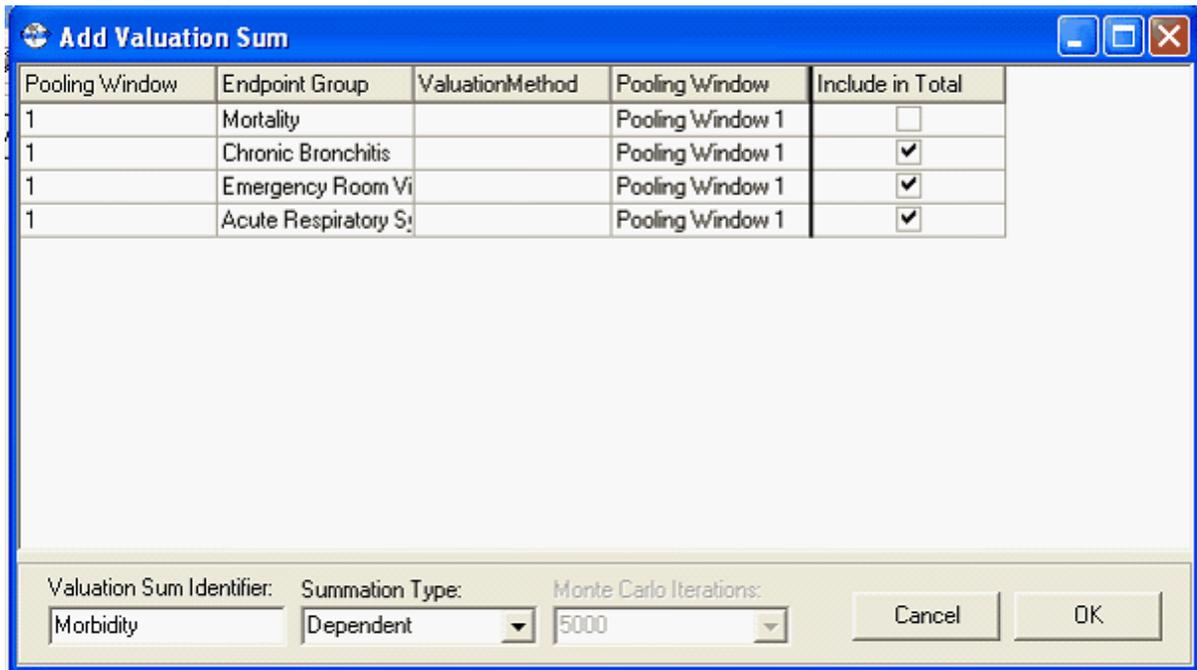
Figure 2-13. An *Add Valuation Sum* window showing the settings for creating the morbidity layer.

(2)Click **OK**, which will return us to the *Valuation Sums Layer* window. In the *GIS Field Name* column, highlight the cell and type in "Morbidity" (Figure 2-14).

> **Do not click OK** in the *Valuation Sums Layer* window. Continue on to step (3).

Pooling Window	Endpoint Group	ValuationMethod	Pooling Window	Include in Total
1	Mortality		Pooling Window 1	☐
1	Chronic Bronchitis		Pooling Window 1	✔
1	Emergency Room Vi		Pooling Window 1	✔
1	Acute Respiratory S		Pooling Window 1	✔

Valuation Sum Identifier: Morbidity   Summation Type: Dependent   Monte Carlo Iterations: 5000   Cancel   OK
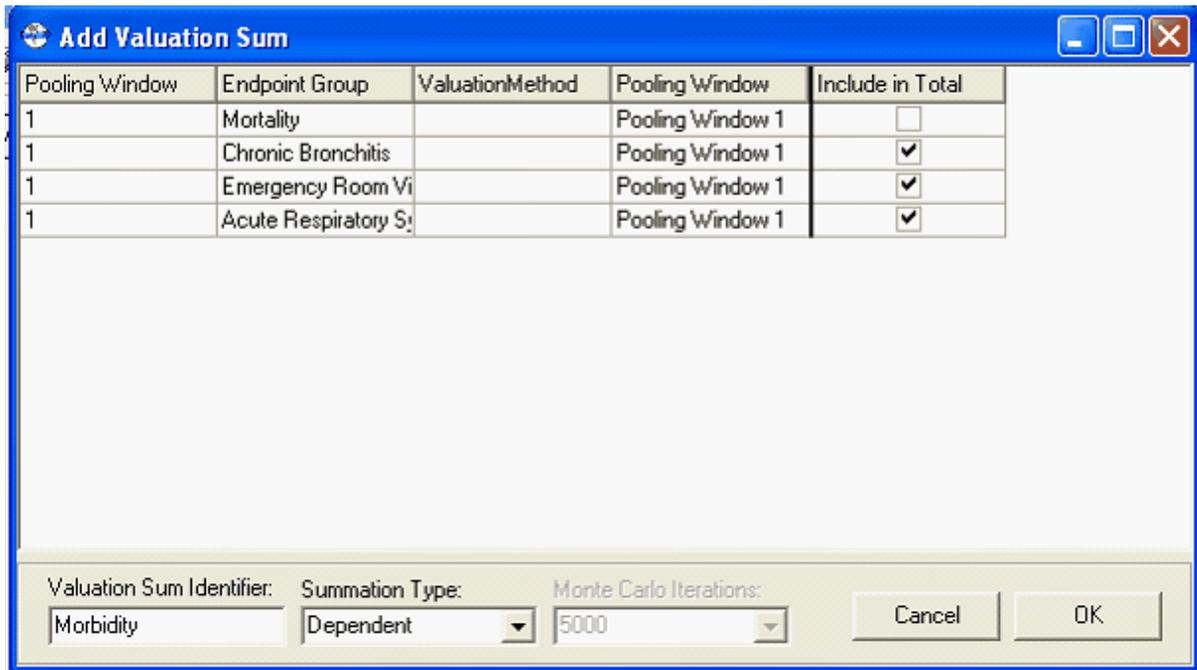
Figure 2-14. *Valuation Sums Layer* **window after the
morbidity sum identifier and GIS field name have been added.**

    (3)Next, add mortality to the *Valuation Sums Layer* window by clicking on the **Add Sum** button again, then checking "Mortality" in the *Add Valuation Sum* window, typing "Mortality" into the *Valuation Sum Identifier* field, and clicking **OK**.

        Return to the *Valuation Sums Layer* window and enter "Mortality" into the *GIS Field Name* cell in the mortality row that has been added to that window. Finally, click **OK** to close the window.

  (e)In the *Layers* area of the GIS window, you will see that two layers have been added: "Pooled Valuation Results Sums" and "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr".

    Double click on the "Pooled Valuation Results Sums" layer and set the display variable to "Mortality" (Figure 2-15).
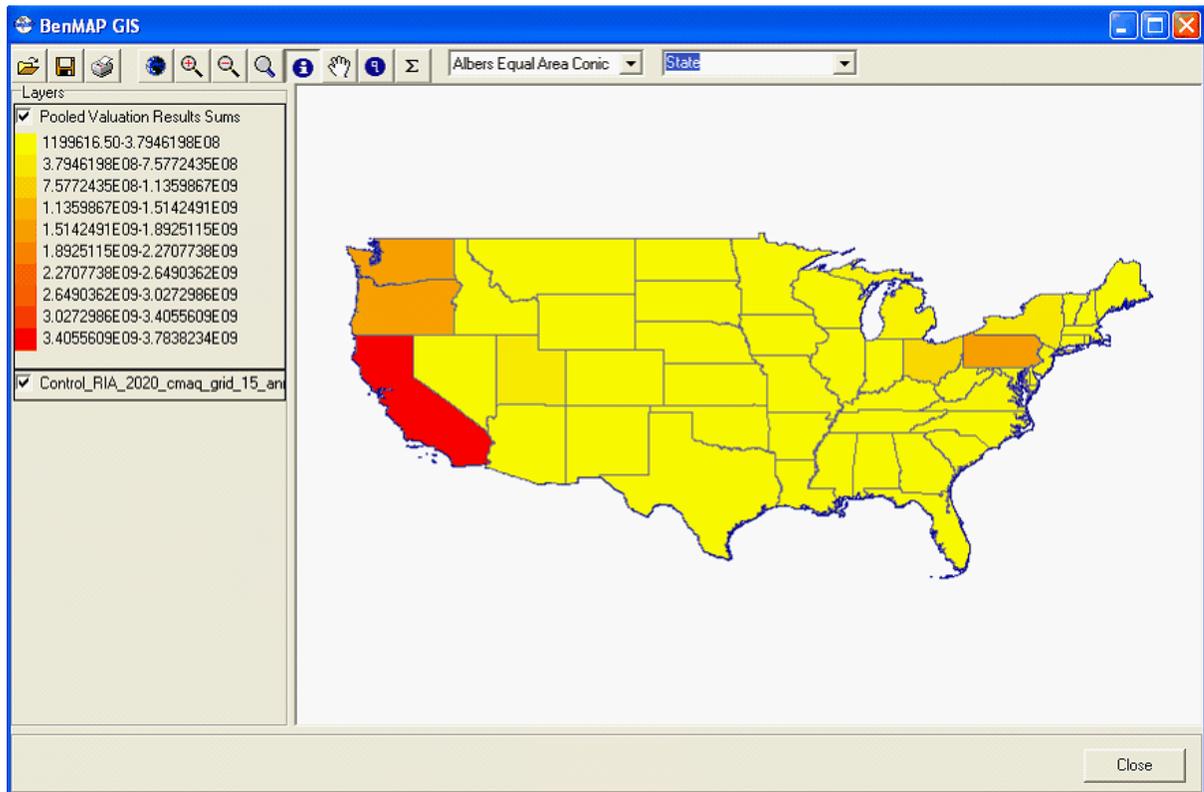
Figure 2-15. Economic benefit due to the reduction in deaths.

(f) Display the morbidity valuation data. Use the **Display info** button to explore some of the morbidity and mortality valuation values for specific states. *Note*: these benefits are in dollars.

(g) In the same GIS window, we will now overlay a pooled and aggregated health incidence layer for the same control scenario. We can use this layer to see the total number of prevented health incidences in each State.

(1) Use the **Open a File** button, select "APV Configuration Results (*.apvr)", then choose "Pooled Incidence Results" (Figure 2-16).
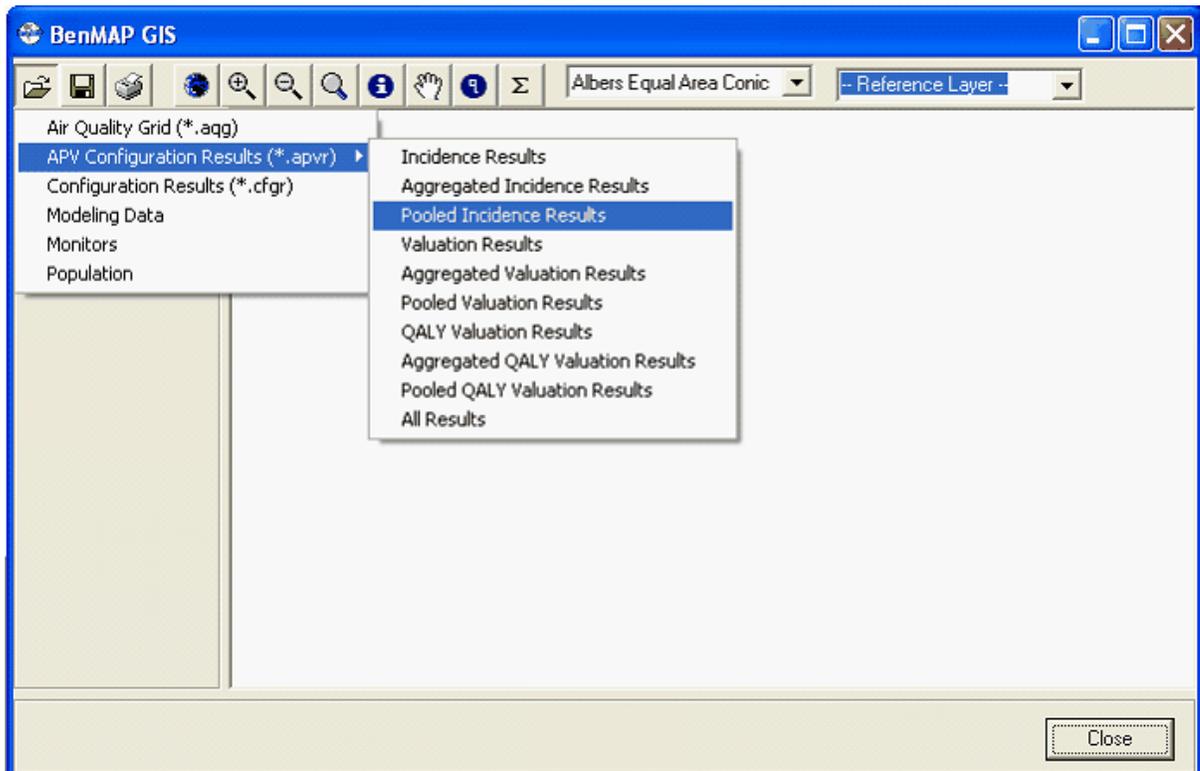
Figure 2-16. Selecting the "Pooled Incidence Results" file.

(2) Under the "Configuration Results" folder, open the same file: "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr".

(3) In the Edit GIS Field Names window, edit the field names as before, then close that window.

Double click the top layer, "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr (Pooled Incidence Results)" and set the variable to "Mortality".

(4) We should now have the pooled and aggregated mortality incidence (number of deaths avoided per state) overlaying the pooled and aggregated mortality valuation (economic benefits per state) (Figure 2-17).
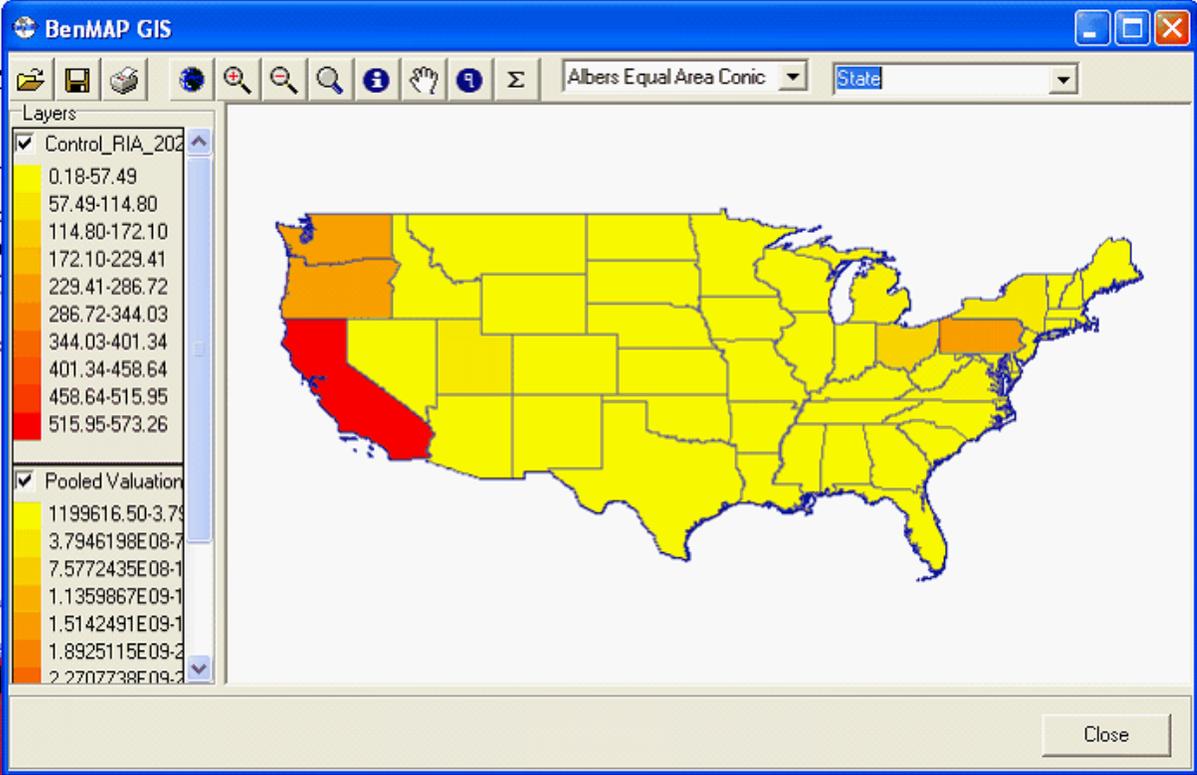
Figure 2-17. State mortality totals overlaying state economic benefit totals.

(h) Exercise (2.11): Compare the number of incidences for mortality and acute respiratory symptoms (one of the components of morbidity). What are the values for Washington State? Now compare the economic valuations for mortality and morbidity in the same state. What values did you get? What does contrasting the incidence numbers with the valuation numbers tell us about the valuation function for mortality versus the one for morbidity? Hint: you may need to switch the active layer by right clicking on the layer of interest in the Layers panel.

Answer:

(i) Close the GIS window. This completes the "Valuation Map" module.

A.1.2.5 Audit Trail

The audit trail is a tool for looking at the headers of files created through BenMAP. In other words, it allows you to explore the metadata (the settings, inputs, and/or configurations) for a BenMAP file. An audit trail is a useful feature to check your work and see which options you selected in your analysis. You can use the audit trail to look at all of BenMAP's output files, including air quality grids, configuration files, and results files.

A.1.2.5.1   Example Health Incidence: Control PM2.5

The goal of this exercise is to use the audit trail tool to explore the metadata of files created in BenMAP.

(a) In the main BenMAP window, click on the Report graphic that is at the bottom of the right-hand panel (under "Custom Analysis"). In the Select Report Type window that appears, select "Audit Trail Reports" (Figure 2-18). Click OK.
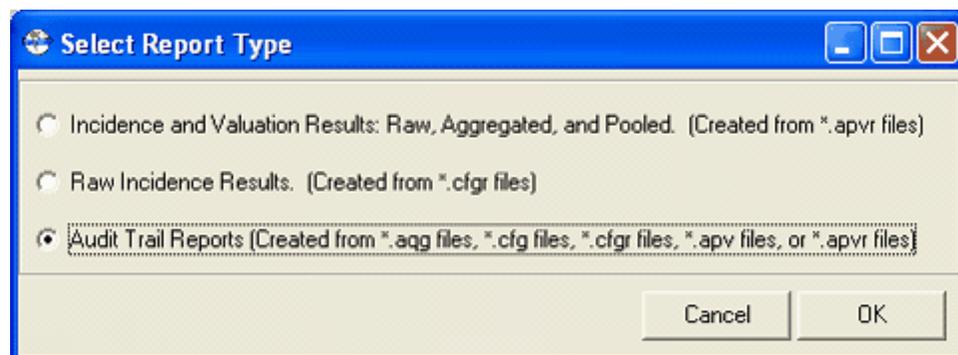
Figure 2-18. *Select Report Type* **window.**

(b) In the Open window, under the "Configuration Results" folder, open the file "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily.cfgr".

(c) An Audit Trail Report window will open that contains a tree structure giving the audit trail information (Figure 2-19). You can expand any section of the tree by clicking on the plus sign next to the heading, or collapse a section by clicking on the minus sign.
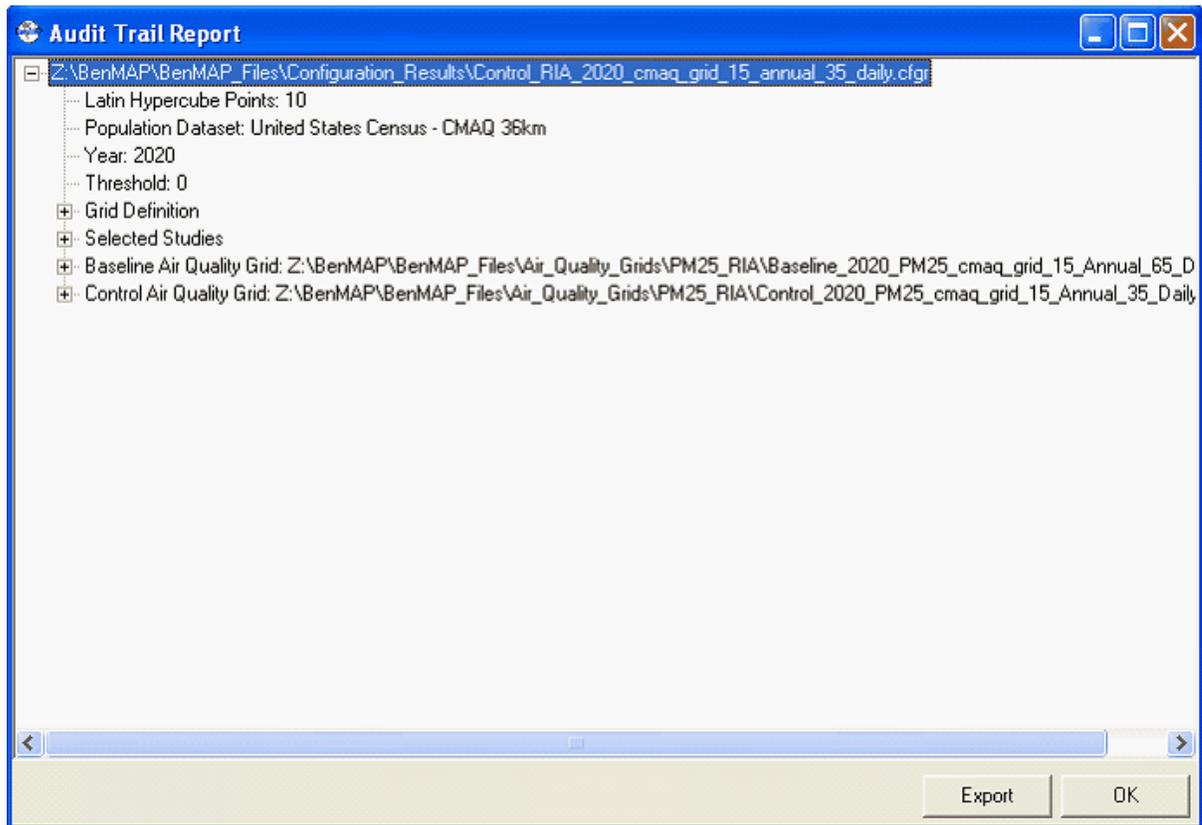
Figure 2-19. *Audit Trail Report* **window.**

(d) Exercise (2.12): What population year was used in this study? What is the name of the grid type?

> Answer:

(e) Exercise (2.13): What is the age range for the emergency room (ER) CR function, and who was the author of the study?

```
Answer:



```

(f) Export the audit trail. Click the Export button. In the Save as window, save the file under the "Reports" folder as "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily". This is a ".txt" file that can be easily opened in Microsoft Word or Notepad.

```
Answer:



```

(g) Close the Audit Trail Report window by clicking OK. This completes the "Audit Trail" module.

A.1.2.6 Report

Reports are a good way to summarize your BenMAP results in a table (columns and rows) and export them to be used in Excel or some other data analysis tool.

A.1.2.6.1   Example Pooled Incidence: Control PM2.5 RIA

The goal of this exercise is to create reports from BenMAP results. In particular, to look at the pooled and aggregated health incidence results from the 2020 RIA control scenario.

(a) In the main BenMAP window, click on the Report graphic under "Custom Analysis".

In the Select Report Type window, select the "Incidence and Valuation Results" item.

In the Open window, under the "Configuration Results" folder, open the file "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr".

(b) In the Choose a Result Type window that appears (Figure 2-20), select "Pooled Incidence Results". Click **OK**.
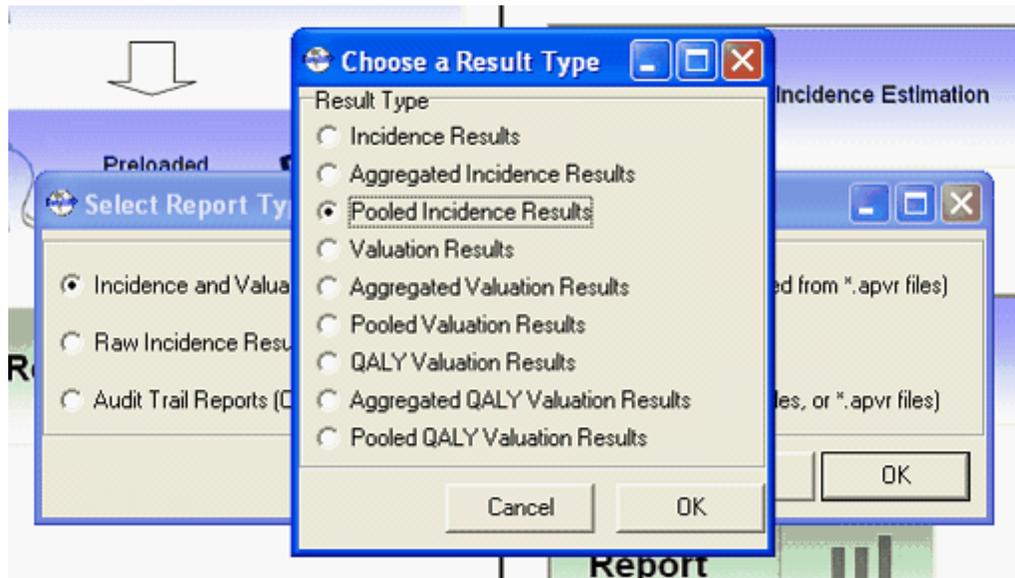
Figure 2-20. *Choose a Result Type* **window.**

(c) In the Configuration Results Report window that appears (Figure 2-21), click on the checkbox for "Endpoint Group" in the "Pooled C-R Function Fields" area, and **uncheck** the checkboxes for "Variance" and "Latin Hypercube" in the "Result Fields" area. In other words, we are selecting which of the many available columns to display in the results table.

Under the "Display Options", the "Elements in Preview" field determines the number of rows included in this preview window (in our example, 25). When you save the report, you will get all the rows.

> *Background*: In the report window, there are two columns in the report that are grid fields ("Column" and "Row"). They are unique identifiers of each minimum spatial unit. For a CMAQ grid, these are merely the column and row number for each cell in the grid. For political grids, their meaning depends on your grid definition. In our example, the column is the state code and the row is the FIPS code.
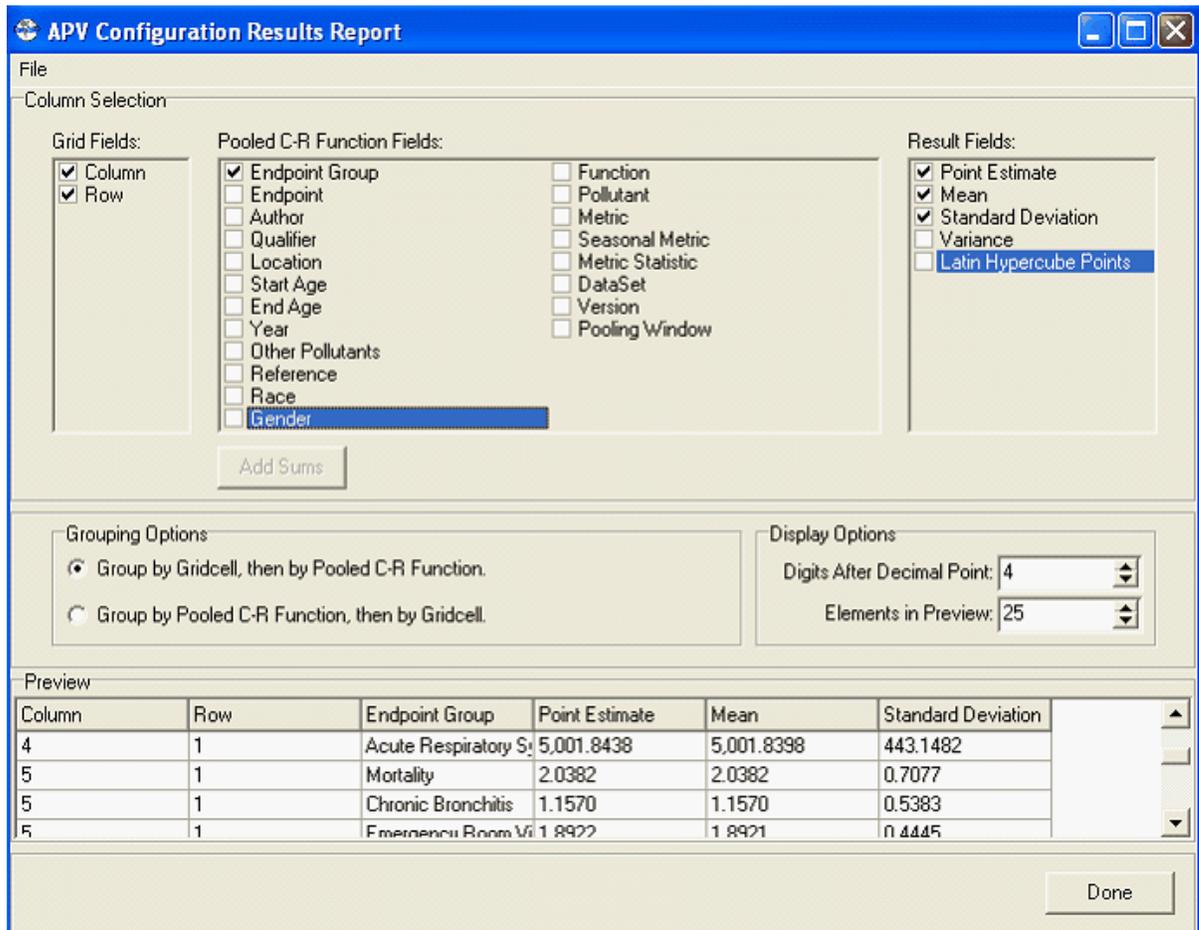
Figure 2-21. *Configuration Results Report* **window.**

(d) Save the report to a file. To do this, you can either drop down the File menu at the upper left-hand corner of the window and choose "Save" or simply hit Ctrl-S on the keyboard.

In the *Save As* window, save the file under the "Reports" folder as "Control_PM25_RIA_2020_15_annual_35_daily_incidence_state.csv", a comma-separated-value file.

(e) Open the .csv file in Excel to review outputs. The full path to this folder is:

"C:\Program Files\Abt Associates Inc\BenMAP 2.4 US Version\Reports"

(f) Close the *Configuration Results Report* window by clicking **Done**. This completes the "Report" module.

This completes the "Mapping Introduction" lab. In the next lab, "One-Step Analysis" we will run the one-step analysis, creating both health incidence results (cfgr) and valuation results (apvr).

A.1.2.7 Additional Mapping Activities

1. Overlay the PM2.5 monitor data over the baseline model data. Look at the similarity and differences in the pattern of monitor versus modeled data. Make sure that the monitors' layer is topmost.

2. Overlay the baseline model data over the control 2020 RIA 14/35 data. Compare the regions in the two layers that are out of attainment. Also, compare the regions that show PM2.5 concentrations greater than 10 µg/m3.

A.1.3 Section 3. One-Step Analysis

In this section, you will create health incidence and valuation results. This one-step analysis uses pre-defined health incidence and economic valuation configurations, which you can combine with your own baseline and control air quality grids. We will run two separate one-step analyses for two similar, but distinct control strategies.

The One-Step Analysis encapsulates all three stages of BenMAP: (1) Calculate the change in pollution concentrations that would be produced by the application of a given set of emissions controls. The concentration change in a pollutant (say, PM2.5) is the difference (the "delta") between the modeling results from an emissions control scenario and the modeling results from the baseline scenario. (2) These deltas, background incidence rates, and a gridded population dataset are then used in concentration-response (C-R) functions to calculate the change in health incidence that would result from this change in pollution. Typically, these health incidence results show the decrease in health incidence (e.g. the decrease in asthma, bronchitis, mortality, etc) due to a decrease in pollution. (3) These health incidence results are then used to calculate economic benefits due to changes in the population's mortality and morbidity. For example, what is the economic valuation of the avoided premature mortality due to the emissions control? The valuation results (and the underlying health incidence) are typically then aggregated to county, state, or national totals. Note: in One-Step Analysis, the health incidence and valuation functions are standard EPA configurations.

A.1.3.1 Example PM2.5 Control 2020 15/35 National

The goal of this exercise is to re-create the control RIA 2020 15 annual, 35 daily µg/m³ health incidence and valuation study that we saw in the "Mapping Introduction" lab. We will produce similar health incidence results (.cfgr file) and valuation results (.apvr file).

Procedures:

(a) In the main BenMAP window, open one-step analysis. Simply click on the left-hand panel graphic under the label "One-Step Analysis". This will open the

One Step Analysis window.

Figure 3-1. One-Step Analysis window.

(b) Set the "Run Name". This is the name that will be used for our cfgr and apvr files. A recommended practice is to be specific and base it on the control. In the "Run Name" field, type: "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county".

Note: we have added "county" to differentiate these results from our previous Control RIA 2020 15/35 results that   had been aggregated to State totals.

(c) Set the output directory. Use the Select button to select the "Configuration Results" directory

(d) In the panel "1. Air Quality Grids", set the baseline and control aqg. In many cases, the control AQG will have lower pollution values than our baseline. This reduction in modeled ambient PM2.5 is the result of our policy scenario that reduced emissions from industrial or other sources. Click the Open button next to the "Baseline File" field.

An Open window will appear. Navigate to the "PM25_RIA" folder under the "Air Quality Grids" folder, select "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg", and click Open.

Repeat these steps for the "Control File", this time selecting

"Control_PM25_RIA_2020_cmaq_grid_15_Annual_35_Daily.aqg".

(e) Map the change in pollution (the deltas) between the baseline and the control. Click on the Map Deltas button. A BenMAP GIS window will appear with 3 layers: "Delta" (the change in PM2.5), "Control Grid" (the control modeled values), and "Baseline Grid" (the baseline modeled values) (Figure 3-2):
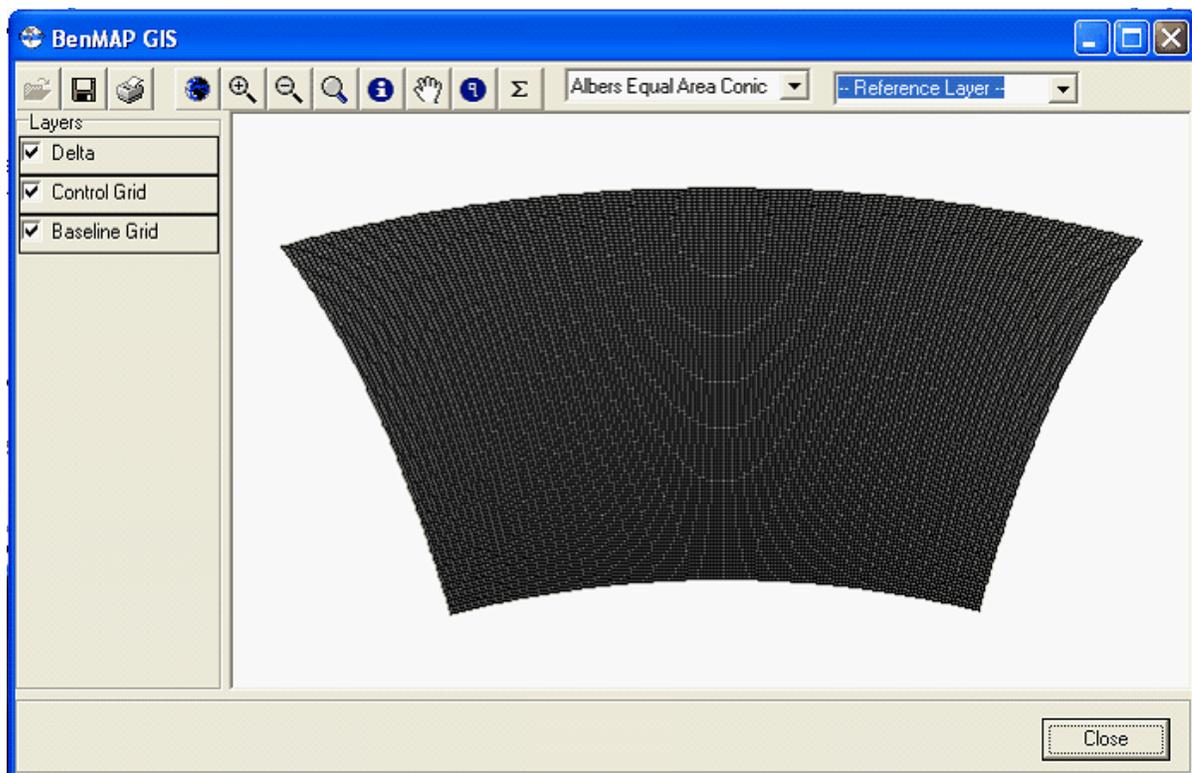
Figure 3-2. GIS window for mapping deltas between the baseline and the control.

(1) Double click the "Delta" layer in the BenMAP GIS window. In the Display

Option window, set the   variable to "QuarterlyMean" and turn off the grid outlines. Click OK.

In the BenMAP GIS window, uncheck the "Control Grid" and "Baseline Grid" layers and overlay a state reference layer.

> *Background*: The "Delta" is a map of the change in AQ between your baseline and control scenarios (i.e. baseline - control). Typically it is a good practice to check the delta. You can use this map to see if the changes in air quality are in the right direction (typically that you are getting positive values, i.e. reductions) and magnitude and that the changes are occurring in the appropriate places. You can also use this window to look at the underlying control or baseline scenarios: "Control Grid" and "Baseline Grid" in the *Layers* panel.
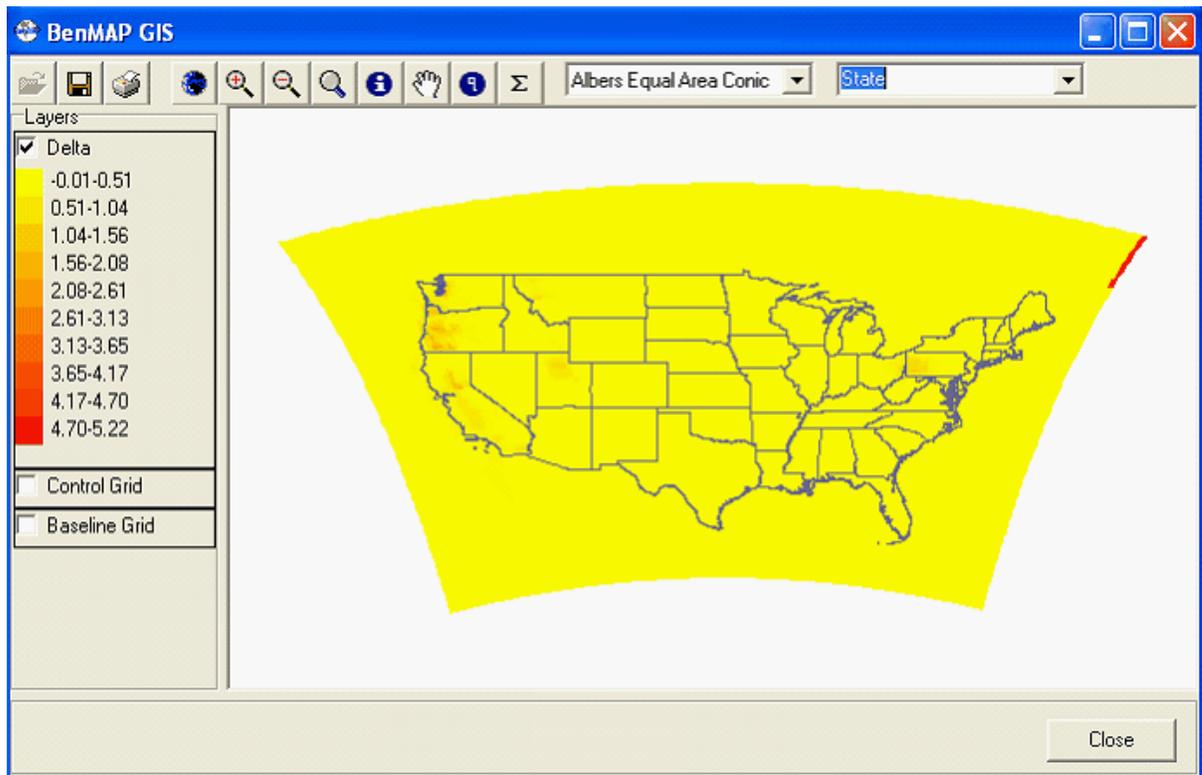
Figure 3-3. The resulting deltas between the baseline and the control.

    (1) **Exercise (3.1)**: In which states are the difference between the baseline and the control greater than 0.2 μg/m3? Hint: use the Build queries button.

```
Answer:



```

    (2) **Exercise (3.2):** Off the coast of California and on the Northeastern edge of the CMAQ grid, there are significant reductions in concentrations. Will these areas change our health incidence and valuation values? Explain your answer.

```
Answer:



```

(f)   Click **Close** in the *BenMAP GIS* window. This will return you to the *One Step Analysis* window.

(g)   *Aggregation.* Aggregation refers to the summing of health incidence results and valuation results to get more meaningful totals. In our case, the baseline and control modeled AQ data are at the fine scale of CMAQ grid cells. We want to aggregate them up to the county level.

    (1) **Exercise (3.3):** Look at the aggregation levels for Incidence and Valuation. What grids are available? If we set the Incidence grid to state, what grids are now available for Valuation?

Answer:

> (2) Set the aggregation levels to "County" under "2. Incidence" and "3. Valuation".

(h) Run one-step analysis. Click the **Go** button on the bottom of the *One Step Analysis* window. A *Progress* window will appear. The run will have completed when the *One Step Results* window appears (Figure 3-4). **Do not close** the *One Step Results* window when it appears.

Running a one-step analysis will take a few minutes. If we were running a larger domain, the grid cells had finer resolution, or our configuration included more functions, the one-step analysis would take much longer. The majority of the computation time is taken up in calculating the valuation results (apvr).

Analysis: The rest of the exercises in Section 3.2 focus on analyzing the results of our BenMAP One-step Analysis run. Specifically, we will look at the newly created health incidence results (cfgr) and valuation results (apvr) files.

(a) One Step Results are a series of customized reports and plots that were designed for EPA's apvr setup. Because we are using a simplified configuration, these reports provide limited results.

Click the **Audit Trail Report** button in the *One Step Results* window.
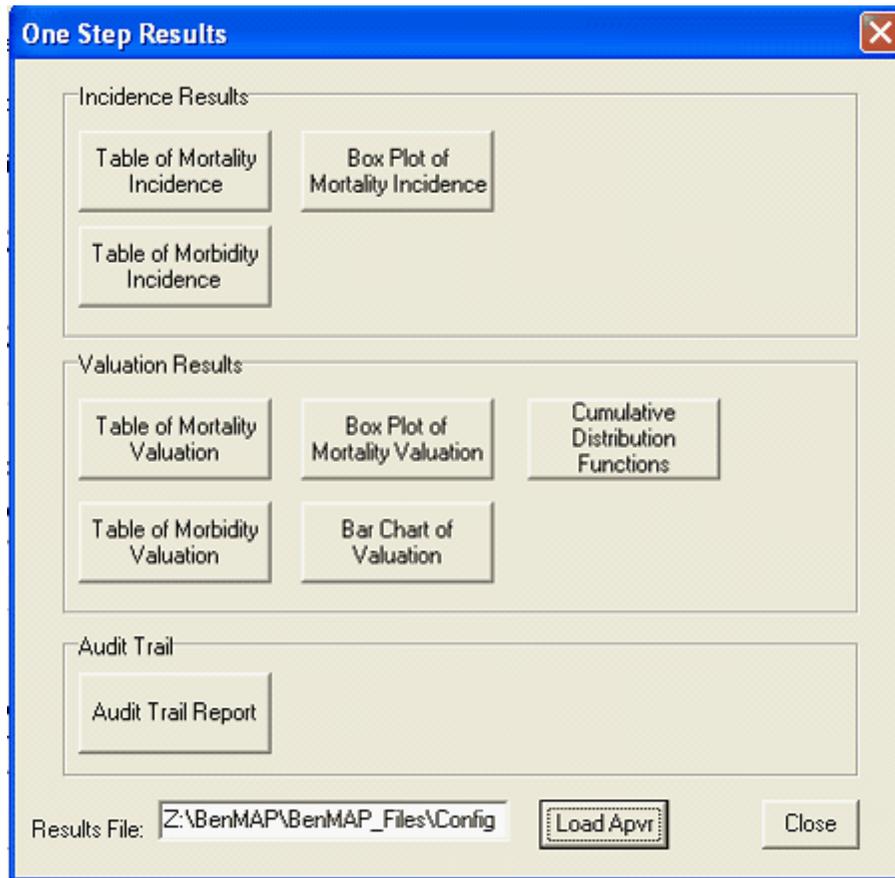
Figure 34. One Step Results window providing custom reports and plots.

Tip: By using the **Load Apvr** button (bottom of the *One Step Analysis* window), you can select any results file (apvr). *Note*, the full capabilities of the *One Step Results* currently only work for a full EPA RIA configuration.

(b) **Exercise (3.4):** In the audit trail report and under "Configuration Results", open "CR Function 0"? What is the function's endpoint? Who is the author of the underlying study? Under "Advanced", what is the aggregation name for incidence and valuation results?

Answer:

(c) In the *Audit Trail Report* window, click **OK** to close the window. Click **Close** in the *One Step Results* window. This will return you to the main BenMAP window.

(d) Open a new *BenMAP GIS* window, click on the **Open a File** button, then select "Configuration Results (*.cfgr)". Under the "Configuration Results" folder, select the newly created file: "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county.cfgr".

(e) In the *Edit GIS Field Names* window, change the field names to: "Mortality", "ChronBronc", "ER", and "AcutResp". Click **OK**.

In the *BenMAP GIS* window, double click the layer "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county.cfgr" and display the "AcutResp" variable.

(f) **Exercise (3.5):** Which grid is used in the health incidence results? Why is this different than the aggregation level in the One-step analysis window? *Note*: you can use either the audit trail or the GIS tool for this exercise.

Answer:

(g) **Exercise (3.6):** Which states have a reduction of more than 4000 acute respiratory incidences? *Hint*: use the **Build queries** button.

Answer:

(h) Overlay the new .apvr (valuation results). Use the **Open a File** button, select "APV Configuration Results (*.apvr)", then choose "Pooled Incidence Results" (Figure 35).

Open the file "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county.apvr". Edit the GIS field names to more meaningful names.
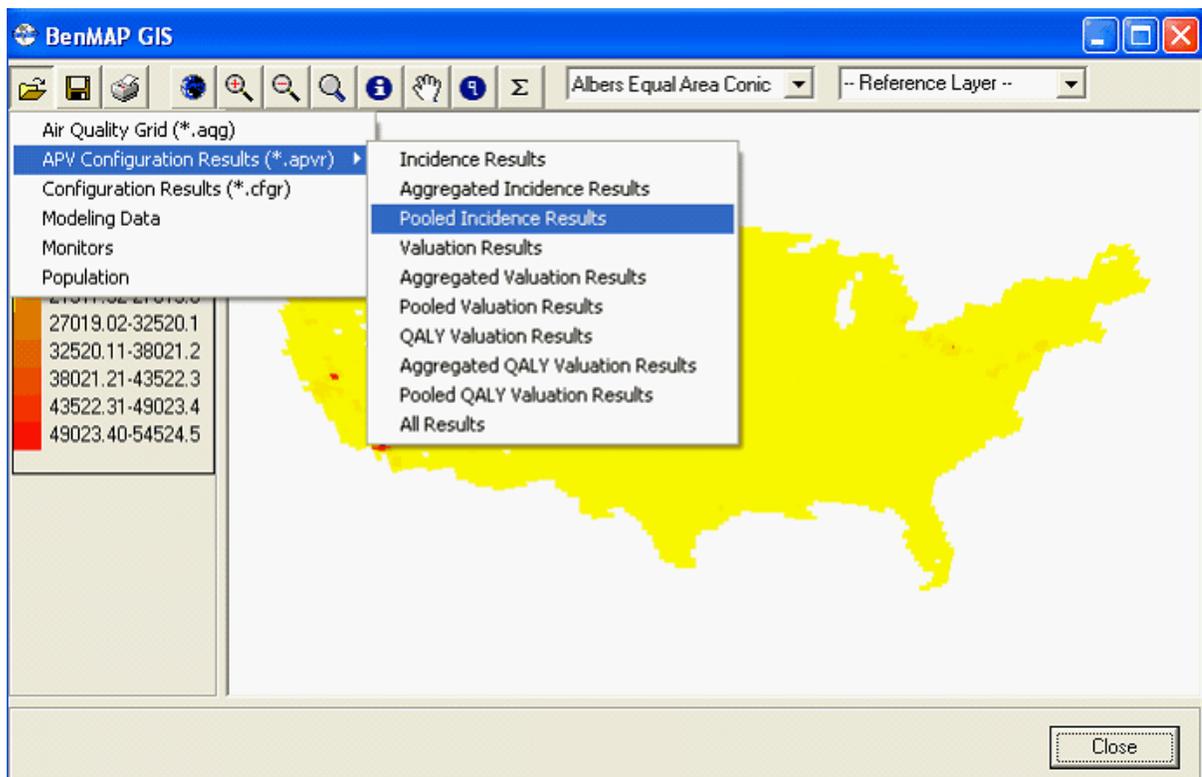
Figure 3-5. Opening pooled incidence results.

(i) **Exercise (3.7):** What grid is used in the pooled incidence results?

Answer:

(j) **Exercise (3.8):** Select mortality for your pooled (and aggregated) incidence results. What is the state (col) and FIPS (row) codes for the county with the highest number of avoided mortalities? What is the value?

Answer:

(k) **Exercise (3.9):** What are the maximum value and sum for avoided mortalities in the pooled incidence results? Make the incidence results ("Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county.cfgr ") the active layer. What are the maximum value and sum for avoided mortalities in the incidence results? *Hint:* Use the **Calculate layer statistics** button.

Answer:

(l) Overlay the pooled valuation results. As a reminder:

· **Open a File**, select "APV Configuration Results (*.apvr)", then select "Pooled Valuation Results". Open "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county.apvr".

· Edit the GIS field names. Use the same four field names ("Mortality", "ChronBronc", "ER", and "AcutResp").

· In the *Valuation Sums Layer* window, click **Add Sum** to create a Mortality and a Morbidity Sum (Figure 3-6). Click **OK** when you have added the two sums. We will view these summed results in the GIS window.
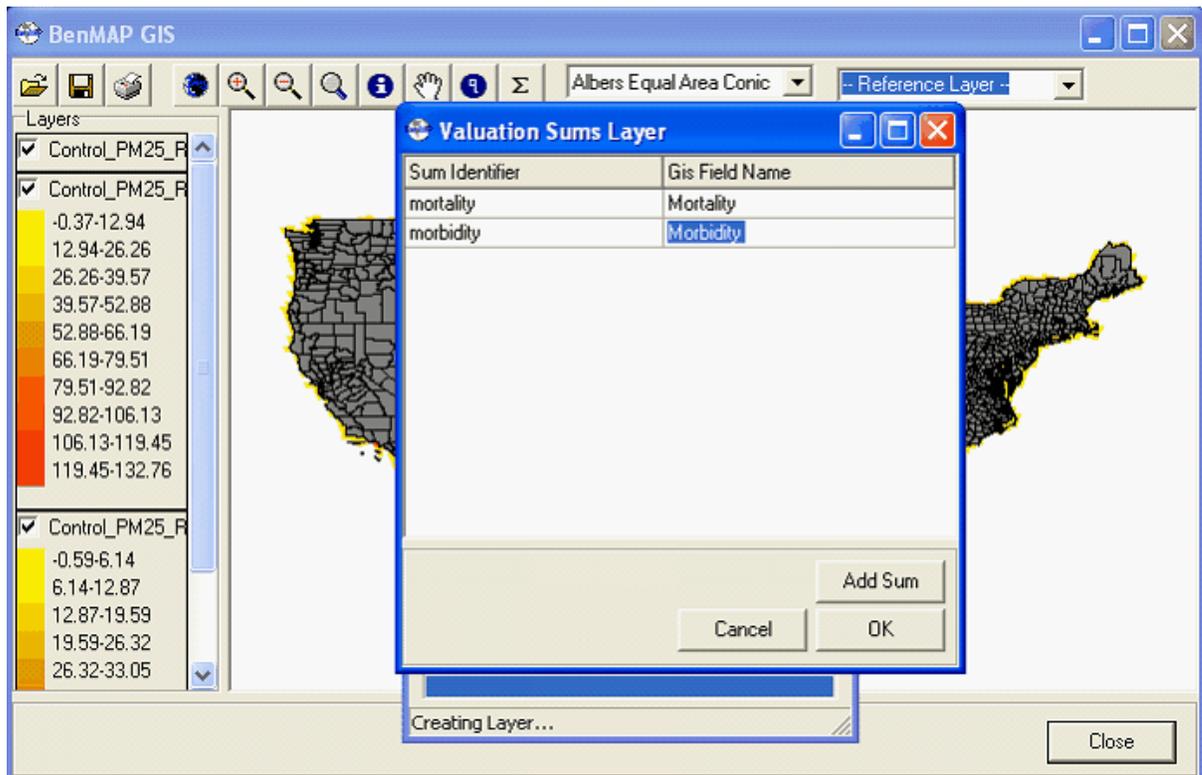
Figure 3-6. Creating valuation sums for the pooled valuation results.

· In the BenMAP GIS window, double click the "Pooled Valuation Result Sum" layer and display the "Morbidity" variable.

(m) **Exercise (3.10):** What is the state (col) and FIPS (row) codes for the county with the greatest benefit (valuation of avoided morbidity and mortality)? What are the morbidity and mortality benefits for this county? What is the sum of the mortality and morbidity benefits across the whole country? *Note*: you could also do this through Reports.

Answer:

(n) Save the pooled valuation results to a shapefile. Click on the **Save Active Layer** icon [save icon]. The *Save the active layer to file* window will appear.

Under the "Reports" folder, save the shapefile as "PM25_RIA_2020_15_annual_35_daily_county_pooled_valuation". Click **Save**. This shapefile can be used as an input for another GIS software program. Click **Close** in the GIS window returning you to the main BenMAP window.

(o) From the main *BenMAP* window, run a standard pooled valuation report on our apvr. As a reminder:

Close the *One-Step Analysis* window. From the main *BenMAP* window, click the **Report** graphic (bottom of the right-hand panel). Select a report type of "Incidence and Valuation Results".

After we have opened the newly created county apvr, select the "Pooled Valuation Results" in the *Choose a Results Type* window.

In the report window, check the "Endpoint Group" in the "Pooled Valuation Method Fields", uncheck the "Latin Hypercube" from the "Result Fields", and reduce the digits after decimal point to 0 (Figure 3-7).
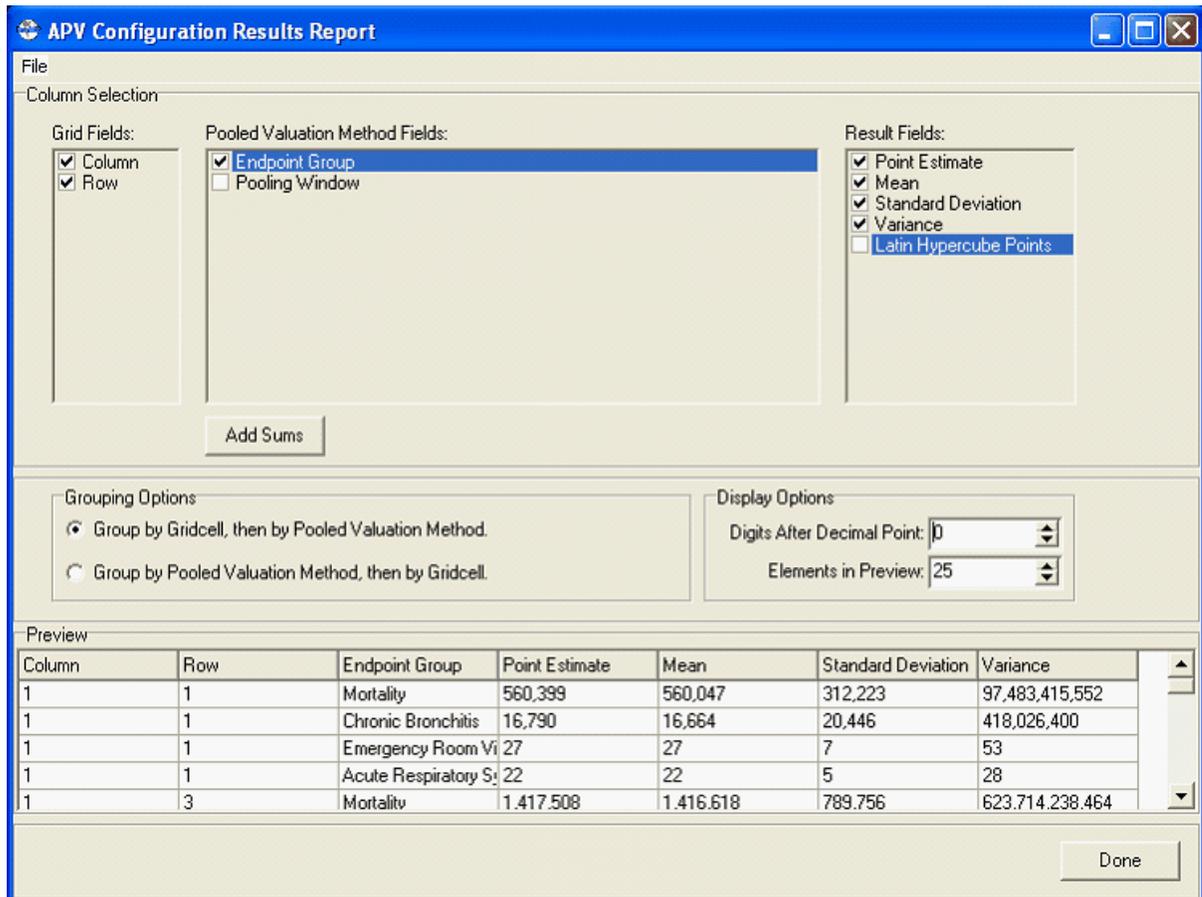
Figure 3-7. Report of pooled valuation results.

(p) **Exercise (3.11):** For col 1 and row 9, which endpoint (i.e. ER visits, Mortality, etc) has the greatest standard deviation? Which endpoint has the greatest coefficient of variation, a.k.a. relative standard deviation (standard deviation/mean)?

Answer:

(q)   Close the *APV Configuration Results Report* window by clicking **Done**.

A.1.3.2 Example PM2.5 Control 2020 14/35 State

The goal of this exercise is to create a control RIA 2020 14 annual, 35 daily µg/m3 health incidence and valuation study. Unlike our previous example, we will aggregate to the state level. Our intuition is that this will have greater benefits than the control 2020 RIA 15/35 µg/m3 analysis because we have a lower annual NAAQS.

Procedures: We have abbreviated the instructions because they are very similar to the steps you just completed in Section 3.2.
(a) Open the One Step Analysis window.

(a) Set the run name to:
    "Control_PM25_RIA_2020_cmaq_grid_14_annual_35_daily_state".

(b) Set the output directory to "Configuration Results"

(c) Set the baseline and control to:
    "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg" and
    "Control_PM25_RIA_2020_cmaq_grid_14_Annual_35_Daily.aqg", respectively.

(d) Go through the process of mapping the deltas. See section 3.2(e) for explicit instructions.

(e) **Exercise (3.12):** How do the deltas for this analysis compare to our previous analysis. For example, you could use the same query (QuarterlyMean > 0.2) to compare the deltas.

Answer:

(f) Set the health incidence and valuation aggregation to the "State" grid and click Go.

Note: the One-step Results will not work for State aggregation. The current One-step Reports are designed only for National or Report region aggregation levels. They also assume that we are using the full EPA configuration. Instead, we will use normal reports and mapping to analyze this run.

Analysis: The rest of the exercises in Section 3.3 focus on analyzing the results of our

BenMAP One-step Analysis run. Specifically, we will look at the newly created health incidence c results (cfgr) and valuation results (apvr) files.

(a) **Exercise (3.13):** Map the pooled incidence results for "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_state.apvr" and "Control_PM25_RIA_2020_cmaq_grid_14_annual_35_daily_state.apvr". *Note*: for detailed instructions see Section 3.2(d) in the analysis section. How do the mortality values for IL compare between the 2 control scenarios? How does the mortality values for CA compare? What are the overall differences between the 2 control strategies?

Answer:

(b) **Exercise (3.14):** Overlay the pooled valuation results for "Control_PM25_RIA_2020_cmaq_grid_14_annual_35_daily_state". Remember to create "Mortality" and "Morbidity" variables in the *Valuation Sums Layer* window. What are the mortality and morbidity values for CA? For IL? How does the sum of mortality and morbidity benefits across the 48 states compare to our results from "Control_PM25_RIA_2020_cmaq_grid_15_annual_35_daily_county" (see Exercise 3.10)? *Note*: for detailed instructions see Section 3.2(l) in the analysis section. You could also use reports to answer this question.

Answer:

(c) After finishing the exercise, close the *BenMAP GIS* window.

(d) This completes the "One-Step Analysis" lab. In the next lab, "Creating Grids" we will create new air quality grids (aqg results) from both monitor and model data.

A.1.4   Section 4. Creating Grids

In this section, you will create air quality grids (aqg) from both air quality (AQ) model data and from monitor data. The overall goal will be to produce baseline and control air quality grids for later estimations of health incidence and valuation.

A.1.4.1 Air Quality Model Grids

Think of AQ models as weather models for air pollution. They produce model results over the whole map on a regular grid. Examples of AQ models are CMAQ and CAMx. These models are post-processed to generate a csv file or a database table and their values are usually averaged to daily or yearly statistics. These post-processed datasets can now be imported into BenMAP.

A.1.4.1.1   Example: PM2.5 Control 2020 15/35 Adjusted

The goal of this exercise is to create a control AQ grid for PM2.5. The CMAQ model data input is similar to the 15 annual, 35 daily µg/m3 control scenario that we saw in the "Mapping Introduction" and "One-Step Analysis" sections, except that it has been post-processed (adjusted) to remove the extreme high values.

Procedures:

(a)Open Air Quality Grid Creation. Click on the graphic on the right-hand panel of BenMAP's main window

(b)Choose "Model Direct" in the Air Quality Grid Creation Method window (Figure 41).
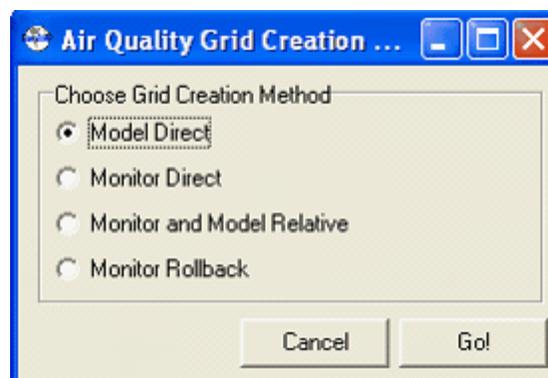
Figure 4-1. *Air Quality Grid Creation Method* **window.**

(c)In the Model Direct Settings window, set the grid type to "CMAQ 36km Nation Overlap" and the pollutant to "PM2.5". Leaving the tab as "Generic Model Databases", click **Browse**. Navigate to the "PM25_RIA" folder under the "Air Quality Grids" folder. In the *Open* window, change the "Files of type" field to "Text files" and select

the model data:
"Control_PM25_RIA_2020_cmaq_grid_15_Annual_35_Daily_adjusted".

(d)Click **OK** to create a new model derived aqg. Save the output under the "PM25_RIA" folder as: "Control_PM25_RIA_2020_cmaq_grid_15_Annual_35_Daily_adjusted". *Note*: this may take a few minutes to run.

A.1.4.1.2   Example: Control PM2.5 RIA 2020 14/35 adjusted

The goal of this exercise is to create a control AQ grid for PM2.5, similar to the 14 annual, 35 daily µg/m3 control that we saw in the Mapping Introduction and One-step Analysis sections. Again, this model data has been post-processed to remove the extreme high values.

Procedures:

(a)Repeat the above steps, now using the 14 annual 35 daily adjusted model data

Analysis:

(a)**Exercise (4.1)**: Open up either of the newly created adjusted model data using the GIS tool. Using the Query button, select regions that have a PM2.5 QuarterlyMean greater than 10 µg/m3. Now change the query to QuarterlyMean greater than 15 µg/m3(i.e. out of attainment). How do these results compare to the non-adjusted aqg's that we analyzed in the "Mapping Introduction" section (see Exercise 2.6)?

Answer:

(b)Close the GIS window. This completes the "AQ Model Grids" module.

A.1.4.2 Monitor Grids

When we convert monitor data into an AQG, we need to interpolate from the monitor locations to all the grid locations. There are two overarching interpolation techniques: Closest Monitor (also known as nearest neighbor) and Voronoi Neighborhood Averaging (VNA). Closest Monitor means that the grid cell will have the same value as the nearest monitor. VNA uses distance to weight the average of monitors in calculating the grid cell's value. There are multiple advanced options to change the distance weighting functions and to apply maximum distance thresholds to these calculations. Additional details are provided in the User's Manual.

A.1.4.2.1 Example: Baseline PM2.5 2004

The goal of this exercise is to generate an AQ grid from PM2.5 monitoring data from 2004. In addition, we will compare closest monitor and VNA interpolation techniques.

Procedures:

(a) Click the **Create Air Quality Grid** button.

(b) Select "Monitor Direct" and click the **Go!** button (Figure 4-2).
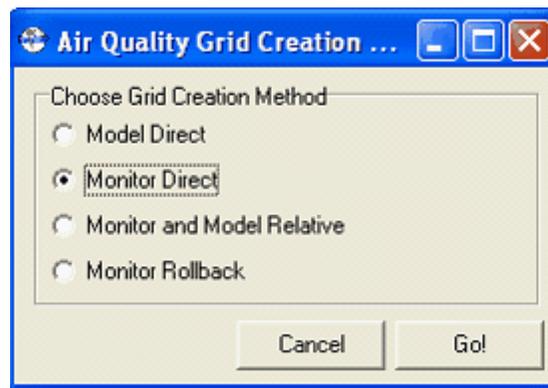
Figure 4-2. Selecting Monitor Direct.

(c) In the Monitor Direct Settings window (Figure 4-3), select "CMAQ 36km Nation Overlap" as the grid type.

Select "PM2.5" as the pollutant ,and "Closest Monitor" in the "Interpolation Method" list.

Make sure that the "Library" tab is selected, and select "EPA Standard Monitors" as the monitor dataset and set the monitor library year to "2004".
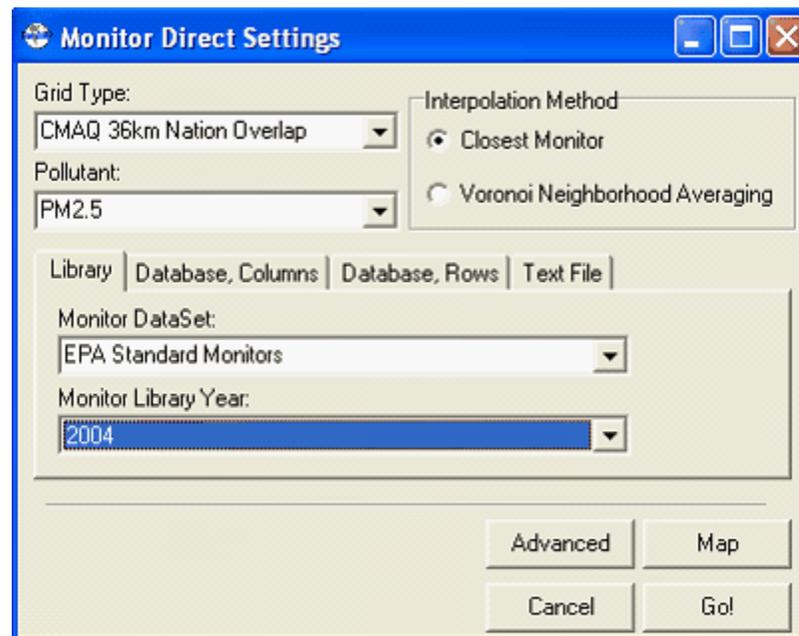
Figure 4-3. *Monitor Direct Settings* **window.**

(d)Click the **Map** button. This will bring up the *BenMAP GIS* window (Figure 4-4).
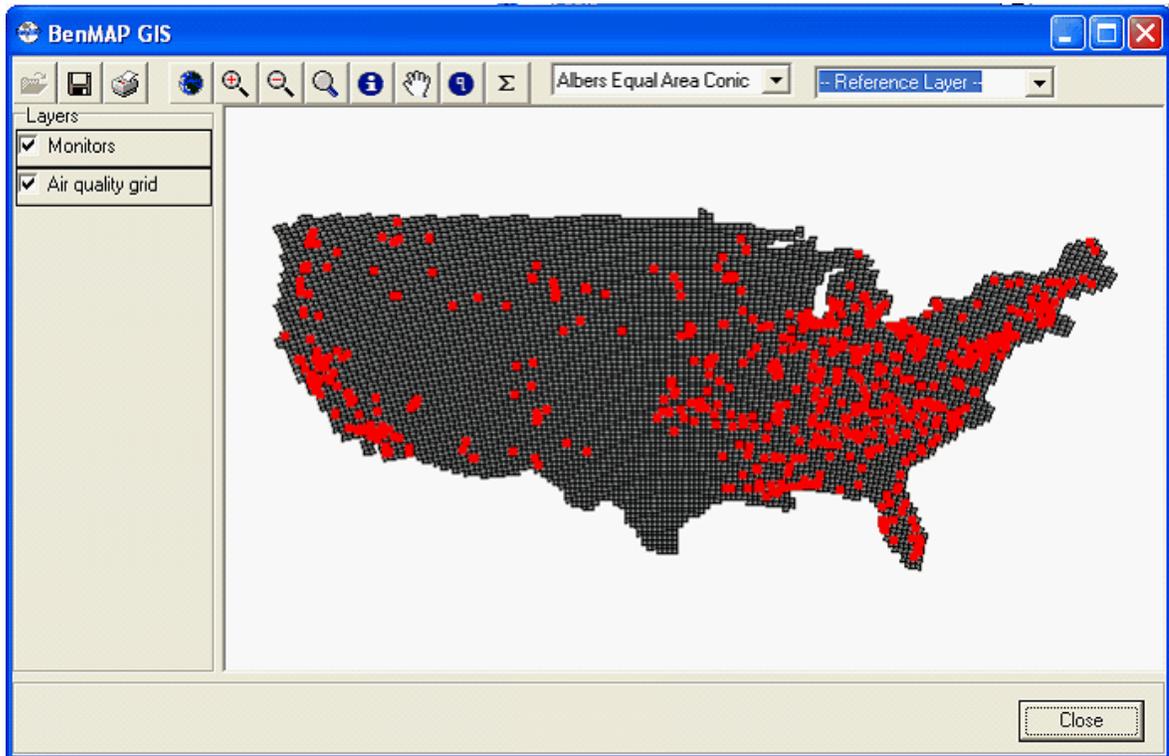
Figure 4-4. Map of monitors and closest monitor.

(e) Double click the "Air quality grid" in the Layers list on the left side of the window. Select "QuarterlyMean" in the Variable list, and uncheck the "Grid Outline" box.

(f) In the GIS window, overlay the state reference layer

(g) Zoom into the California-Nevada border area and look at the pattern of the AQG.

(h) Click Close on the GIS window. This will return you to the Monitor Direct Settings window.

(i) **Exercise (4.2)**: Change the interpolation technique to VNA. Remap the data and look at the California-Nevada border. How does the VNA AQG compare to the Closest monitor AQG? Note: this will likely take significantly longer to calculate.

> Answer:
>
>
>
>
>
>
>
>

(j) Return to the *Monitor Direct Settings* window (by closing the GIS window), and click **Go!**

> *Tip*: If you click Cancel at this point, you will not have created an aqg. Only by clicking **Go!** do you actually create a new aqg file. The **Map** button gives you a preview of the interpolated aqg, it does not actually produce the file.

(k) In the *Save as* window, create a new folder, "PM25_Monitor" under the "Air Quality Grids" folder. Save the aqg as: "Baseline_PM25_2004_VNA_CMAQ_grid".

(l) This completes the "Monitor Grid" module.

A.1.4.3 Monitor Rollback

Monitor rollback is a method for creating control scenarios. The rollback technique has three main steps: (1) the pollutant and the monitoring data are chosen, (2) the chosen monitoring data is reduced through one or more of the three rollback approaches, and (3) the rolled back monitors are then interpolated to the chosen grid type. The three rollback approaches are: percentage, incremental, and rollback to a standard. Note: in constructing the rollback, we differentiate between the rolled back grid and the control grid. The rollback grid is the grid type under which we want to perform the rollback. This typically is a political grid (e.g. State or county). In contrast, the monitor data will be interpolated to our control grid type.

A.1.4.3.1   Example: Control PM2.5 2004 Percentage Rollback

The goal of this exercise is to create a control by performing a 10% rollback of monitors in the West coast and in Pennsylvania (remember, these regions had the largest deltas in the One-step analysis section). With this rollback approach, each daily value above the background level is rolled back by 10 percent. We will produce a county grid aqg.

Procedures:

(a) Click the **Create Air Quality Grid** button. This will bring up the Air Quality Grid Creation Method window.

(b) Select "Monitor Rollback" and click the **Go!** button (Figure 4-5).
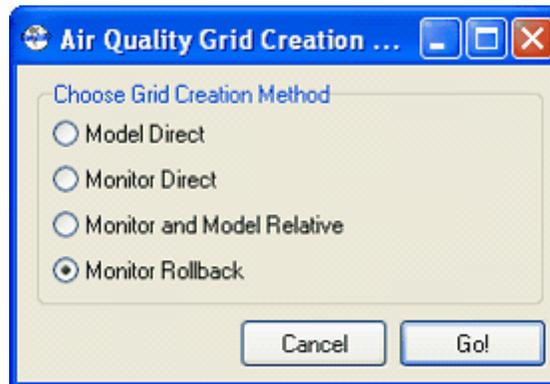
Figure 4-5. Selecting Monitor Rollback.

(c)This will bring up the Monitor Rollback Settings: (1) Select Monitors window (Figure 4-6). Select "PM2.5" from the pollutant list.

Make sure that the "Library" tab is selected, and select "EPA Standard Monitors" from the monitor dataset list. Set the monitor library year to "2004". Select "State" from the rollback grid type list.
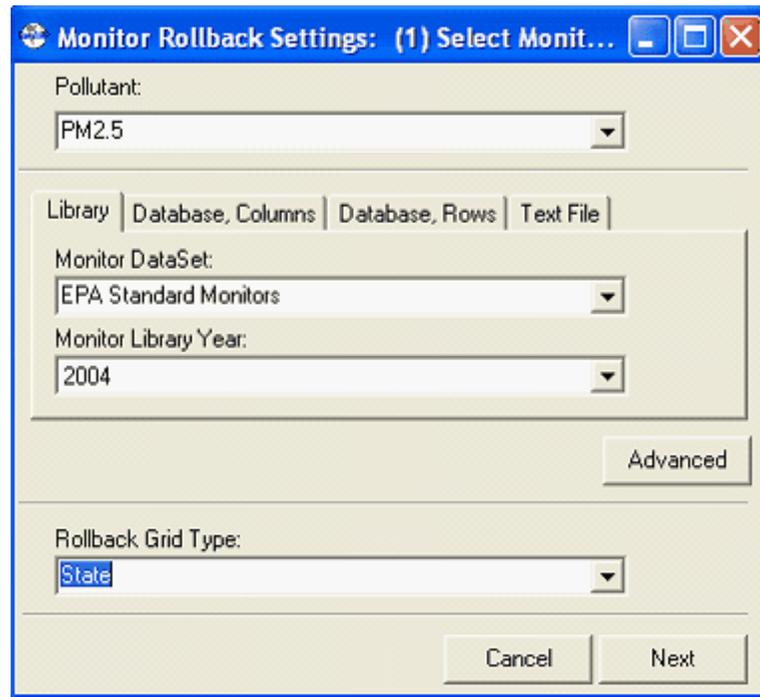
Figure 4-6. *Monitor Rollback Settings: (1) Select Monitors* **window.**

(d) Click **Next**, which will bring up the *Monitor Rollback Settings: (2) Select Rollback Regions & Settings* window.

(e) In that window, click Add Region. A Select Region Rollback Type window will appear (Figure 4-7). Throughout this region, we will apply one rollback technique. In this case, we will be using a percentage rollback.

Select "Percentage Rollback" and click OK. You will be returned to the Monitor Rollback Settings: (2) Select Rollback Regions & Settings window.
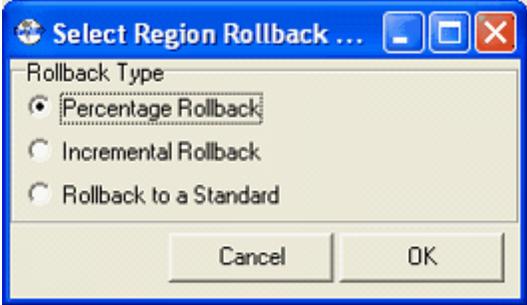
Figure 4-7. *Select Region Rollback Type* **window.**

(f) In region 1, set the rollback to 10% and the background to 3 µg/m³. In other words, all monitor values that are greater than 3 will have a 10% reduction in value.

(g) Apply the controls to the West coast states and PA (four states in total). Use your mouse to click on the map to select a state. For example, we have selected California in Figure 4-8.
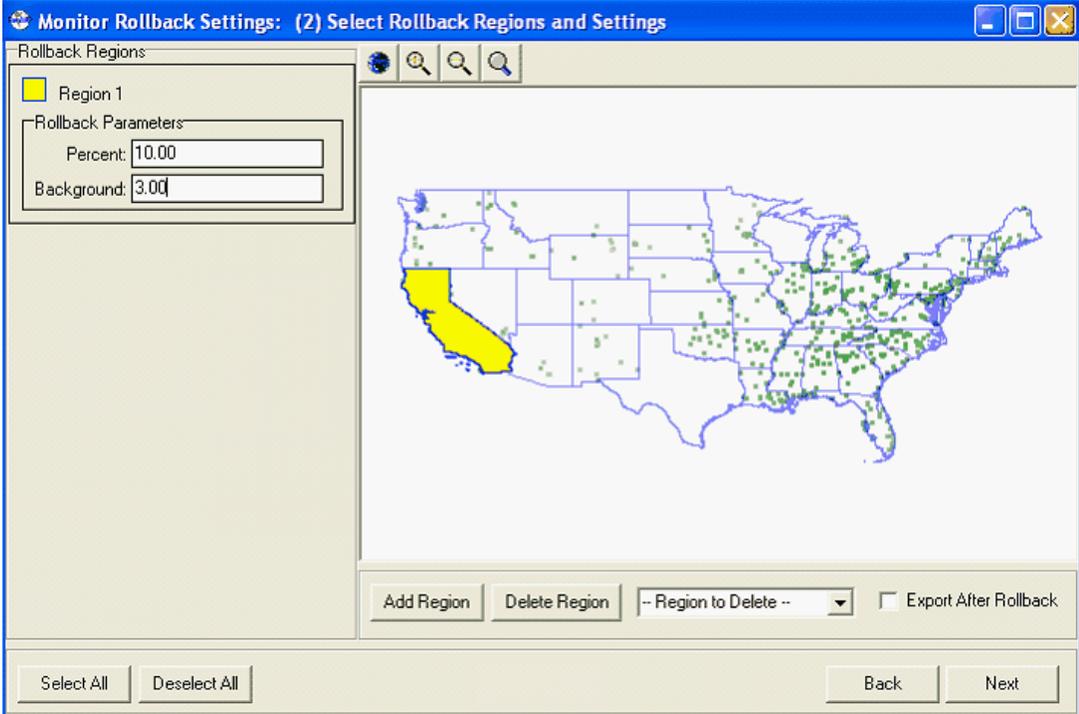
Figure 4-8. *Monitor Rollback Settings (2) Select Rollback Regions and Settings* **window. Selecting a region for a 10% monitor rollback.**

(h) Click **Next**

(i) In the *Monitor Rollback Settings (3) Additional Grid Settings* window (Figure 4-9), we want to use VNA interpolation and county grid type (i.e., the control grid). Uncheck "Make Baseline Grid (in addition to Control Grid)" because we did this in a previous step.
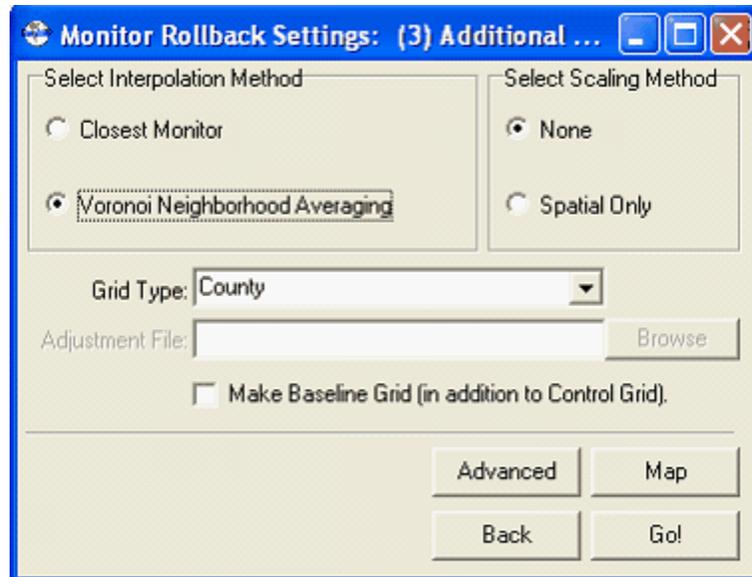
Figure 4-9. *Monitor Rollback Settings (3) Additional Grid Settings* **window. Setting county grid type.**

(j) Click *Go!* and save the aqg under "PM25_Monitor": "Control_PM25_2004_VNA_county_10pct_rollback_3_background"

Analysis:

(a) Exercise (4.3): Create an audit trail report on our new monitor rollback aqg. Under "Advanced", what is the neighbor scaling type? Under "Monitor Rollback", what are the four states (names and codes) that have been rolled back? What is the rollback method and value?

> Answer:
>
>

(b)Exercise (4.4): Map our new monitor rollback aqg. What states have a QuarterlyMean greater than 15 µg/m3?

> Answer:
>
>

A.1.4.3.2   Example: Control PM2.5 2004 Multiple Rollback Techniques

The goal of this exercise is to combine multiple rollback techniques into one control scenario. We will rollback the West coast incrementally. On the East coast, we will rollback to a standard using peak shaving. In other words, on the West coast, we will decrease all monitors by a fixed amount, while on the East coast we will define a standard and only reduce those monitors that exceed that standard, for only those hours over the standard. The aqg will have a CMAQ grid type.

Procedures:

(a)Go through the same steps as above to setup a monitor rollback for PM2.5 for 2004. Again select "State" as the rollback grid type.

(b)Add the first region, then select "Incremental Rollback" in the Select Region Rollback Type window (Figure 4-10) and click OK.
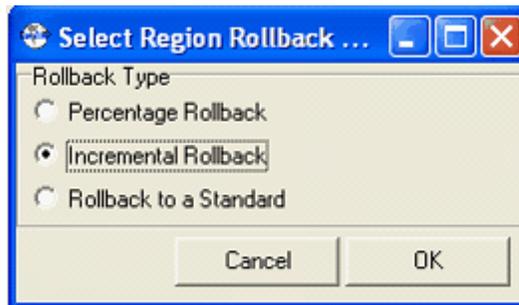
Figure 4-10. Selecting incremental rollback.

(c) Set the rollback to 4 µg/m³ and the background to 3. Select all the West coast states.

(d) Add a second region. In the *Select Region Rollback Type* window, set the type to "Rollback to a Standard". Here we have many more options in defining what our standard is and how we want to reduce the monitors so that they match that standard. In our case, we will set a standard that no monitor should have a daily mean value greater than 35 µg/m³.

   (1) Set daily metric to "D24HourMean". Leave the seasonal metric blank and the annual statistic type blank.

   (2) Set the standard to 35 and leave the ordinality as 1.

   > *Background*: Ordinality refers to the number of monitor values in the season that can exceed your standard. For example, if we had set the ordinality to 4, than a monitor can have as many as three daily averages >35 µg/m³ without violating the standard (i.e. it would not be rolled back). If it has more than 3 daily averages in an exceedance of the standard, then the rollback technique will be applied to that monitor.

   (3) Set the rollback method to "Peak Shaving", and the background to 3 µg/m³.

   (4) Select all the East coast states. Figure 4-11 shows the selection of the first state, Maine. *Note:* You may need to use the zoom button to get some of the smaller states.
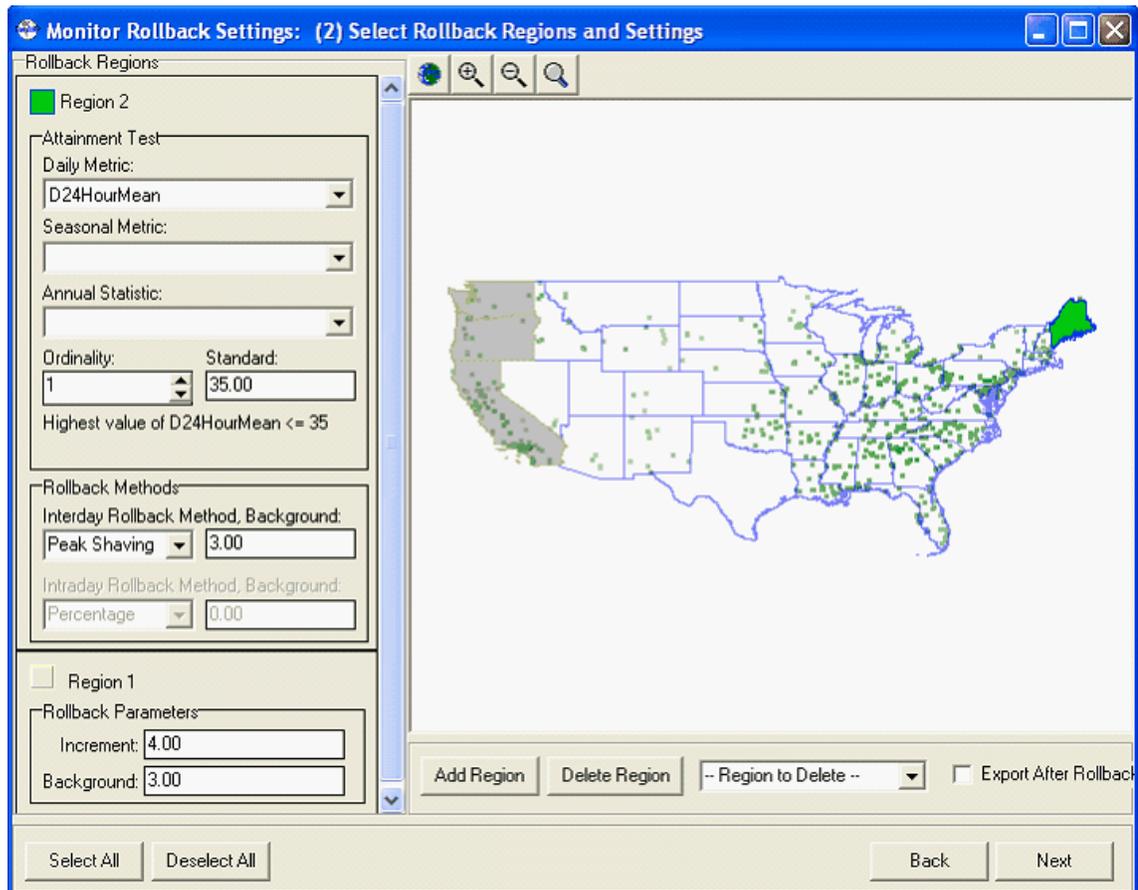
Figure 4-11. *Monitor Rollback Settings (2) Select Rollback Regions and Settings* **window. Adding a second rollback region, rollback to standard.**

(e) Click **Next** to go to the *Monitor Rollback Settings: (3) Additional Grid Settings* window. This time, set the grid type to "CMAQ 36km Nation Overlap" and the interpolation to "VNA".

Uncheck "Make Baseline Grid (in addition to Control Grid)" because we did this in a previous step (Figure 4-12).
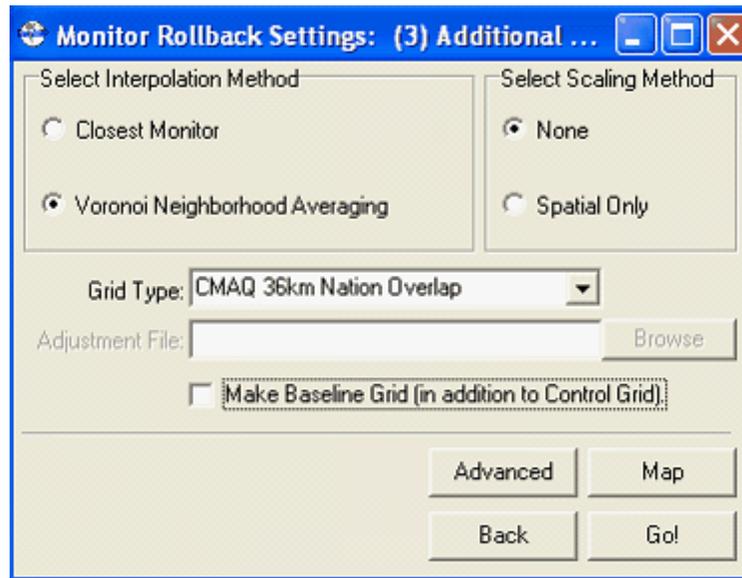
Figure 4-12. *Monitor Rollback Settings (3) Additional Grid Settings* **window. Setting CMAQ grid type.**

(f) Click **Go!.** Save the results under "PM25_Monitor" as: "Control_PM25_2004_VNA_CMAQ_4_incremental_35_daily_3_background"

Analysis:

(a)**Exercise (4.5):** Create an audit trail report on our new monitor rollback aqg. Under "Monitor Rollback", describe the two rollback regions, focusing on their techniques?

Answer:

(b)**Exercise (4.6):** Map the deltas between the Baseline $PM_{2.5}$ VNA 2004 and this multiple rollback control that we just created. How do the East Coast and West Coast deltas differ? How do you explain the pattern on the eastern boundary of CA, WA, and OR? *Hint*: Use the One-step analysis window to map the deltas. You don't need to perform a full One-step analysis. Instead, in the "1. Air Quality Grids" panel, select your newly created "Baseline_PM25_2004_VNA_CMAQ_grid.aqg" as the "Baseline File" and your newly created control as the "Control File". Click **Map Deltas.**

Answer:

(c)Close any open BenMAP windows. This completes the "Monitor Rollback" module.

(d)This completes the "Creating Grids" lab. In the next lab, "Health Incidence" we will take our aqg results and calculate the corresponding health incidence results due to the change in air quality.

A.1.5 Section 5. Health Incidence

In this section, you will modify an existing health incidence configuration and use it to create new health incidence results. You will create two separate sets of health incidence results based on the same configuration and two similar, but distinct, control strategies.

Creating health incidence results has three main stages:

(1)Select baseline and control air quality grids (AQG) and other general settings. These general settings include population dataset and analysis year, air quality threshold, and statistical parameters. The delta between the baseline and control AQGs is combined with the population data as a major input to the health impact functions.

(2)Select specific health impact functions and modify default parameters. Some of these parameters include demographics (e.g., gender, race, or age ranges) and incidence and prevalence rates.

(3)Save all the settings from the first two stages as a configuration (cfg), which can be re-used later with other baseline/control pairs if desired. Finally, run the health incidence configuration, which will create a new health incidence results file (cfgr).

Health impact functions relate the change in number of observed, adverse health effects in

a given population to a given change in concentration for a given pollutant. These functions are usually based upon epidemiological studies with specific sub-populations, and have a baseline estimate of the health effect. Having estimated changes in air quality levels from monitored or modeled data in previous labs, in this section, we will learn how to use these changes in concentrations and health impact functions to estimate changes in health incidence.

In this lab, we start from an existing configuration to help you develop a new configuration. After working through this lab, however, you should also be able to create a new configuration from scratch.

Note: For many of the health endpoints (e.g., mortality), there are many different health impact functions that you could choose to include in your configuration. This lab has you select specific functions and teaches you how to differentiate these functions and modify the input parameters. This lab does not teach you how to determine which functions are the best ones for a particular study. To determine the best choices for a particular analysis, we recommend that you read the appendices accompanying BenMAP that describe the specific epidemiological studies that correspond to the specific health impact functions and/or that you discuss your choices with an epidemiologist. If you want to use EPA's standard set of functions, you can use the configurations that are pre-loaded for the one-step analysis when you download BenMAP.

A.1.5.1 Example: PM2.5 Control 2020 14/35

The goal of this exercise is to create a new configuration by modifying the health incidence configuration that was used in the One-step analysis (Section 3) and create health incidence results for the control scenario RIA 2020 14 annual, 35 daily µg/m3. We will add the following health impact functions to the previous configuration: mortality for infants, hospital admissions due to respiratory problems, and acute myocardial infarctions (AMIs), also known as heart attacks.

Procedures:

(a)In the main BenMAP window, begin a health incidence estimation by clicking on the graphic titled "Incidence Estimation" located in the right-hand panel, under the "Step 2" heading. This will open the Configuration Creation Method window (Figure 5-1).
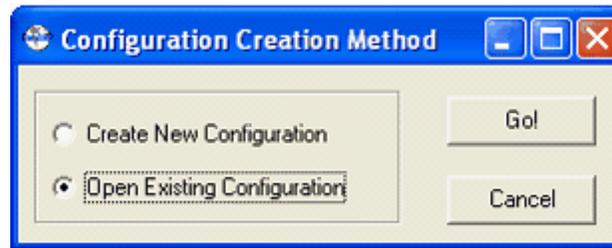
Figure 5-1. Health incidence *Configuration Creation Method* window.

(b) In this window, select "Open Existing Configuration". This means that we are starting from some already selected health impact functions (i.e., we are not "starting from scratch"). Click **Go!**. An *Open* window will appear. In the "Configurations" folder. Select the "PM25_RIA_2020_course.cfg" file and click **Open**.

(c) The *Configuration Settings* window (Figure 5-2) is used to set the baseline and control AQG files, and also to set some general parameters that will be used by all the health impact functions. In this window, we will change only the baseline and control AQG files. The rest of the settings (Latin Hypercube Points, Population DataSet and Year, Point Mode, and Threshold) will be left unchanged.

(d) In the "Select Air Quality Grids" panel, next to the "Baseline File" field, click **Open**. The *Open* window will appear. Navigate to the "PM25_RIA" folder under the "Air Quality Grids" folder, select "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily.aqg", and click **Open**.

(e) Repeat these steps for the "Control File", this time selecting "Control_PM25_RIA_2020_cmaq_grid_14_Annual_35_Daily.aqg".
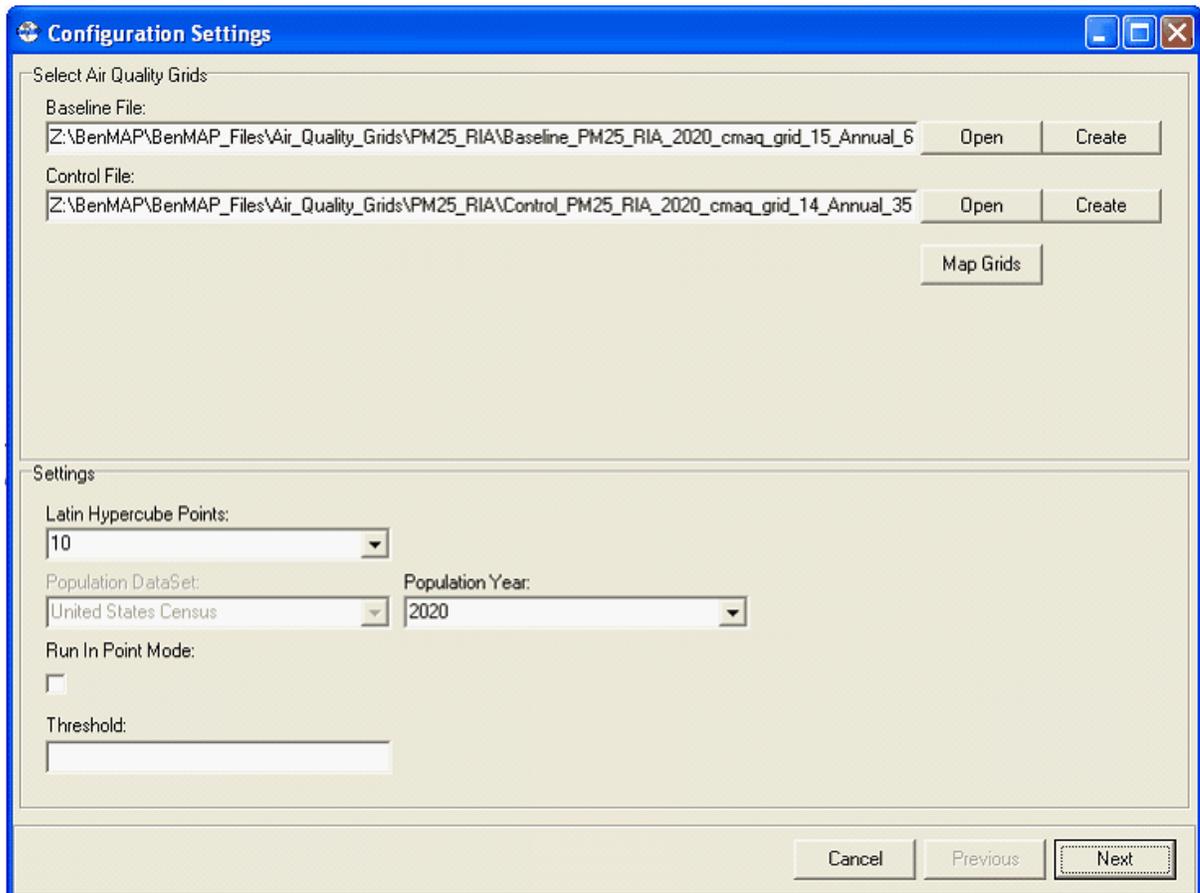
Figure 5-2. *Configuration Settings* **window - General parameters.**

Background: The other settings in this window may be changed when you are doing your own studies. The *Point Mode* versus the *Latin Hypercube Points* options allow you to generate either an average incidence estimate or a range of results. With the *Point Mode* option, BenMAP uses the mean values of the inputs to the health impact functions, and generates a single "point estimate" of the change in adverse health effects; this is useful for quick analyses. With the *Latin Hypercube Points* option, BenMAP estimates a distribution of incidence results that expresses the variability in the incidence estimates; this option reports specific percentiles along the estimated incidence distribution. The greater the number of chosen points, the greater the number of estimates and hence the greater the time needed by BenMAP to process the results. The *Population DataSet* and *Population Year* specify the population data that will be used in the health impact functions. The Population DataSet should match the grid definition of your baseline and control files. The *Threshold* indicates the air quality level below which the program will not calculate benefits. That is, air quality metrics below the threshold will be replaced with the threshold value.

(f) Before continuing on to select health impact functions, we should do some quality assurance (QA) on the aqg files. Specifically, we want to look at the AQ deltas. In the *Configuration Settings* window (Figure 5-2), click on **Map Grids**. This will open a new *BenMAP GIS* window.

(g) In the *BenMAP GIS* window, uncheck the "Control Grid" and "Baseline Grid" in the "Layers" panel (on the left side). Double-click on the "Delta" layer, which will open the *Display Options* window.

> In the *Display Options* window, uncheck "Grid Outline" and set the Variable to "QuarterlyMean". Click **OK**.

> Back in the *BenMAP GIS* window, overlay a State reference layer. Your GIS window should now look similar to Figure 5-3.
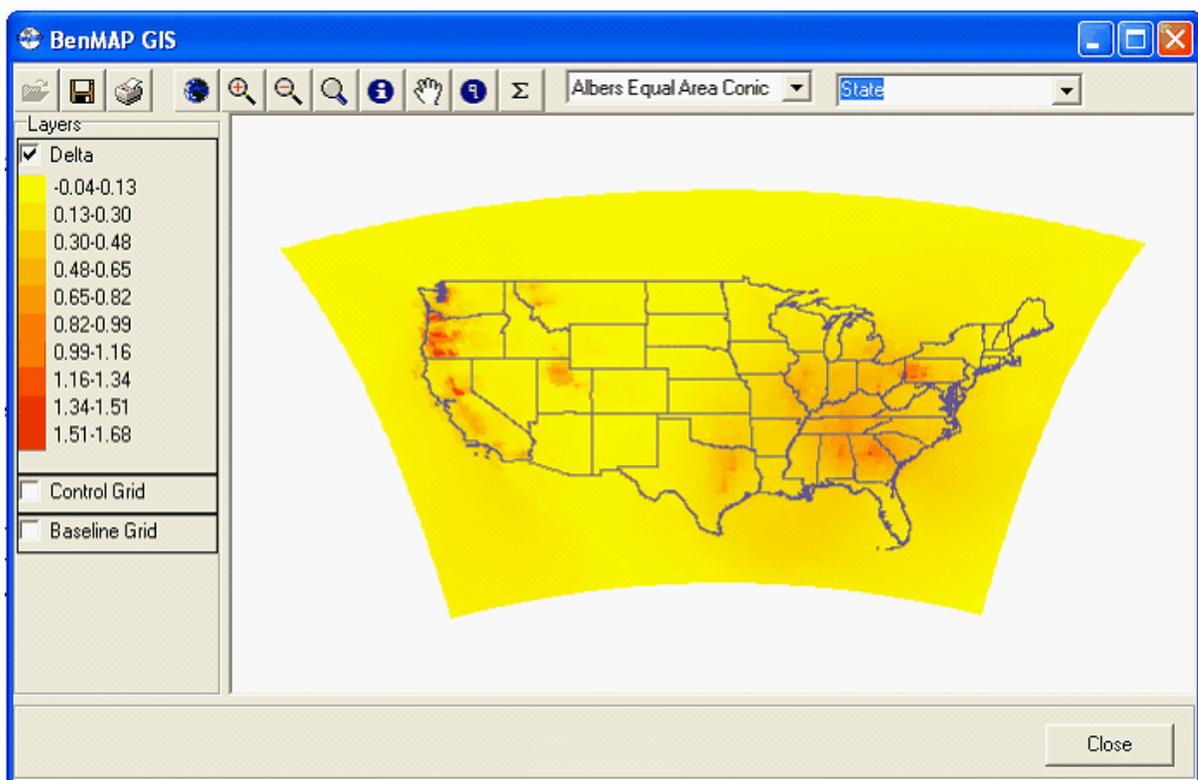
Figure 5-3. Mapping delta AQ

(h) **Exercise (5.1):** Using the **Create Layer Statistics** button, what are the maximum and mean for the QuarterlyMean of $PM_{2.5}$? Using the **Build query** button, which states have a QuarterlyMean greater than 1.0 $\mu g/m^3$?

Answer:

(i) After you have completed the above exercise, click **Close** in the *BenMAP GIS* window to return to the *Configuration Settings* window. At this point we have set all the general parameters for our configuration. *Note*: we did not modify the latin hypercube points, population year, or threshold. Click **Next**.

(j) This will open a new *Configuration Settings* window (Figure 5-4). This window describes the specific health impact functions that were selected in the configuration file. We will also use this window to select new health impact functions for our new configuration. This window is divided into two panels:

Available CR Functions: This describes all the available functions that are appropriate for this type of AQ data. In our case, these are all the $PM_{2.5}$ functions. The left-hand panel is the "Tree". It describes the hierarchy between endpoint groups (major groupings of adverse health effects) and endpoints (specific adverse health effects). You can expand any section by clicking on the plus sign next to the heading, or collapse a section by clicking on the minus sign next to a heading. To see the specific health functions, you need to expand the endpoint heading of choice. The right-hand panel is the "Data" panel. The data are all the details of the specific health functions: the author of the study, location where the study was done, the specific function that BenMAP uses to calculate that adverse health effect, whether there is a qualifier to the health function, etc. A complete description of each column can be found in the User's Manual. The scroll bar at the bottom of the "Available CR Functions" panel is for panning across the "Data" columns.

Selected CR Functions: This describes the functions that have been chosen for this specific configuration. The left-hand panel is "Function Identification", which contains all the columns necessary to uniquely identify the function. The scroll bar at the bottom of the "Selected CR Functions" panel is for panning across the "Function Identification" columns. We recommend focusing on *Endpoint, Author, Year* and *Qualifier*. The right-hand panel is "Function Parameters". These parameters are used by the health functions and some of them can be edited by the user. For example, you might change the age range ("Start Age" and "End Age") for a specific health impact function if you were interested in studying the impact on a certain demographic.
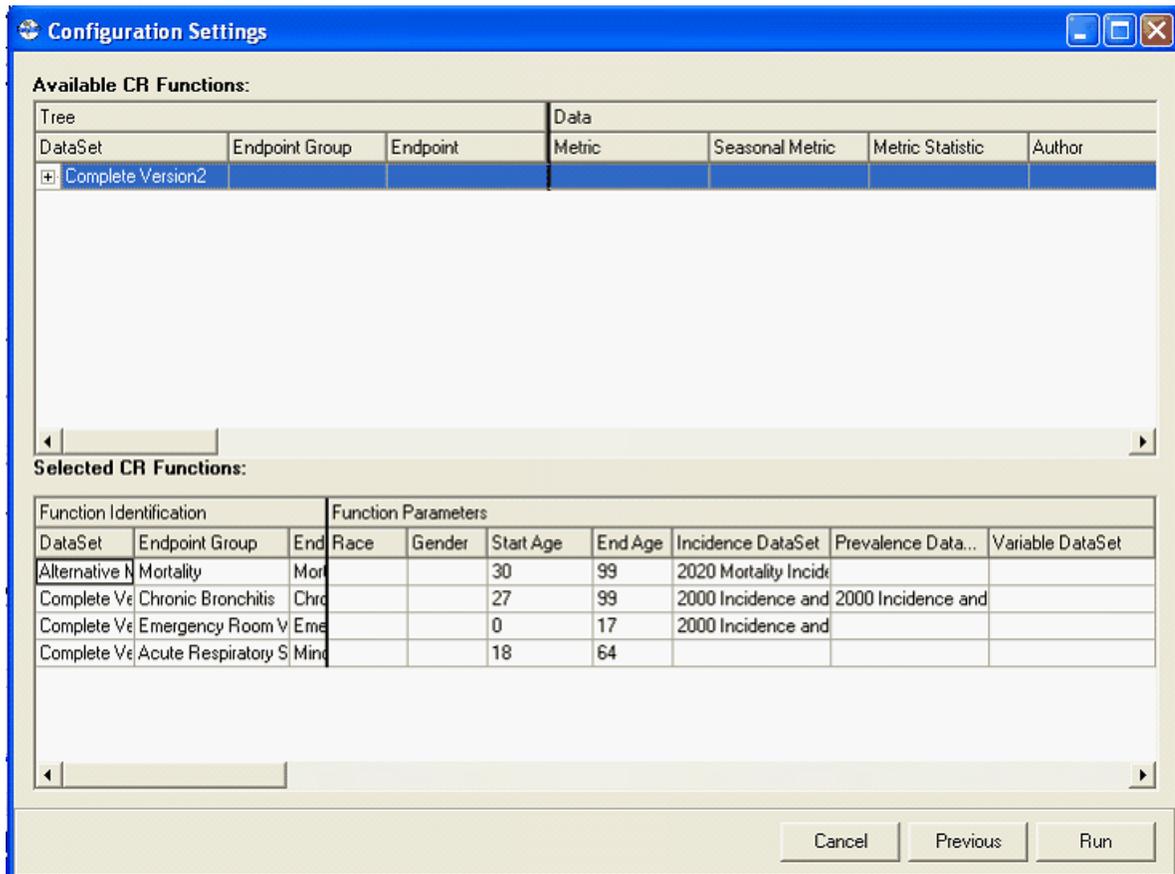
Figure 5-4. *Configuration Settings* **window - Initial functions.**

Tip: When looking at the CR functions, it is often useful to change the order and/or width of the columns. To reorder the columns, click and hold your cursor on the variable name at the top of a column, then drag the column either to the right or to the left. Release your cursor when you have moved the column over the desired location – BenMAP will then rebuild the tree structure using the newly specified variable order. To resize a column, either double click on the right edge of the desired column (maximizes the column) or click on the edge of the column and drag it to the right or left until it is the appropriate size.

(k) Before adding new functions to our configuration, we will look at the current set of selected health impact functions:

(1) In the "Tree" panel, expand the DataSet "Complete Version2". This should reveal the available endpoint groups. You can expand any endpoint group to see the

available endpoints within that endpoint group. You can also expand a specific endpoint so that you can see the specific health functions for that particular endpoint. For example, expand the endpoint group "Emergency Room Visits, Respiratory" and the revealed endpoint "Emergency Room Visits, Asthma" to reveal the specific health impact functions.

(2)**Exercise (5.2):** How many health functions are there for the "Emergency Room Visits, Asthma" endpoint? Who are the author(s) for these functions/studies? What are the differences between these functions? (After you have completed this exercise, you might want to collapse the "Emergency Room Visits, Respiratory" endpoint group. This will reduce the clutter in the top panel.)

Answer:

(3)**Exercise (5.3):** What are the endpoints under the "Hospital Admissions, Respiratory" endpoint group? How many functions are there under the "HA, Chronic Lung Disease" endpoint and who are their authors? *Note*: You might have to expand the width of the *Endpoint* column to make sure you have the right one.

Answer:

(4) **Exercise (5.4):** Who is the author of the mortality function used in our current configuration? *Hint*: Look at the "Function Identification" panel. If you look at the available mortality functions, how many of them are by this author? Looking at the *Qualifier* column, what differentiates this author's functions? Which of the functions are we using in our current configuration? *Hint*: We recommend reorganizing the columns in the "Data" panel so that you have *Author, Start Age, End Age, Qualifier*, and *Function* next to each other (see previous Tip).

<div style="border:1px solid black; height:300px;">

Answer:

</div>

(l) In the next series of steps, we will add new health functions to our configuration. We will add an additional mortality function, some hospital admission functions, and some acute myocardial infarction (AMI) functions. You will note that in many cases there will be multiple functions for the same endpoint by the same author. In most cases these will be differentiated by the content of the *Qualifier* column. In all of our cases, we will select the function that does not have a threshold or other qualifier.

(1) We will start by adding a new "Mortality, All Cause" endpoint health function. Select the function that has Woodruff as the author and no qualifier (no threshold). To add it to our configuration, simply click on the specific row and drag it to the "Selected CR Functions" panel (Figure 5-5).
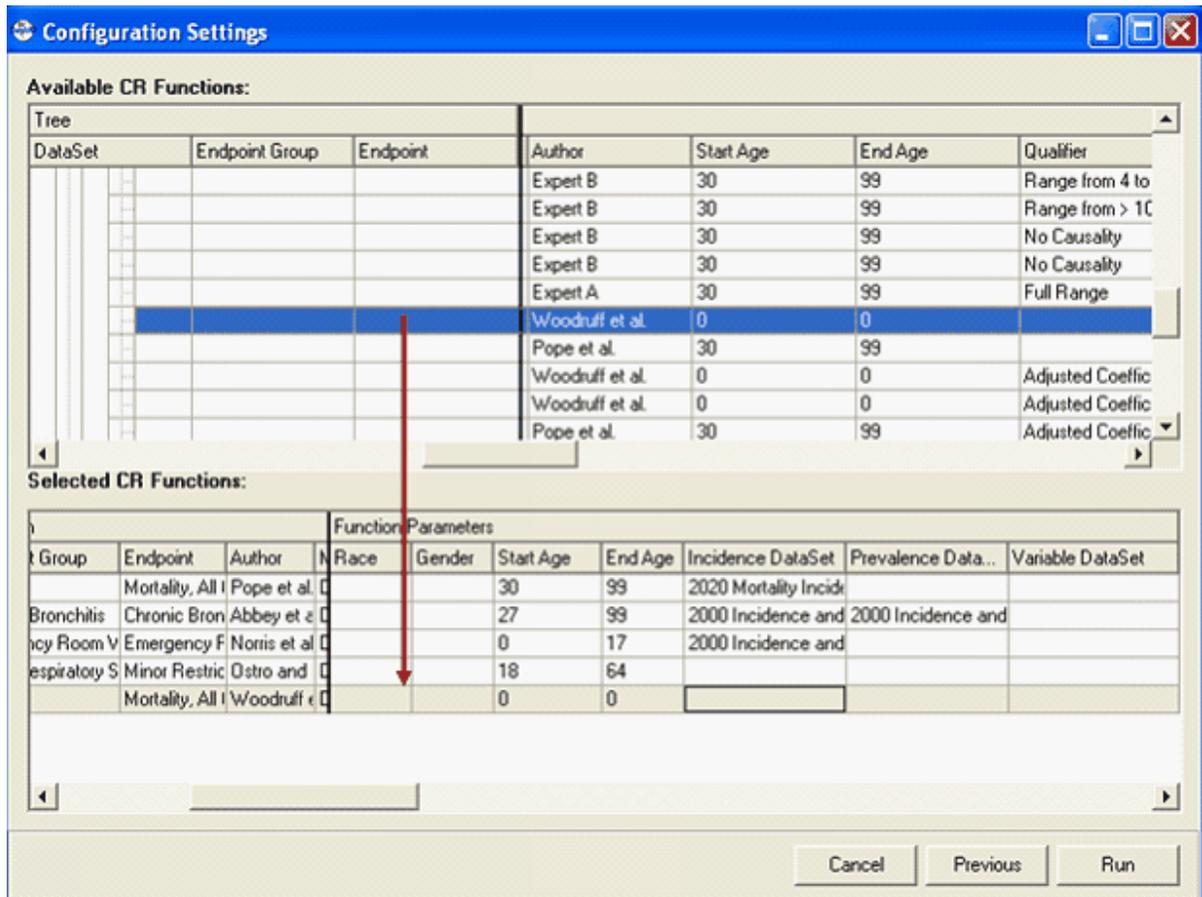
Figure 5-5. Adding a new health impact function.

Tip: If you mistakenly add the wrong health impact function, you can delete it by highlighting the particular function in the "Selected CR Functions" panel and clicking **Delete** on your keyboard.

(2) After moving a function into the configuration, you need to decide whether you want to change any of the "Function Parameters", i.e., the inputs to the health function. In the Woodruff case, we want to set the "Incidence DataSet" (i.e., the background incidence rate for mortality). Click on the "Incidence DataSet" cell for the Woodruff study. Use the drop-down menu to select "2020 Mortality Incidence" (Figure 5-6). In other words, we are using the incidence rate for mortality that has been projected to the year 2020.

Note: In this case there is an available incidence rate dataset for 2020. For

other endpoints (e.g. Chronic Bronchitis), the background incidence rates are only available for 2000 because they have not been projected to 2020.
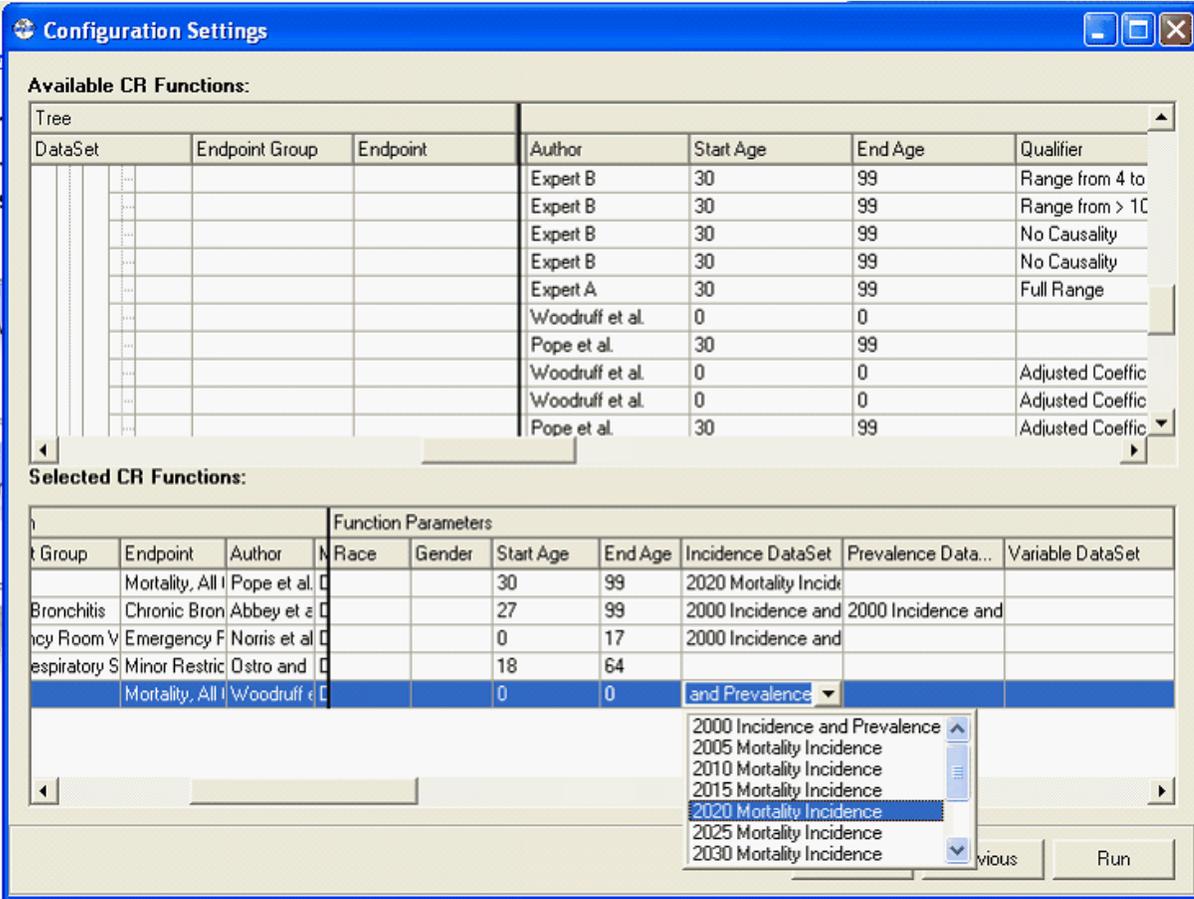
Figure 5-6.Changing Function Parameters - the incidence rate.

(3) **Exercise (5-5):** Compare the Pope and Woodruff mortality studies in the "Available CR Functions" panel. What is the age range for each study? Write down the functions for each study. *Note*: compare the respective Pope and Woodruff functions that you selected.

Answer:

(4) Next we will add a health function for hospital admission due to pneumonia. Under the "Hospital Admissions, Respiratory" endpoint group, expand the "HA, Pneumonia" endpoint. Select the function by Ito without a threshold shown in the qualifier column. Simply click on the appropriate function with your mouse and drag it to the "Selected CR Functions" panel.

> *Note*: In this case the incidence dataset has already been selected. If you try to select another dataset you will see that 2000 is the only available dataset. In other words, the incidence and prevalence for hospital admissions due to pneumonia has not been projected to 2020.

(5) Next we will add health functions for hospital admissions due to chronic lung disease. Under the endpoint "HA, Chronic Lung Disease", add the functions by Ito and Moolgavkar (without threshold). *Note*: Make sure you are **not** using the endpoint "HA, Chronic Lung Disease (less Asthma)".

(6) **Exercise (5-6):** Compare the Ito and Moolgavkar studies. How do their specific functions compare? What is the beta (regression coefficient) for each study? **Bonus question**: Which function is more sensitive to changes in AQ?

Answer:

(7) Finally, we will add health functions for AMI. Under the "Acute Myocardial

Infarction" endpoint group and "Acute Myocardial Infarction, nonfatal" endpoint, select the Peters study (no threshold). Here, we will do something slightly different. Drag *four* copies of the Peters function to the "Selected CR Functions" panel. The age range for this function is 18 to 99 years. We want to look at the effect on a more specific demographic—that is, we want to break this into four age ranges.

> Starting with the first Peters function, select the "End Age" cell in the "Function Parameters" panel. Edit the cell to 44. The new age range for the first Peters function is now 18 to 44, inclusive.

> For the next Peters function, modify the *Start Age* to 45 and the *End Age* to 54. For the remaining Peters functions, change the age range to 55-64 and 65-99, respectively. After modifying the age ranges, your *Configuration Settings* window should look similar to Figure 5-7.
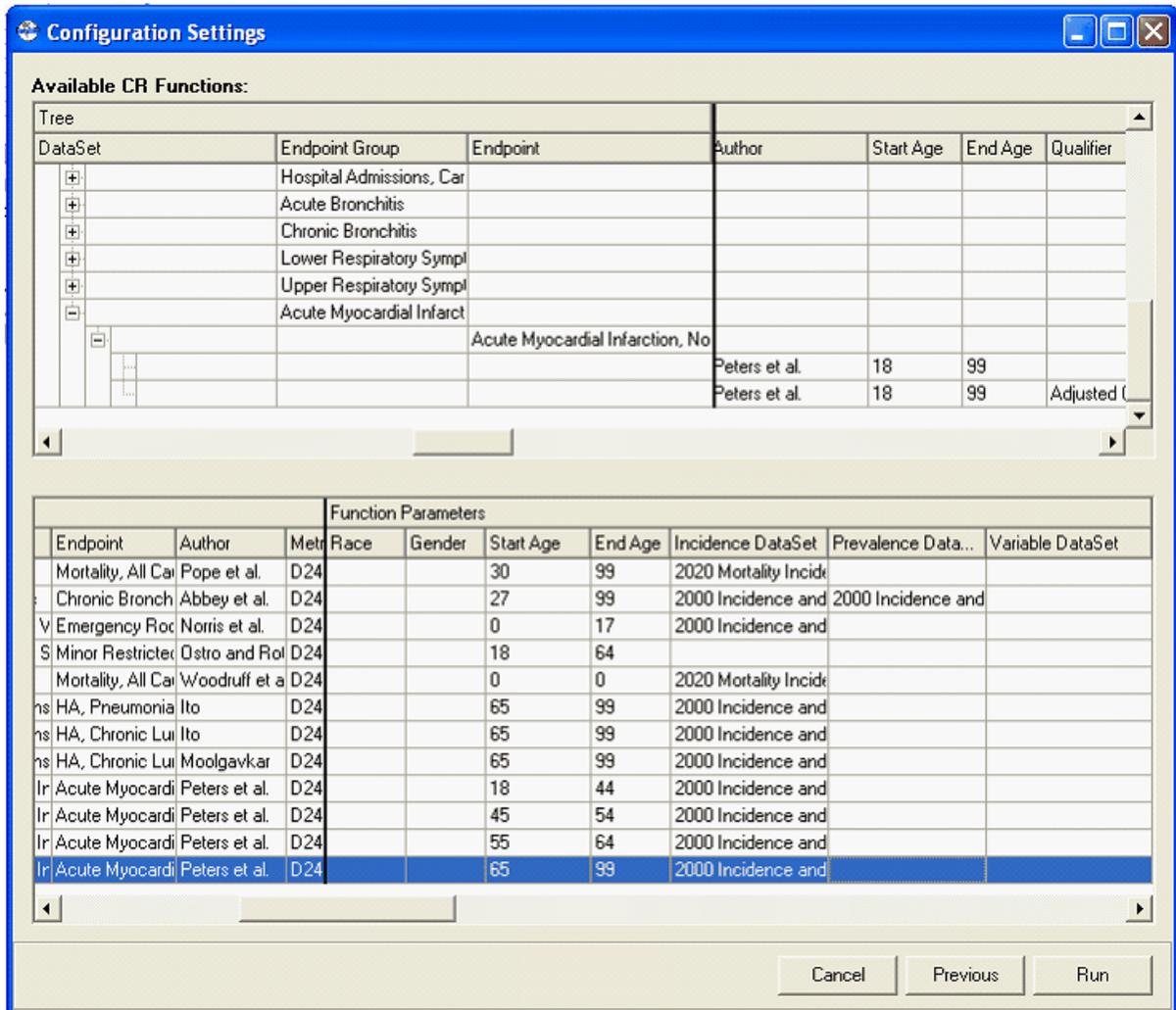
Figure 5-7. Four additional AMI functions with modified age ranges.

(8) You have completed adding all the new health functions to your configuration. You should now have 12 functions in your "Selected CR Functions" panel (Figure 5-7). Double check to be sure all of the selected functions have an empty qualifier cell (check in the "Function Identification" panel by moving the bottom scroll bar until the *Qualifier* column is visible, note that the Pope function will say "no threshold").

(m) When you have set up all the health impact functions as instructed above, you are ready to save the new configuration and generate the health incidence results. Click **Run**. This will bring up a *Save Configuration* window (Figure 5-8).

We want to save this new configuration (so that it can be re-used), so click **Save**. This will bring up a *Save As* window. Under the "Configurations" folder, in the

"File name" field, type in the new configuration file name,
"PM25_RIA_2020_course_modified", and click **Save**. A *Progress* window will
appear.
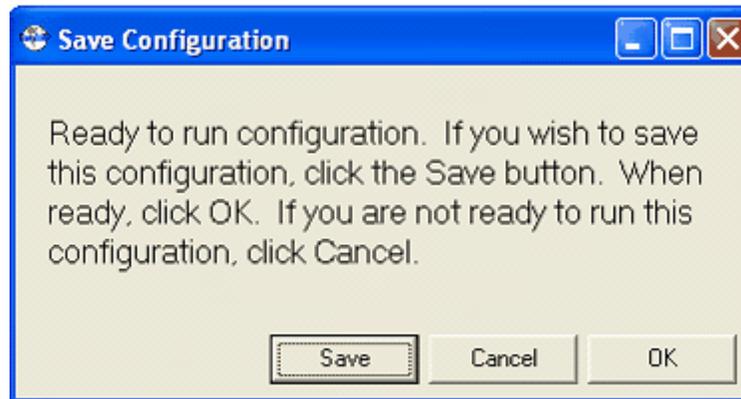
Figure 5-8. *Save Configuration* **window.**

(n) After the configuration is saved, you will be returned to the *Save Configuration* window
(Figure 5-8). Now run the configuration by clicking **OK**. This will bring up another
Save As window. Here we will save the configuration results (cfgr). Under the
"Configuration Results" folder, save the results as
"Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily".

The calculation of the results will begin and a *Progress* window will appear. The
calculation of the results may take a few minutes. When the calculations are
finished, you will be returned to the main BenMAP window.

Tip: If you are not ready to run this configuration, click **Cancel**. If you generated
a onf gurat on you an o en t e onf gurat on and run t to generate a ua
re u t at a ater t e.

Analysis:

The rest of the exercises in Section 5.2 focus on analyzing the results of our BenMAP run.
Specifically, we will look at the newly created health incidence configuration file (cfg) and
results file (cfgr). A good habit to get into is to quality-assure both your configuration and
your results. Some things to check include whether you have the right functions and
parameters selected and whether the results seem reasonable.

(a) First, we will look at the newly created configuration file (cfg) using the audit trail.
From the main BenMAP window, click on the "Report" graphic in the right-hand panel.

Select the "Audit Trail Reports" in the *Select Report Type* window and click **OK**. Under the "Configurations" folder, select the new configuration, "PM25_RIA_2020_course_modified.cfg" and click **Open**.

(a) **Exercise (5.7):** What is the population year used in this configuration? How many health functions (CR functions) were used? What are the location and the incidence rate dataset used in the Woodruff mortality function? *Hint*: if you added the functions in the same order as the lab, the Woodruff study should be "CR Function 4". When you are done with this exercise, click **OK** to close the *Audit Trail Report* window.

Answer:

(b) Using the *Tools* menu in the main BenMAP window, open a *BenMAP GIS* window. Click on **Open a File**, then select "Configuration Results (*.cfgr)". Under the "Configuration Results" folder, open our newly created file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily.cfgr".

(c) In the *Edit GIS Field Names* window, provide more meaningful names for your health incidence results, then click **OK**. Here are suggested names: MortPope, ChronBronc, ER, AcutResp, MortWood, HAPneum, HAChrIto, HAChrMool, AMI18, AMI45, AMI55, AMI65. *Note*: BenMAP GIS field names cannot exceed 10 characters in length and cannot include commas.

(d) **Exercise (5.8):** Compare the Pope and Woodruff mortality results. How do the maximum values compare? The Pope and Woodruff health impact functions are calculating avoided mortalities for different subgroups within the population. What is the Pope result measuring versus the Woodruff result? Do they have similar spatial patterns?

Answer:

(e) **Exercise (5.9):** Compare the Ito and Moolgavkar Hospital Admissions due to Chronic Lung Disease results. How do the maximum values compare? How do their respective populations compare (by "population", we mean the population variable POP1, POP2, etc. that correspond to the appropriate heath incidence result)? The spatial patterns of the two results are identical while their magnitudes differ. How do you explain this? *Hint*: Look back at Exercise 5.6.

Answer:

(f) **Exercise (5.10):** Compare the AMIs (heart attacks) for various age groups. What are the maximum values for the AMI18 and AMI65 functions? What are the population maximum values for these two functions? Why do you think that the value of AMI18 is less than the value of AMI65? After you are done with the above exercises, close the *BenMAP GIS* window by clicking **Close**.

Answer:

(g) From the main BenMAP window, click on the "Report" graphic in the right-hand panel. Select "Raw Incidence Results" in the *Select Report Type* window and click **OK**. Under the folder "Configuration Results", open our results file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily.cfgr".

In the "C-R Function Fields" panel within the "Column Selection" panel of the *Configuration Results Report* window (Figure 5-9), select Endpoint, Author, Start Age, and End Age. In the "Result Fields" panel, deselect Delta and Variance.

As you select or deselect items in the "Column Selection" panel, your choices are reflected in the "Preview" panel in the bottom half of this window, which displays a preview of the columns that will be included in the results report when you save the file.
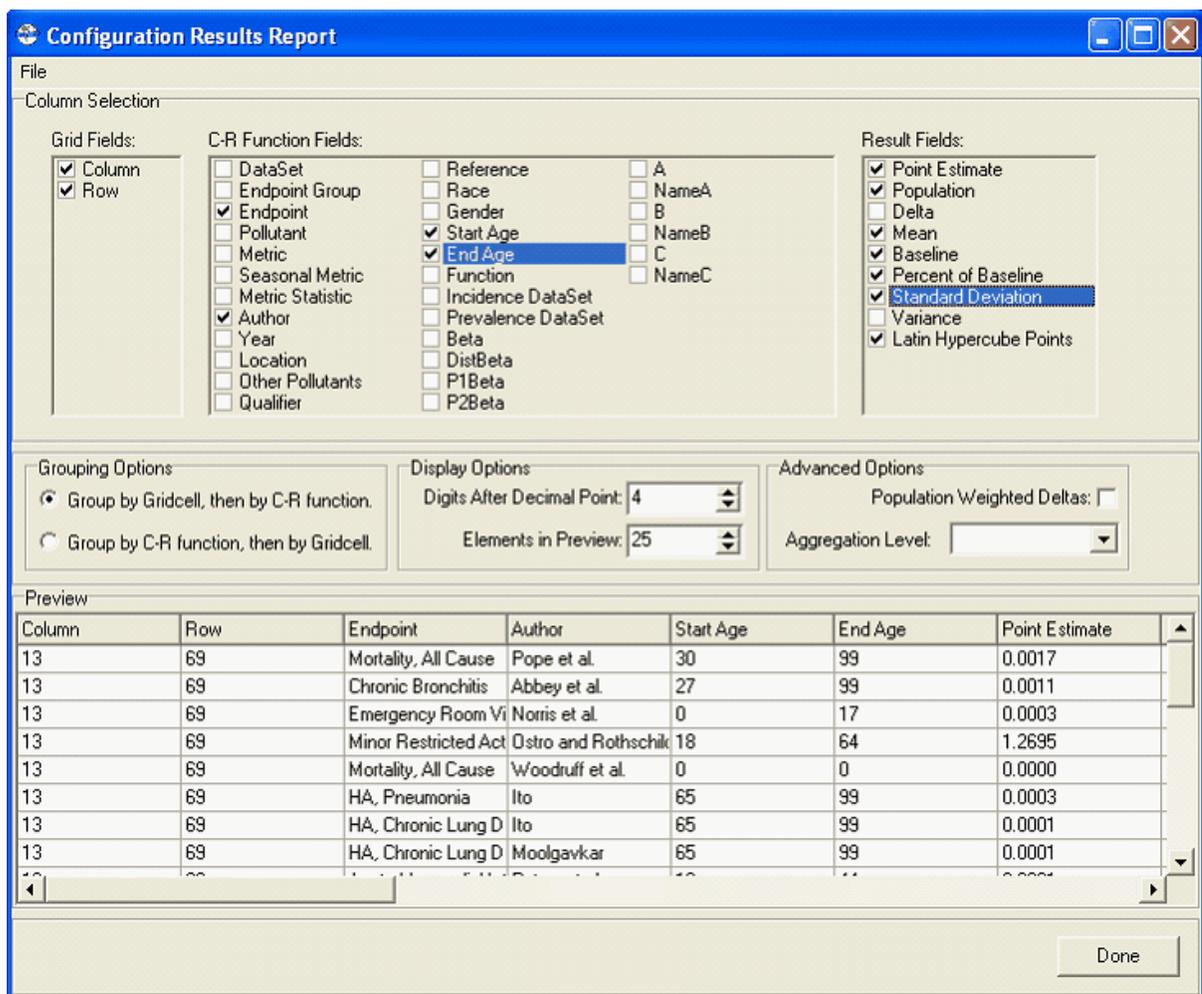
Figure 5-9. Report window for raw incidence results.

Background: Some of the key columns that can be included in results reports are the following: the "mean" is the mean of the Latin Hypercube points for this result; the "point estimate" is the single point estimate for this result; "percentiles" are the individual Latin Hypercube points for this result; the "baseline" is the number of individuals experiencing this adverse health effect due to all causes (typically incidence rate × population); the "percent of baseline" is the relative change in adverse health effects due to the control scenario we are considering (point estimate/baseline); the "population" is the population used in the particular health function at this grid cell.

(h) Now that you have chosen all the columns to include in your results report, you can save the report. From the *Configuration Results Report* window, type **Ctrl-S**. Under the "Reports" folder, save the file as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily". After the file has been saved, close the window by clicking **Done**.

> Use Windows Explorer to navigate to the "Reports" folder. Double-click on the newly created csv file and it should open in Excel. *Note*: if you are using an earlier version of Excel, it may give you an error message saying that it could not load the complete file. Click **OK** in the warning window.

(i) **Exercise (5.11):** All of the following questions refer to the grid cell 14, 63 (column, row) for the endpoint "Minor Restricted Activity Days" (within the endpoint group "Acute Respiratory Symptoms"). What is the background incidence (the total number of minor restricted activity days due to all causes) for this particular grid cell? *Hint*: the column is called "Baseline". What is the change in adverse health effects (i.e., the number of minor restricted activity days avoided) under the control scenario? We know that the heath impact functions have an underlying statistical function that gives us a range of results. What is the estimate of the change in adverse health effects at the 5th percentile? What is the estimate at the 95th percentile? When you are finished with this exercise, close the Excel window.

Answer:

A.1.5.2 Example: PM2.5 Control 2020 14/35 Adjusted

The goal of this exercise is to re-use our newly modified health incidence configuration (from Section 5.2) and create health incidence results for the adjusted control scenario RIA

2020 14 annual, 35 daily µg/m3. Note: in the following instructions, we refer to "adjusted" and "nonadjusted" results, AQ data, and/or configurations. The "nonadjusted" data and configuration are from Section 5.2. The "adjusted" are from this Section (5.3).

Procedures: The following is a list of the necessary steps. We have abbreviated the instructions because they are very similar to the steps you just completed in Section 5.2.

(a) Open "Incidence Estimation" from the main BenMAP window, Custom Analysis side.

(a) Select "Open Existing Configuration", and Click **Go!**. Choose the newly created configuration "PM25_RIA_2020_course_modified.cfg" under the "Configurations" folder and open it.

(a) In the *Configuration Settings* window, change the baseline and control aqg files. Under the "PM25_RIA" folder, select the "Baseline_PM25_RIA_2020_cmaq_grid_15_Annual_65_Daily_adjusted.aqg" as the baseline and the "Control_PM25_RIA_2020_cmaq_grid_14_Annual_35_Daily_adjusted.aqg" as the control. Leave the rest of the settings the same.

(a) Quality-assure the aqg files. Click on the **Map Grids** button. In the new GIS window that opens, select only the "Delta" layer and in the *Display Options* window display the QuarterlyMean.

(a) **Exercise (5.12):** What are the maximum and mean for the delta QuarterlyMean of $PM_{2.5}$? Which states have a delta QuarterlyMean greater than 1.0 µg/m³. Compare your answers to Exercise 5.1. Which scenario (adjusted or nonadjusted) do you predict will have greater number of avoided adverse health incidences? Explain your answer.

Answer:

(a) After returning to the *Configuration Settings* window, click **Next**.

(a) Review your selected health functions. There should be 12 of them. Make sure they are the same functions you used in the last example then run the example by clicking **Run**.

(a) **Do** *not* **save the configuration**. There is no need to save it because the only changes were to the baseline and control files. Do create the health incidence results by clicking **OK**. Under the "Configuration Results" folder, save the results as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_adjusted". The calculation of the results will begin and a *Progress* window will appear. The calculation may take a few minutes.

Analysis: Now we will look at the newly created health incidence results file (cfgr). Because we have already checked the cfg file in Section 5.2, we do not need to quality-assure this configuration file again.

(a) Click on the right-hand "Report" graphic in the main BenMAP window. Use the *Select Report Type* window to open an audit trail for the new health incidence results file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_adjusted.cfgr".

(b) **Exercise (5.13):** What is the population year used in this cfgr? How many health functions (C-R functions) were used? How many Latin Hypercube points are used in calculating the C-R functions' statistical distribution?

Answer:

(c) Open a *BenMAP GIS* window then open the same health incidence results file under the "Configuration Results" folder.

(d) **Exercise (5.14):** Compare the Pope and Woodruff mortality results. How do the maximum values compare? How do these results compare to the nonadjusted results (see Exercise 5.8)?

Answer:

(e) **Exercise (5.15):** Compare the AMIs for various demographics. What are the maximum values for the AMI18 and AMI65 functions? What are the population maximum values for the two functions? How do these results compare to the nonadjusted results (see Exercise 5.10)?

Answer:

(f) From the main BenMAP window, follow the steps needed to create a raw incidence results report from the same health incidence results file. Select the same results columns as before (see Figure 5-8).

Save the report under the "Reports" folder as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_adjusted". Then open the new csv file in Excel.

(g) **Exercise (5.16):** All of the following questions refer to the grid cell 14, 63 (column, row) for the endpoint "Minor Restricted Activity Days" (within the endpoint group "Acute Respiratory Symptoms"). What is the background incidence (the total number of minor restricted activity days due to all causes) for this particular grid cell? What is the change in adverse health effects (i.e. the number of minor restricted activity days avoided) under the control scenario? We know that the heath impact functions have an underlying statistical function that gives us a range of results. What is the estimate of the change in adverse health effects at the 5th percentile? What is the estimate at the 95th percentile? How do these results compare to the nonadjusted results (see Exercise 5.11)?

Answer:

(h) Close the Excel file and any open BenMAP windows. This completes the "Health Incidence" lab. In the next lab, "Aggregation, Pooling, and Valuation," we will take our health incidence results and calculate the corresponding monetized benefits due to these health effect changes.

A.1.6 Section 6. Aggregation, Pooling, and Valuation

In this section, you will create an aggregation, pooling, and valuation (APV) configuration and use it to produce new valuation results. You will create two separate sets of valuation results based on the same configuration and two similar, but distinct, control strategies.

Creating valuation results has four main stages:

(1) Select a health incidence results file (cfgr) and set up pooling for similar results (i.e. combining similar results together into one result).

(2) Select specific valuation functions and pool similar valuations.

(3) Select additional parameters for the valuation functions, and decide on the aggregation (e.g. summing results from county level to state level) for the health incidence results and the valuation results.

(4) Save all the settings from the first three stages as a configuration file (apv), which can be re-used later with other health incidence results if desired. Finally, run the aggregation, pooling, and valuation configuration, which will create a valuation results file (apvr).

A reduction in air quality level is usually associated with lowering the risk for adverse health effects for a population. This reduction in risk is usually not the same for all individuals, and there is a need to translate the reduction in risk to a quantifiable economic value. BenMAP relies upon published studies where the unit value of such a reduction in risk has been calculated for various health effects. Since multiple studies are sometimes available for a given health incidence, the user needs to choose between them, or adopt techniques to pool (statistically combine) the different functions in an appropriate manner. In this section, we will learn how to pool the results and monetize the reduction in risk for adverse health effects due to changes in air quality levels.

Note: For many of the health endpoints (e.g., mortality), there are many different valuation functions that you could choose to include in your configuration. In addition, there are multiple ways to pool your health incidence and valuation results. This lab has you select specific functions and teaches you how to differentiate these valuation functions and modify certain parameters. This lab does not teach you how to determine which valuation functions and pooling options are the best for a particular study. To determine the best choices for a particular analysis, we recommend that you read the appendices accompanying BenMAP that describe the specific studies that correspond to the specific valuation functions and/or that you discuss your choices with an economist.

A.1.6.1 Example: PM2.5 Control 2020 14/35

The goal of this exercise is to create a new aggregation, pooling, and valuation configuration (apv) and produce valuation results for the control scenario RIA 2020 14 annual, 35 daily μg/m³.

Procedures:

(a) In the main BenMAP window, begin the process of creating an apv configuration by clicking on the graphic titled "Pooling, Aggregation, and Valuation" in the right-hand panel, under the "Step 3" heading. This will open the *APV Configuration Creation Method* window (Figure 6-1).
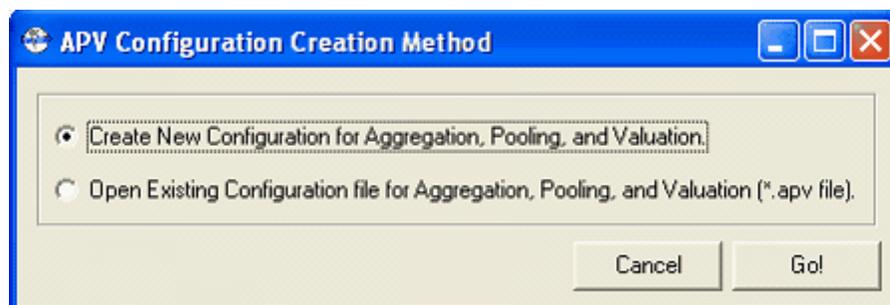
Figure 6-1. *APV Configuration Creation Method* **window**

(b) In the *APV Configuration Creation Method* window (above), select "Create New Configuration for Aggregation, Pooling, and Valuation". Click **Go!**. An *Open* window will appear.

We first have to select the health incidence results that this apv will be based on: the results that we created in Section 5.2. In the "Configuration Results" folder, select "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily.cfgr" file and click **Open**. This will open the *Incidence Pooling and Aggregation* window (Figure 6-2).
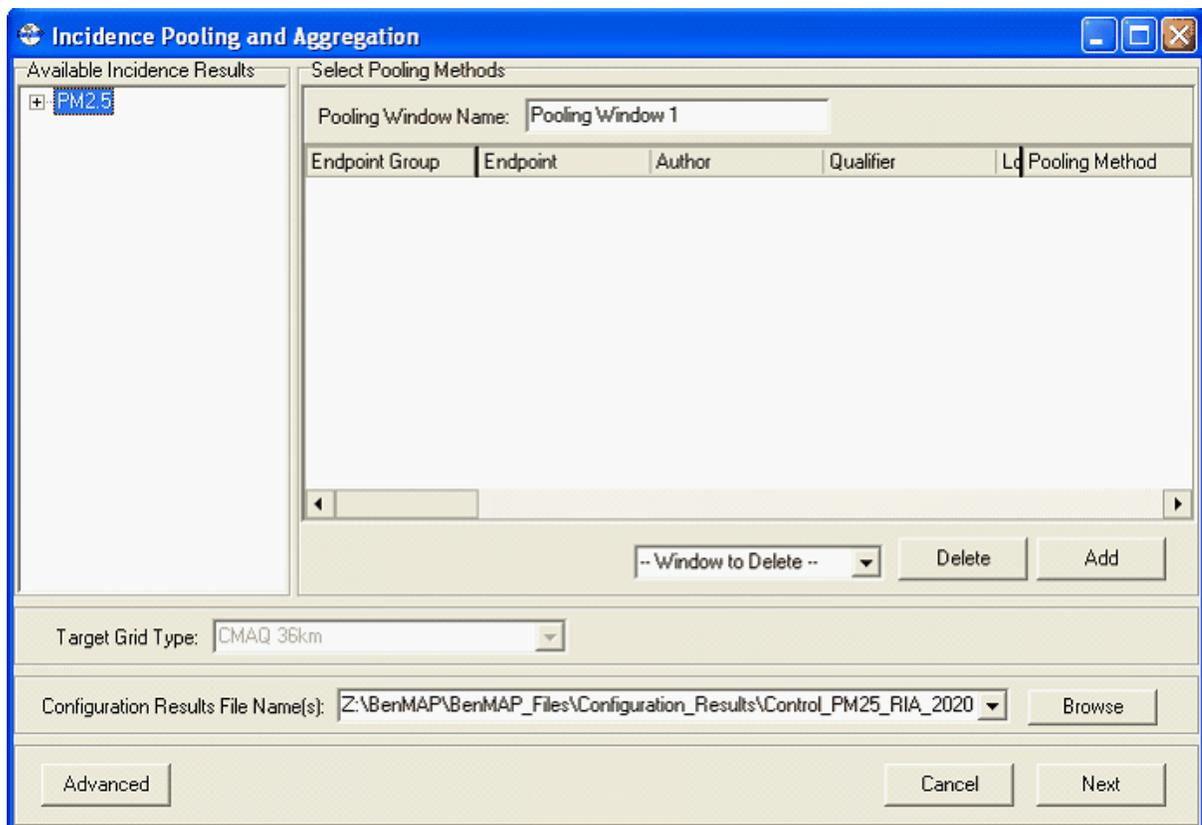
Figure 6-2. APV *Incidence Pooling and Aggregation* window

(c) This window is where we will combine (pool) similar health incidence results together. Before we begin pooling, let us look at the main features of this window. First, the "Configuration Results File Name(s)" field near the bottom of the window shows the health incidence results (cfgr) that you just selected for this apv. If you want to use a different cfgr, you would use the **Browse** button to locate it. We will do this later, in the second example for this lab.

The "Target Grid Type" is the grid definition used to generate the selected health incidence results; it always matches the cfgr input's grid. The "Available Incidence Results" panel has an expandable hierarchical tree that lists each of the health incidence results in the cfgr file. The "Select Pooling Methods" panel has one or more pooling windows. The pooling windows are where you will select individual health incidence results that you may want to pool. These selected results (whether or not they have been pooled) will then be available for valuation (next window). The **Add** and **Delete** buttons add new pooling windows or remove pooling windows that have already been created.

(d) First, change the name of the pooling window. (For our configuration, there will be only one pooling window.) Click in the "Pooling Window Name" field within the "Select Pooling Methods" panel. Change the text "Pooling Window 1" to "Main Pooling Window".

(e) In the "Available Incidence Results" panel, expand "PM2.5" by clicking on the plus [+] sign. This will reveal the available endpoint groups. You can expand any endpoint group to see the available endpoints within that endpoint group. In turn, you can expand a specific endpoint so that you can see the specific health incidence results available for that endpoint.

We will begin by expanding the "Chronic Bronchitis" endpoint group and endpoint. Notice that the health incidence result is identified in the tree using the reference that describes the study used to derive the underlying health incidence function. Click on the chronic bronchitis result "Abbey D. E., B. E. Ostro, F. Pertersen and R. J. Burchette. 1995. …" and drag it to the "Main Pooling Window" (Figure 6-3). That is all it takes—you have added the first health incidence result.
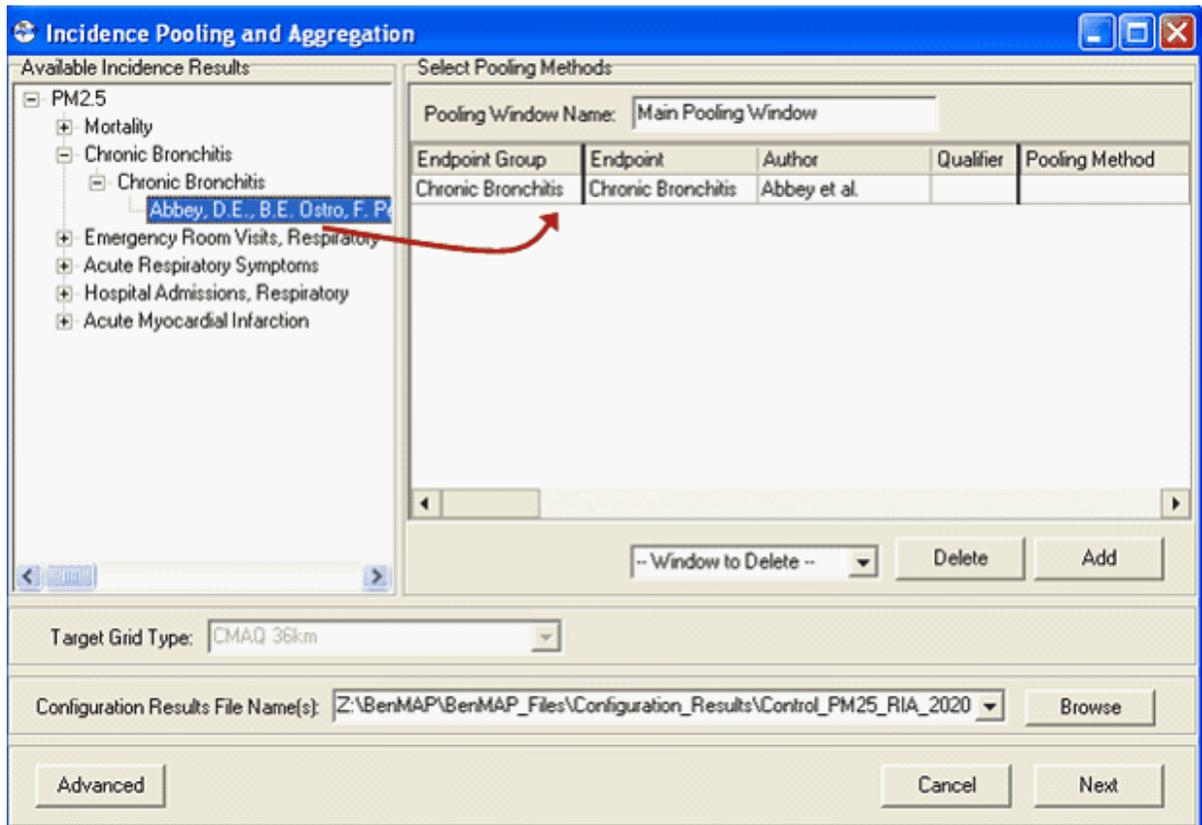
Figure 6-3. Adding a health incidence result to a pooling window

---

Tip: By hovering the cursor over the Available Incidence, you can see a more complete description of these studies.

(f) For the "Emergency Room Visits, Respiratory" endpoint group, add the health incidence result "Norris G., et al. …" to the "Main Pooling Window". Again, simply click on the specific result from the "Available Incidence Results" panel and drag it to the pooling window.

(g) For the "Acute Respiratory Symptoms" endpoint group, add the health incidence result "Ostro, B. D. and S. Rothschild …" to the "Main Pooling Window". At this point you should have three endpoints with their corresponding three health incidence results in the pooling window (Figure 6-4).
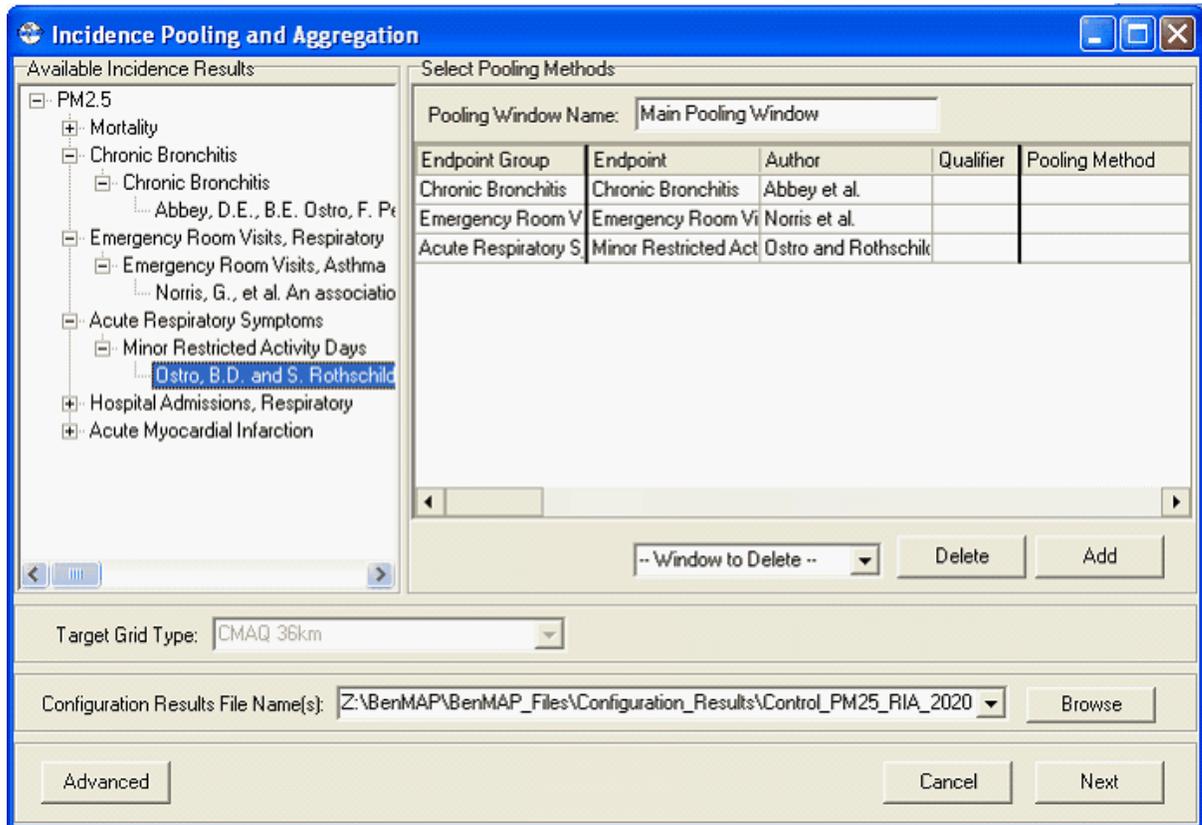
Figure 6-4. Initial three health incidence results in the pooling window

Tip: If you mistakenly add the wrong health incidence result, you can delete it by highlighting the particular function in the "Select Pooling Methods" panel and clicking **Delete** on your keyboard. The **Delete** button in the *Select Pooling Methods* panel will remove the whole pooling window. Typically, you would not want to do this.

(h) For the rest of the endpoint groups, we will add multiple health incidence results per group. In the "Available Incidence Results" panel, expand the "Mortality" endpoint group. Drag the "Pope et al.,…" result into the "Main Pooling Window". Repeat for the "Woodruff, T.J., J. Grillo…" result (Figure 6-5).
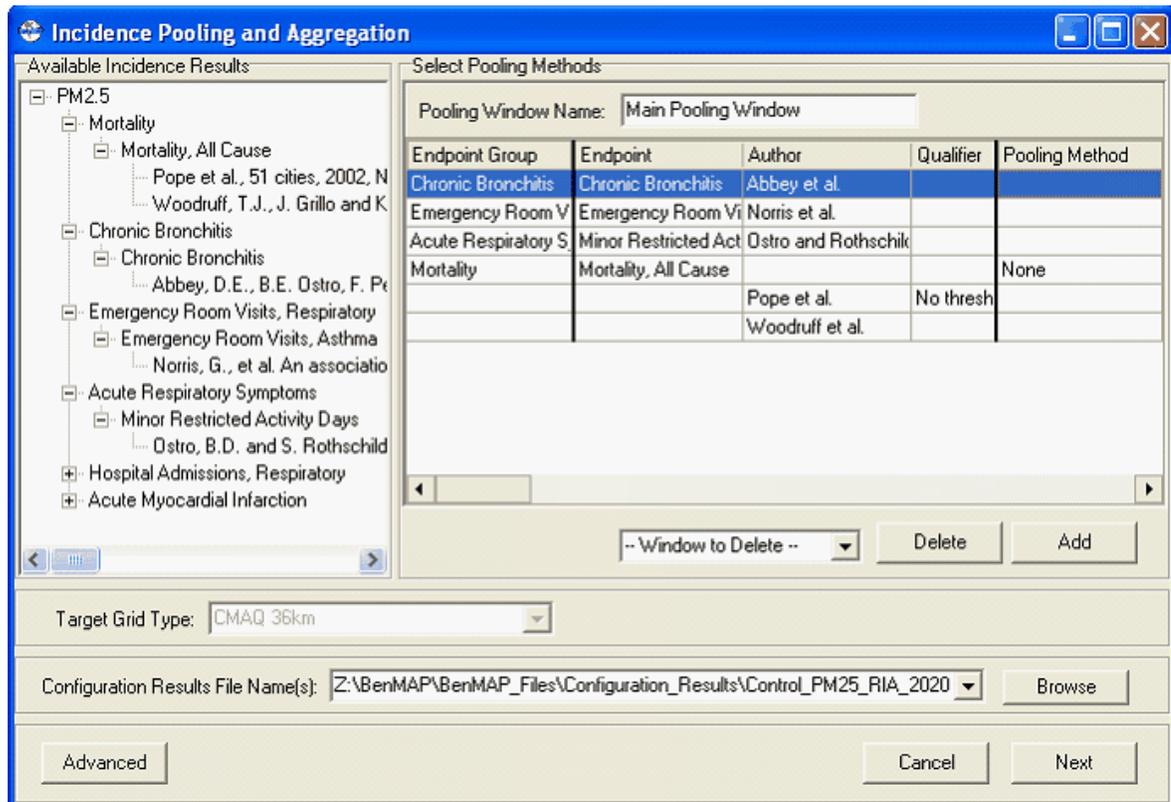
Figure 6-5. Addition of mortality health incidence results to the pooling window

(i) Now, please note a couple of points about the present state of the pooling window. First, we have begun to create a hierarchy. Under the "Mortality" endpoint group we have one endpoint, "Mortality, All Cause". Under this endpoint we have two health incidence results, "Pope" and "Woodruff". You can use the scroll bar to pan across the various columns of data describing the specific results (similar to what you did in the "Available CR Functions" panel of the health incidence "Configuration Settings" window (Figure 5-7)).

Second, a "None" appeared in the *Pooling Method* column of the "Main Pooling Window" in the "Mortality" row. When there are multiple health incidence results under a single endpoint group, BenMAP allows you to pool the results. Unlike the other endpoint groups, the "Mortality" group has two results; therefore, you have the option of pooling them. By default, the pooling is set to "None"—that is, no pooling. We do want to use pooling, but before setting up the pooling we are going to add the rest of the health incidence results for the remaining endpoint groups.

> *Tips*: (1) You can grab a whole group of endpoints and add all of their respective results to the pooling window at once. (2) You do not need to add all of the available health incidence results to your pooling window(s). If you are not interested in pooling or valuing a particular result, do not add it to a pooling window.

(j) For the "Hospital Admissions, Respiratory" endpoint group, add these three items to the pooling window:

the "Ito" results for the "HA, Pneumonia" endpoint

Both the "Ito" and the "Moolgavkar" results for the "HA, Chronic Lung Disease" endpoint.

(k) For the "Acute Myocardial Infarction" endpoint group, add the four age-specific "Peters" results to the pooling window (Figure 6-6). *Note*: although the four functions seem to be missing in Figure 6-6, they are in the pooling window. You need to use the scroll bar in the right hand panel to see the specific "Start Age" column.
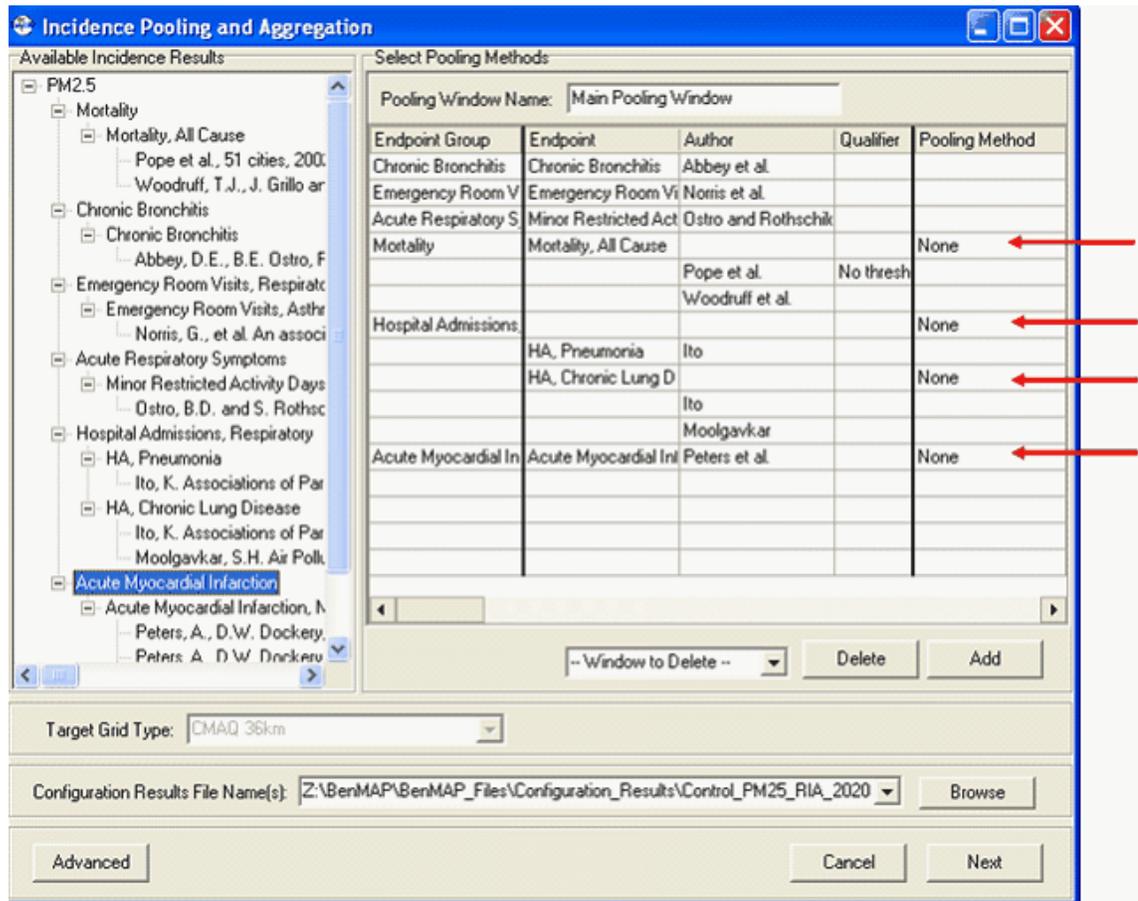
Figure 6-6. All health incidence results added to the pooling window, with no pooling

(l)   The "Main Pooling Window" has four places we could potentially pool the results (indicated by the red arrows in Figure 6-6).

(m) Starting with the "Mortality" endpoint group, click on the "None" in the corresponding *Pooling Method* column. A drop-down menu will appear that lists all the possible pooling methods. Select the "Sum (Dependent)" method (Figure 6-7). This will cause the results from the Pope and Woodruff studies to be summed together to create a single mortality result.
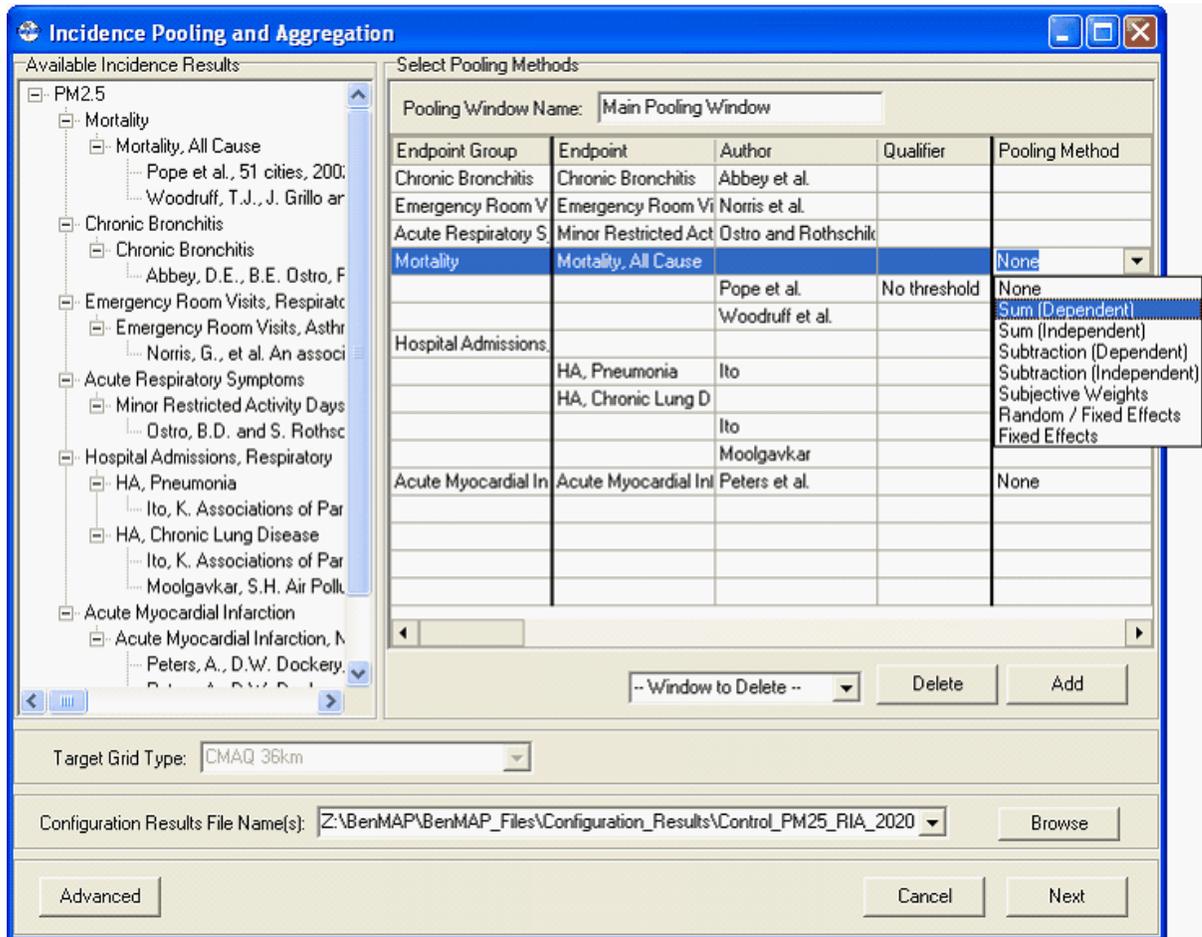
Figure 6-7. Adding a pooling method for the mortality results

Background: In the mortality case, we are using *Sum* because the results from the two studies are distinct from each other. The number of avoided adult mortalities (Pope) does not overlap with the number of avoided infant mortalities (Woodruff)—i.e., they are distinct age groups. See "Pooling Approaches" in the appendices.

(n) Now we will add pooling to the "Hospital Admissions, Respiratory" results. There are two levels of pooling that can be done in this endpoint group, because it contains multiple endpoints ("HA, Pneumonia" and "HA, Chronic Lung Disease"), and one of those endpoints contains multiple results ("Ito" and "Moolgavkar"). Starting with the bottom of the hierarchy, we will combine the "HA, Chronic Lung Disease" endpoint's results ("Ito" and "Moolgavkar"). After combining those results into a single result, we will take this new, pooled result and combine it with the "HA, Pneumonia" endpoint to create a single, combined result for the whole

endpoint group.

First, click the "None" corresponding to the "HA, Chronic Lung Disease" endpoint. Select the "Random / Fixed Effects" pooling method from the drop-down menu. This combines the two results into one pooled result for the "HA, Chronic Lung Disease" endpoint.

Next, click the "None" next to "Hospital Admissions, Respiratory" endpoint group. Select the "Sum (Dependent)" pooling method (Figure 6-8). With this step, we have pooled the different endpoint results into one result for the endpoint group.
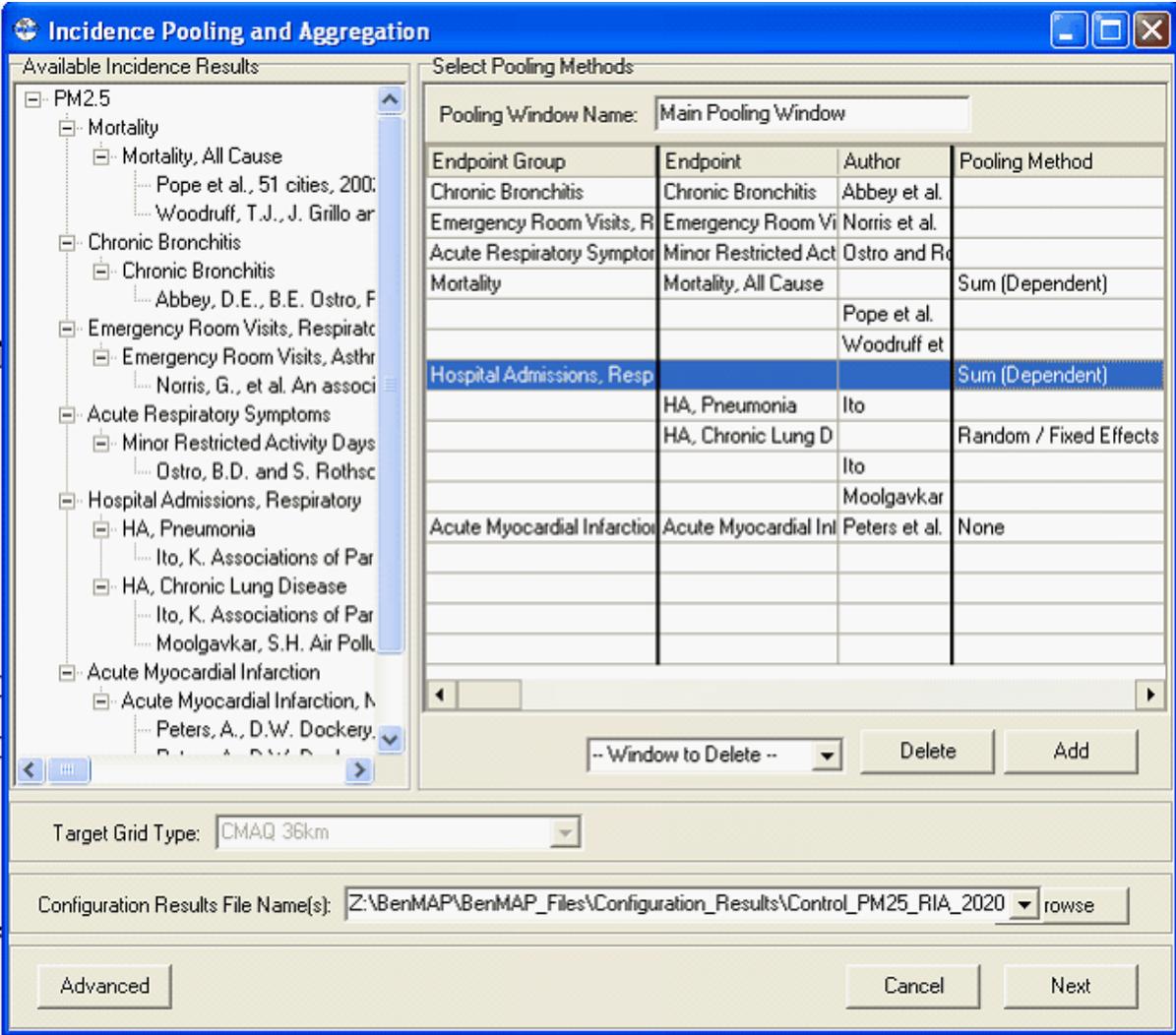
Figure 6-8. Adding pooling methods for the hospital admission results

Background: In calculating the combined HA chronic lung disease results, we are looking at overlapping populations (same endpoint, same age range); therefore, we want to use *Random/Fixed Effects* to combine their distributions. In contrast, when we pool the HA endpoint group, we are looking at nonoverlapping populations: pneumonia versus chronic lung disease (i.e., different endpoints). Therefore, we want to pool the pneumonia and chronic lung disease distributions by doing a *Sum*. See "Pooling Approaches" in the appendices.

(o)   The "Random / Fixed Effects" pooling method has advanced settings. To access these settings, double-click on the "Random / Fixed Effects" cell in the *Pooling Method* column. This will bring up an *Advanced Pooling Options* window (Figure 6-9). There are multiple options for customizing this type of pooling (See the user's guide for specifics). We will not change the existing settings in this window. Click **OK** to close it.
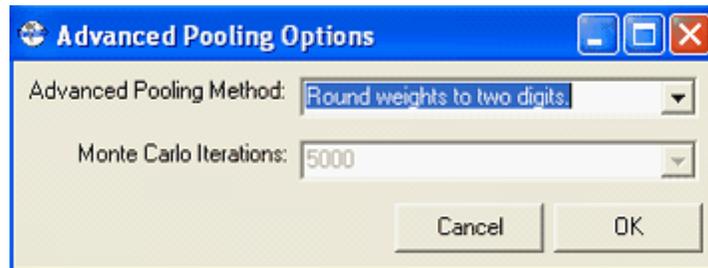
Figure 6-9. *Advanced Pooling Options* **window**

(p)   The "Acute Myocardial Infarction" endpoint group is the final group for which pooling could be done. In this configuration, however, we will not pool the four age-specific AMI results. Leaving the results unpooled will allow us to take advantage of age-specific valuation functions in the next stage of the APV configuration setup. In other words, we will have four separate AMI results to value instead of one single, combined result. To not pool the results, leave the "Pooling Method" for the AMI endpoint group as "None".

We have now finished setting up the health incidence results pooling, which is the first of the four stages in creating valuation functions (see the stages list at the beginning of Section 6). Your pooling window should look like Figure 6-8.

(q)   **Exercise (6.1):** If you were going to pool the four AMI results, which pooling method would you use? Why?

Answer:

(r) Now we move on to the second stage of creating valuation functions: choosing the specific valuation functions and pooling similar valuations. Click **Next** in the *Incidence Pooling and Aggregation* window. This will open the *Select Valuation Methods, Pooling, and Aggregation* window (Figure 6-10).
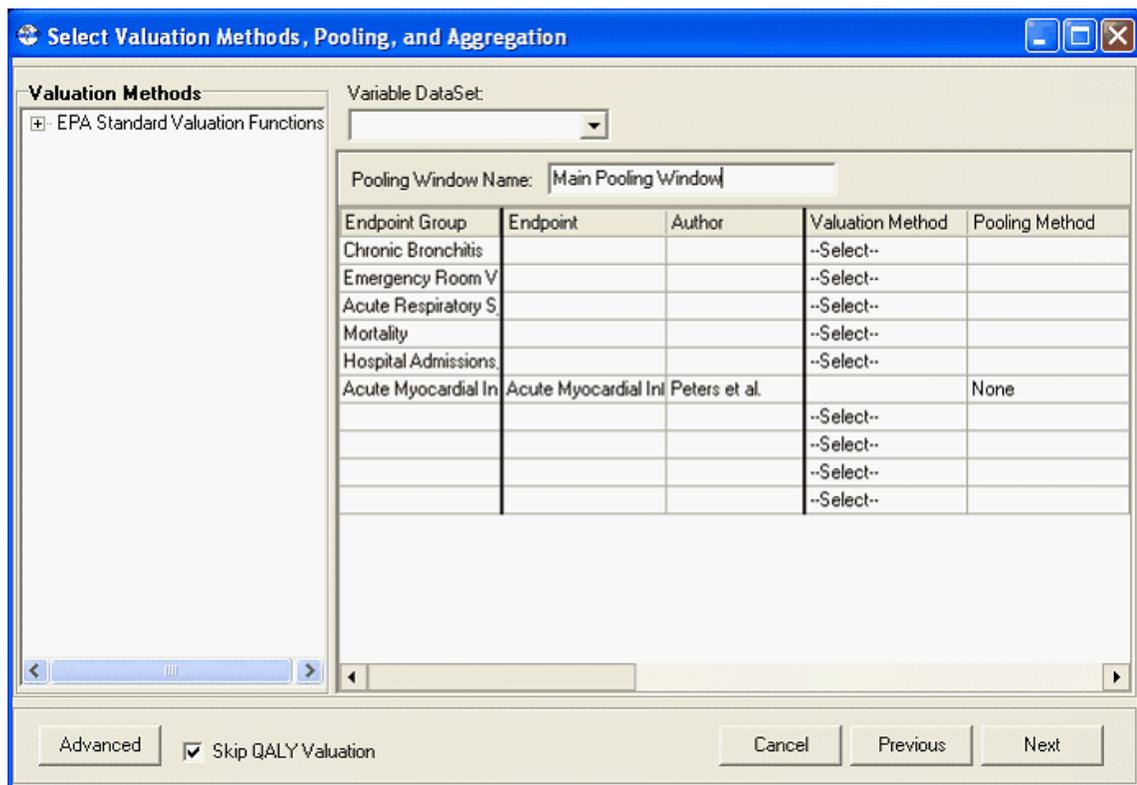
Figure 6-10. *Select Valuation Methods, Pooling, and Aggregation* **window**

(s) This window is where we will apply valuation functions to our health incidence results and combine (pool) similar valuation results together. Before beginning, let us look at the main features of this window. The "Valuation Methods" panel has an expandable hierarchical tree that lists the available valuation functions, based

on the incidence results in this configuration. The right-hand panel contains the pooling windows that were defined in the previous step. In our case, there is only one pooling window, "Main Pooling Window". If there had been three pooling windows created in the previous health incidence pooling step, then we would have three pooling windows in the *Select Valuation Methods, Pooling, and Aggregation* window.

 The "Variable DataSet" field at the top of the window defines specific variables used in the valuation functions. The "Skip QALY Valuation" checkbox at the bottom of the window determines whether or not we will configure and run QALY (Quality Adjusted Life Years) functions. In our case we will not be calculating QALY since we are interested in quantifying the benefits in dollars.

Background: In the context of air pollution benefit analysis, the QALY represents the combined mortality and morbidity benefits of some air quality change. This combined metric is calculated by counting: (1) the number of life years gained; (2) the number of life years lived without some chronic condition. In step 2, the life years are weighted according to the severity of the condition (the "quality" of that year), such that a year in near perfect health might be counted as 0.9, but a year lived with chronic bronchitis might be counted as 0.5.

(t)   In the "Valuation Methods" panel, expand "EPA Standard Valuation Functions". This will reveal the available endpoint groups. You can expand any endpoint group to see the available endpoints within that group. In turn, you can expand a specific endpoint so that you can see the specific valuation functions available for that endpoint.

We will begin by expanding the "Chronic Bronchitis" endpoint group and endpoint. You will see six COI (Cost Of Illness) and one WTP (Willingness To Pay) functions. We will add the WTP function for average bronchitis severity ages 30 to 99. Click on the "WTP: average severity | 30-99" and drag it to the "Main Pooling Window" to the "Chronic Bronchitis" endpoint group's row (Figure 6-11). That is all it takes—you have added the first valuation function.
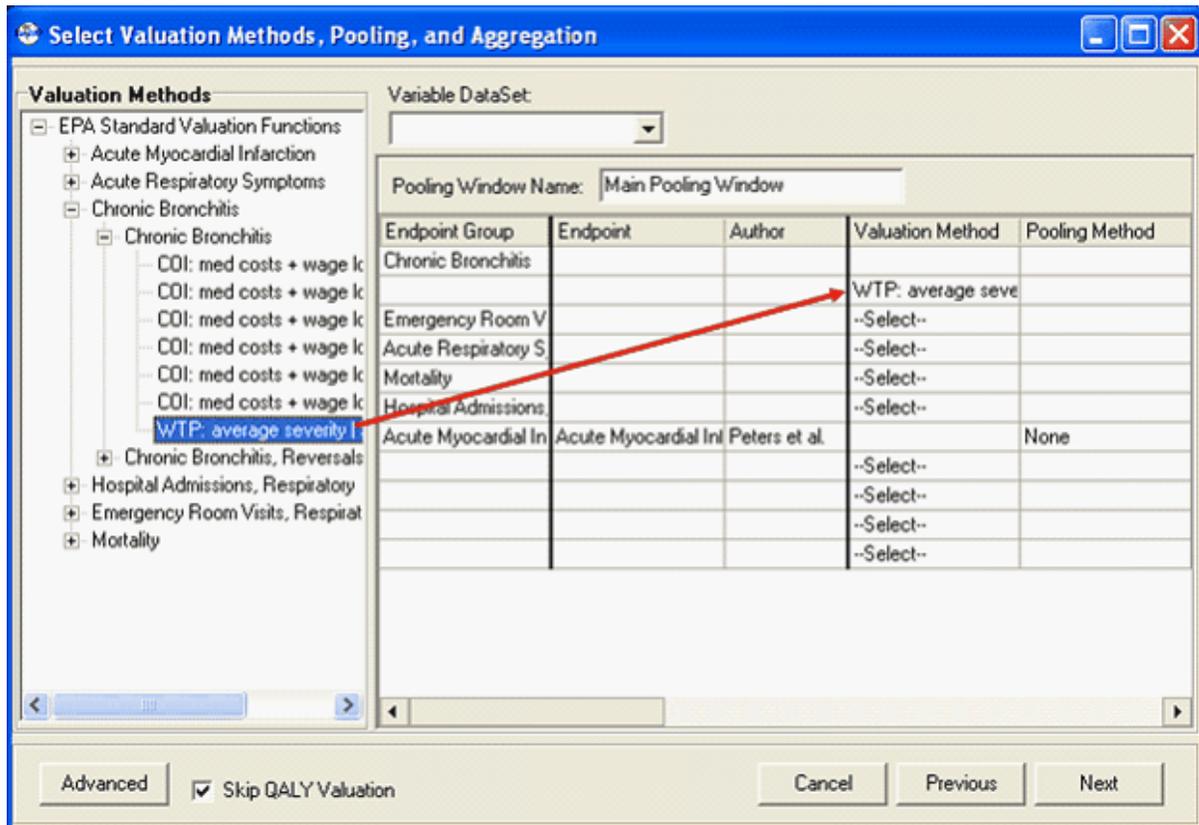
Figure 6-11. Adding a valuation function for chronic bronchitis to the pooling window

> *Tip*: By hovering your cursor over the valuation functions, you can read their full names. If you drag and drop a function and nothing happens, then you probably tried to drop it into the wrong endpoint. Retry with the correct endpoint.

(u) Next we will add a valuation function for acute respiratory symptoms. Under the "Acute Respiratory Symptoms" endpoint group in the *Valuation Methods* column, expand the "Minor Restricted Activity Days" endpoint. Add the WTP valuation function for one day of lost work based on a contingent valuation (CV) study for ages 18 to 99 by clicking on the "WTP: 1 day, CV studies| 18-99" function and dragging it to the corresponding row in the "Main Pooling Window" (i.e. the "Acute Respiratory Symptoms" endpoint group).

(v) Next we will add a valuation function for mortality. Under the "Mortality" endpoint group, expand the "Mortality, All Cause" endpoint. Add the value of a statistical life (VSL) function that ranges from $1-10 million with a normal

distribution by clicking and dragging the function "VSL, based on range $1 to $10 million, normal distribution, | 0-99" to the pooling window (Figure 6-12).
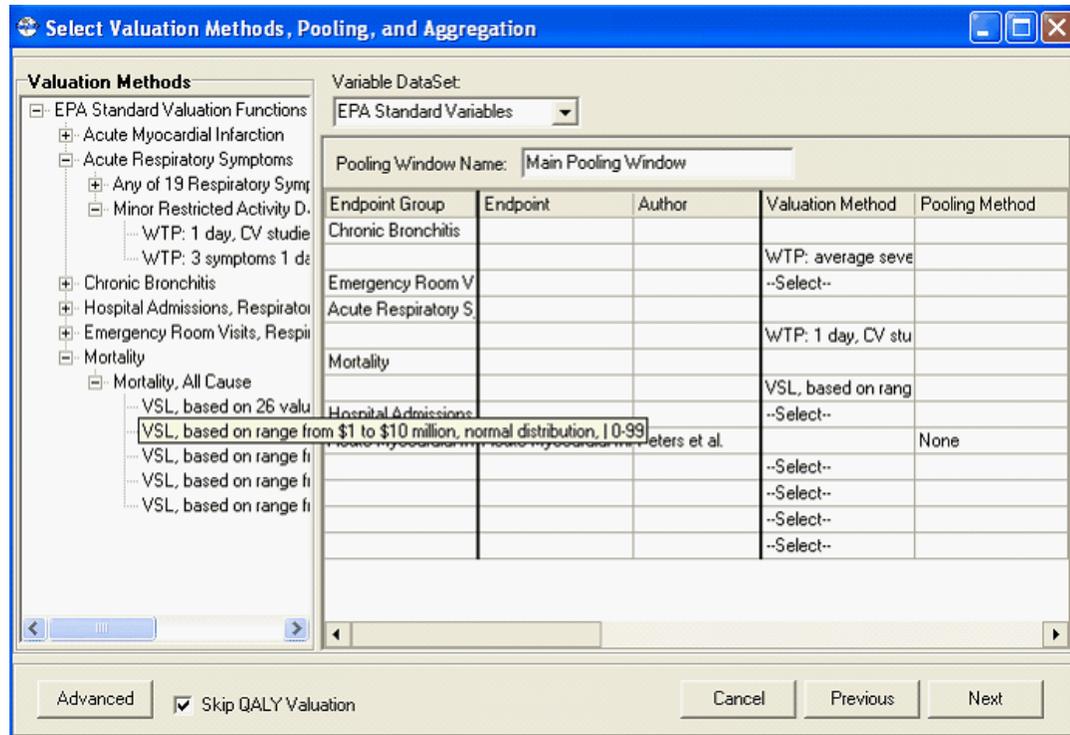
Figure 6-12. Initial three valuation functions in the pooling window

Background: Recall that in Section 6.2(m), we pooled the two underlying mortality results to make this single, combined result to which we have just assigned a valuation function. Because we pooled the health incidence results, we need only one valuation function to calculate the value of mortality. If we had not pooled the results, we would need at least one valuation function for each of the mortality results. The same holds true for the HA endpoint group (see next step).

(w) Under the "Hospital Admissions, Respiratory" endpoint group, expand the "HA, All Respiratory" endpoint. Add the valuation function for cost of illness (COI) medical costs and wage loss ages 65 to 99. Recall that in Section 6.2(n) we combined the individual HA chronic lung disease and pneumonia results to create a pooled result containing all respiratory results for ages 65 to 99.

(x)

> *Tip*: You can click **Previous** to return to the *Incidence Pooling and Aggregation* window if you want to inspect the underlying health incidence results or the incidence pooling. You will not lose your present valuation configuration. When you are done exploring the incidence results, click **Next** to return to the *Select Valuation Methods, Pooling, and Aggregation* window.

(y) For the rest of the endpoint groups, we will be adding multiple valuation functions per group. Expand the "Emergency Room Visits, Respiratory" endpoint group and then the "Emergency Room Visits, Asthma" endpoint. Add the COI Smith valuation function. Repeat for the COI Standford study (Figure 6-13).

Note: A "None" appeared in the *Pooling Method* column across from the endpoint group "Emergency Room Visits, Respiratory". Similar to the incidence pooling window, this indicates that we could potentially pool these two valuation functions. We will do this pooling in a later step.
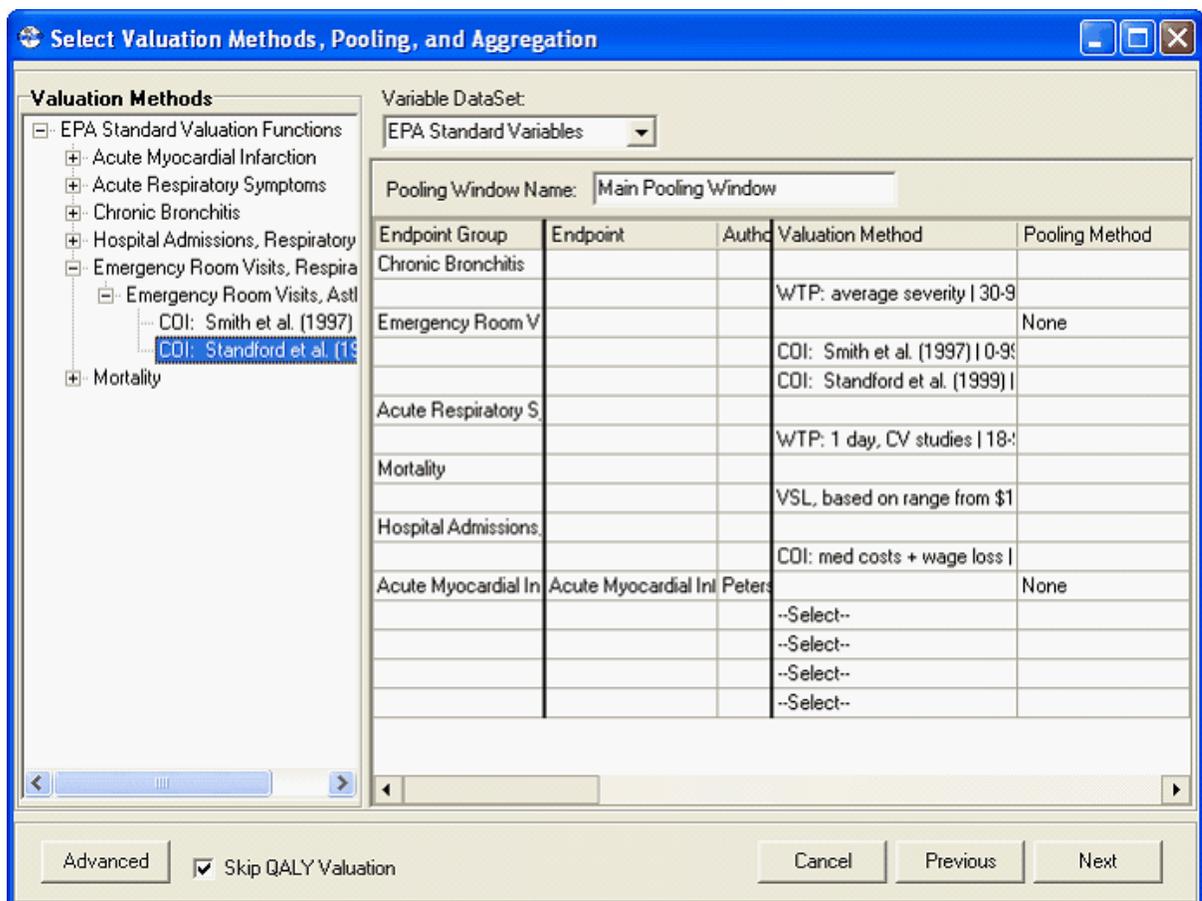
Figure 6-13. Adding two emergency room visits valuation functions to the pooling

window

(z) Next we will add valuation functions to the AMI health incidence results. We will add age-specific valuation functions to each of the age-specific health incidence results. To see the age-specific results, use the scroll bar for the pooling window to pan until you can see the *Start Age* column.

In the "Valuation Methods" panel, expand "Acute Myocardial Infarction". We will add two valuation functions per age range: one function by Russell and the other by Wittels. We will use the COI function for 5 years of medical costs and 5 years of wages loss with a 3% discount rate (DR). Add the valuation function "COI: 5 yrs med, 5 yrs wages, 3% DR, Russell (1998) | 25-44" to the AMI health incidence result for ages 18 - 44 (i.e., start age = 18).

Repeat for "COI: 5 yrs med, 5 yrs wages, 3% DR, Wittels (1990) | 25-44" (Figure 614).

Note: As happened with the endpoint group "Emergency Room Visits, Respiratory", "None" appeared in the *Pooling Method* column next to the 18 to 44 result. Again, this indicates that these two valuation functions could be pooled.
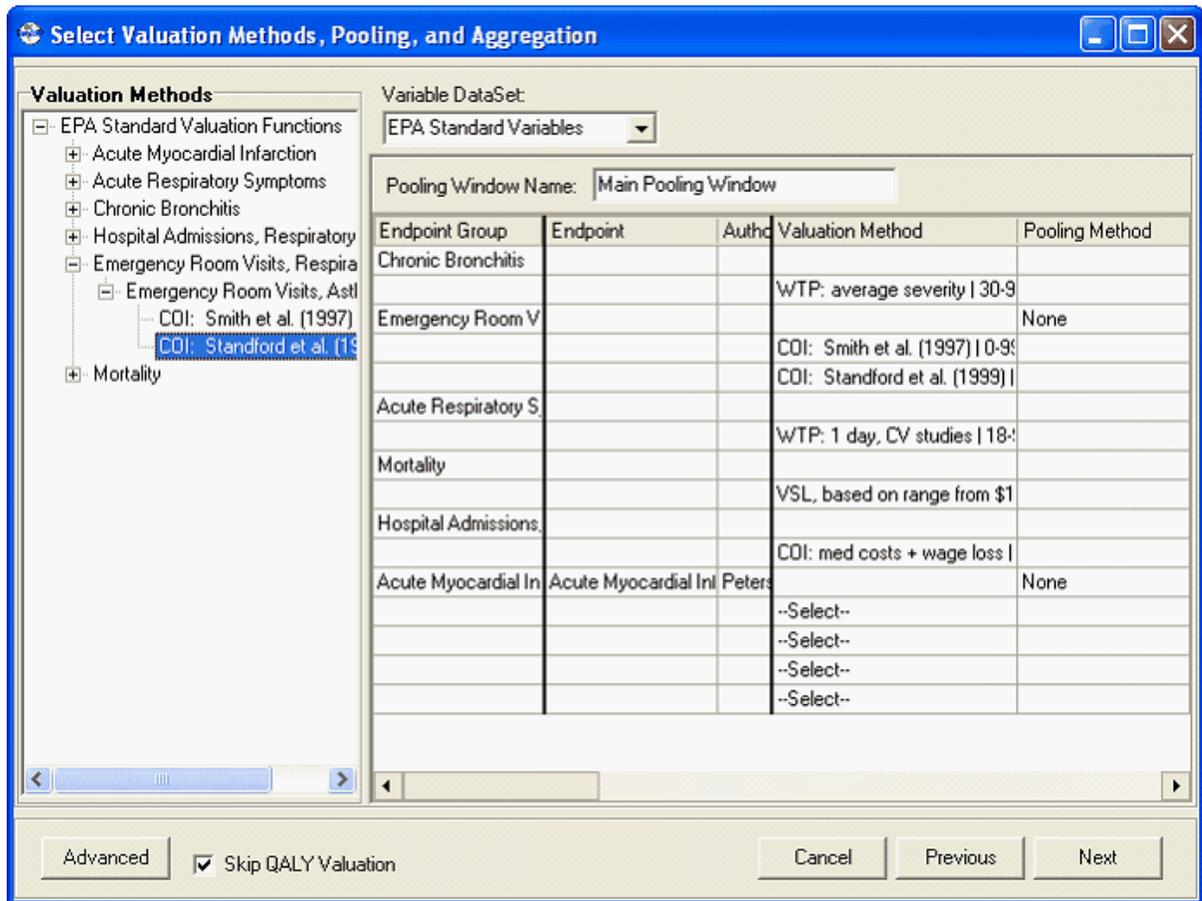
Figure 6-14. Adding two age specific AMI valuation functions to the pooling window

Tip: If you mistakenly add the wrong valuation function, you can delete it by highlighting the particular function in the right hand pooling window panel and clicking **Delete** on your keyboard.

(aa) Continue to add one Russell and one Wittels function for each of the remaining age ranges (i.e., 45-54, 55-64, and 65-99) (Figure 6-15).
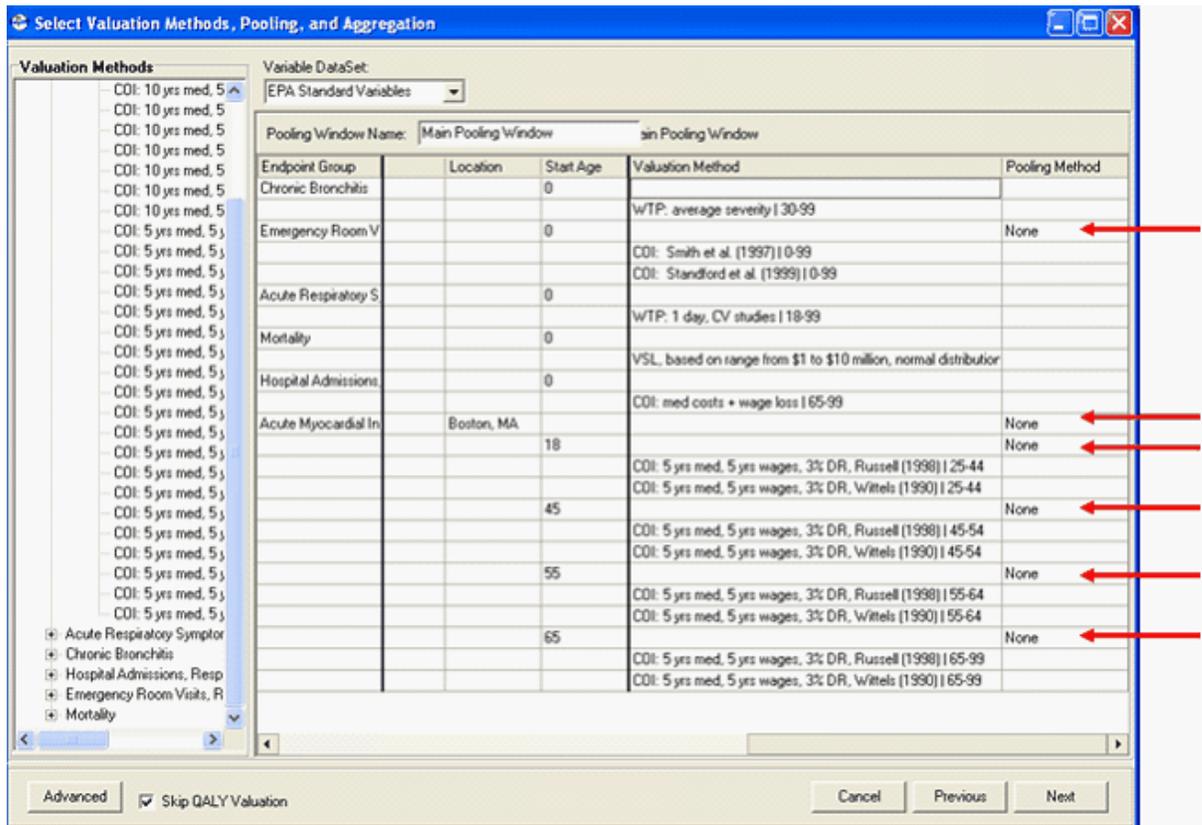
Figure 6-15. All valuation functions added to the pooling window

(ab) The "Main Pooling Window" now has six places we could potentially pool the results (indicated by the red arrows in Figure 6-15). The pooling procedure is the same as the one we used earlier in the *Incidence Pooling and Aggregation* window.

(ac) Starting with the "Emergency Room Visits, Respiratory" endpoint group, click on the "None" in the corresponding *Pooling Method* column. A drop-down menu will appear that lists all the possible pooling methods. Select the "Subjective Weights" method.

Background: You generally use subjective weights when you have overlapping populations (same endpoint and same age range). Unlike the random/fixed effects method, you explicitly determine the relative weights of the two (or more) distributions in calculating the combined distribution. For example, if you want to emphasize one study over another, you would give it a greater weight. We will assign these weights later, after all pooling methods have been selected.

(ad) For each of the AMI age ranges, we will pool together the Russell and Wittels age-specific valuation functions. Click on the "None" for the start age of 18 in the *Pooling Method* column (*make sure that you are **not** selecting the "None" in the row above this one, which is for the entire the endpoint group*), then select "Subjective Weights" (Figure 6-16).

Repeat this step for each of the other three age ranges, choosing "Subjective Weights" in every case.
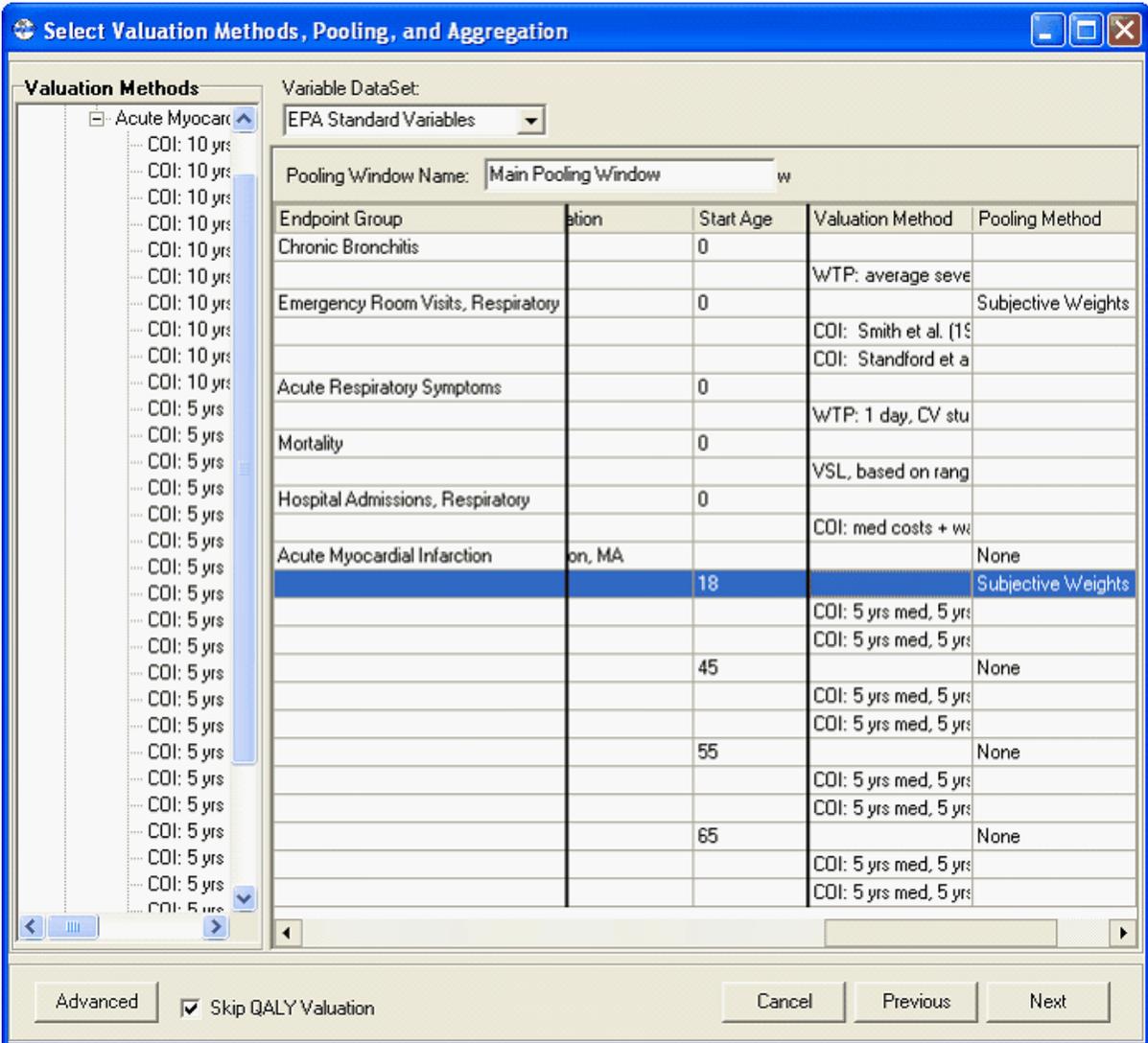
Figure 6-16. Adding "Subjective Weights" pooling to the specific AMI valuation functions

(ae) At this point we have combined the two valuations for each age range. Now we want to "move up a level" in the hierarchy and combine the four age ranges. Click the "None" in the same row as the AMI endpoint group, then select "Sum (Dependent)" (Figure 6-17). This will create one valuation result for the entire endpoint group.
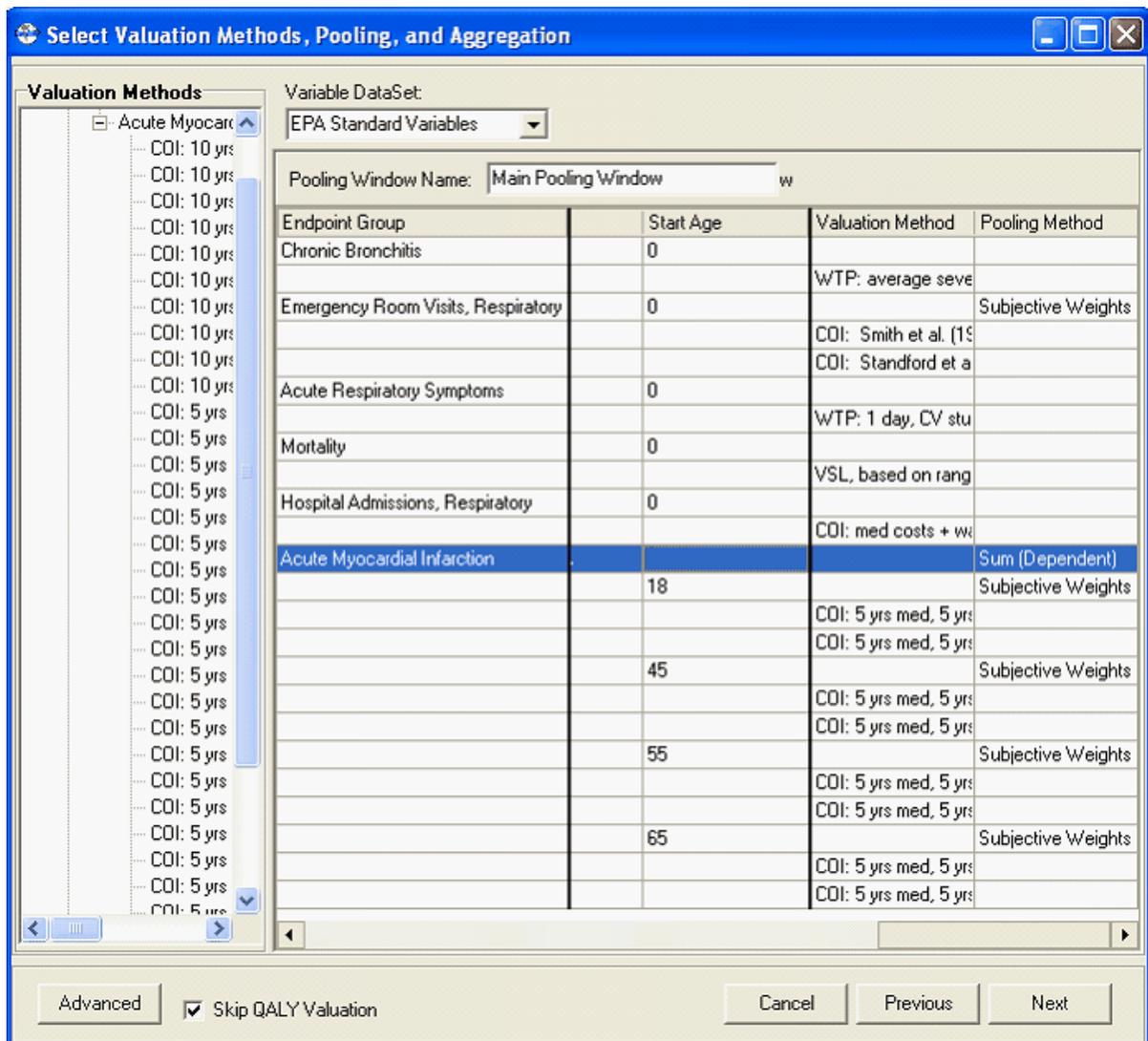
Figure 6-17. Adding "Sum (Dependent)" pooling to the AMI endpoint group

(af) Finally, we need to set the specific weights to be used in the subjective weights pooling methods. Double-click on any of the "Subjective Weight" pooling method cells. The *Select Subjective Weights* window will open (Figure 6-18).

The default weight for each of the component functions is 0.5, i.e., the valuation functions are being weighted equally. We will leave the default weights for AMI's four poolings.

For the emergency room (ER) pooling, we want to give more weight to the Smith valuation function. Edit the *Weights* column so that the Smith function is 0.60 and the Standford is 0.40. Click **OK** to set these weights and return to the *Select Valuation Methods, Pooling, and Aggregation* window.
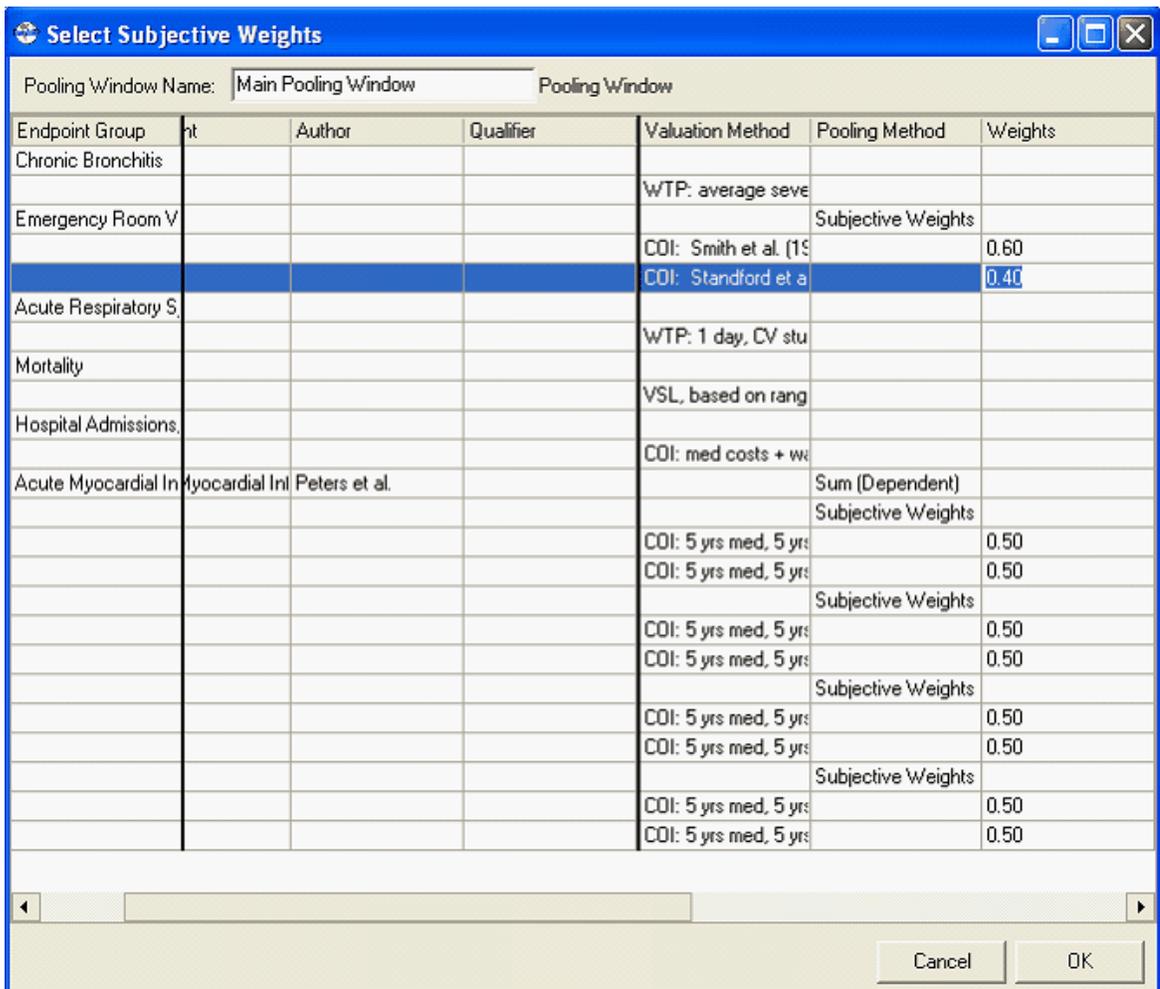
Figure 6-18. Changing the weights for "Subjective Weights" pooling of ER visits

(ag) **Exercise (6.2):** Why did we use "Sum (Dependent)" pooling for the AMI endpoint group instead of "Random/fixed Effects" pooling?

Answer:

(ah) At this point we have set up the pooling for health incidence results (Stage 1 in creating valuation functions), and have set up the specific functions and the pooling for similar valuations (Stage 2). Moving on to Stage 3, we need to set some additional parameters for the valuation functions, and decide on the aggregation levels for the health incidence results and the valuation results. We will then save all these settings from the three stages as a configuration (apv).

In the *Select Valuation Methods, Pooling, and Aggregation* window, click the **Advanced** button. This will open the *APV Configuration Advanced Settings* window (Figure 6-19).

Under the "Aggregation and Pooling" tab, use the drop-down menus to select "State" as the aggregation level for the incidence and valuation results. In other words, the results will be aggregated from CMAQ 36 km grid cells to states (minimum spatial unit).
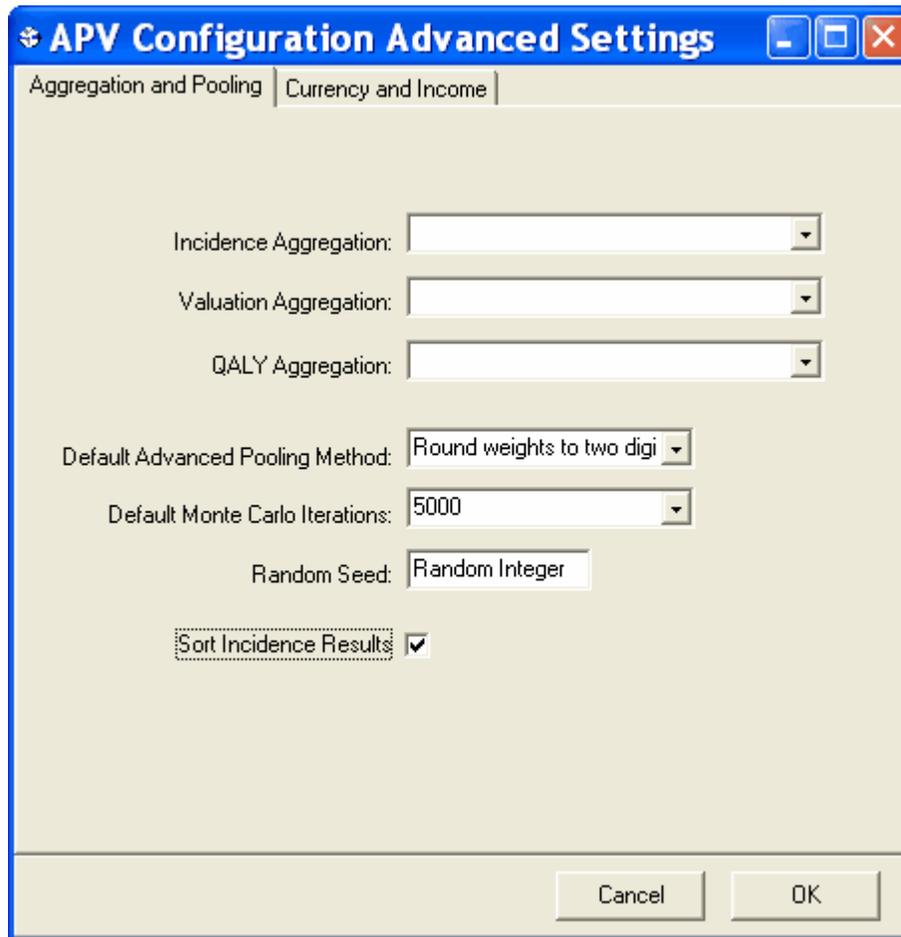
Figure 6-19. Advanced settings: Changing the aggregation level and the inflation dataset

(ai) Under the "Currency and Income" tab set the "Inflation DataSet" to "EPA Standard Inflators" and make sure the "Currency Year" is set to 2000. Then, set the "Income Growth Adjustment DataSet" to "Income Elasticity (3/21/2007). Change the "Year" to 2020, the same year used in our aqg files created earlier in the training. Then select all of the "Endpoint Groups" by clicking on the first group, holding down the *Shift* key and clicking on the last group (Figure 6-20).
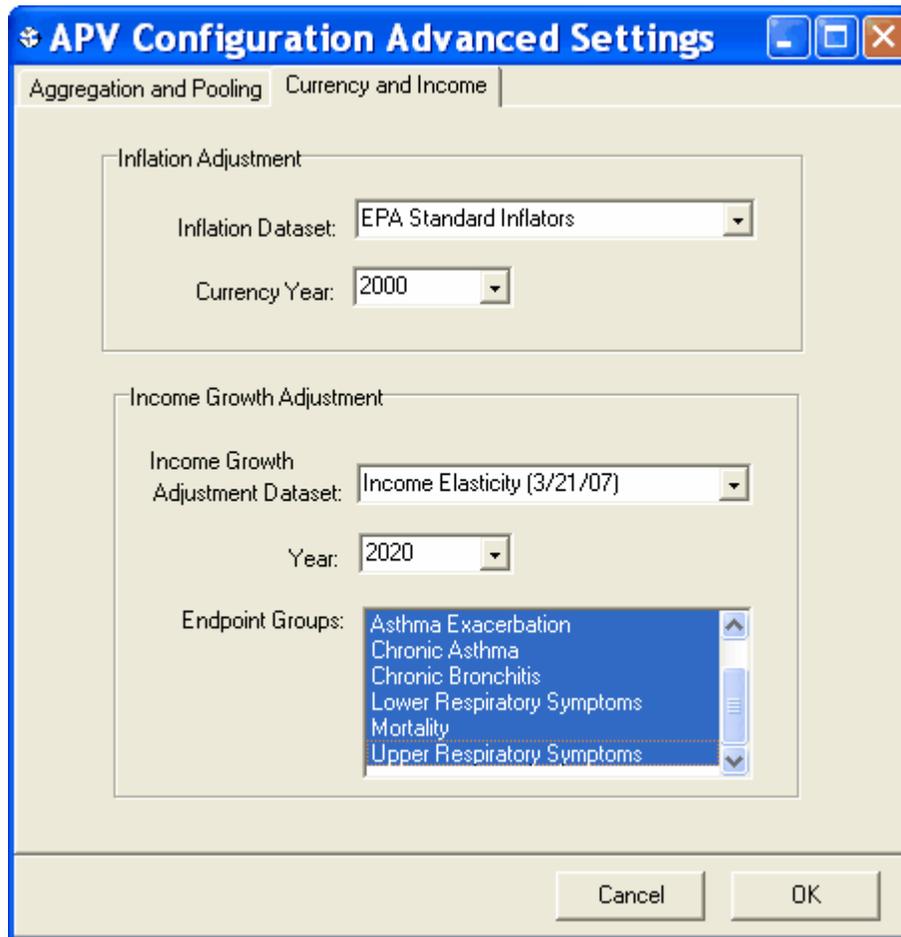
Figure 6-20. Advanced settings: Setting the income growth adjustment parameters

(aj) Since, as noted earlier, we are not doing anything in this training with QALY weights (the third tab in the window), click **OK**. This will return you to the *Select Valuation Methods, Pooling, and Aggregation* window.

The last parameter you need to set is the "Variable DataSet" at the top of the window. Use the drop-down menu to select "EPA Standard Variables" (Figure 6-21). *Note*: If you do not set the "Variable DataSet" you cannot save and run the configuration (next step).
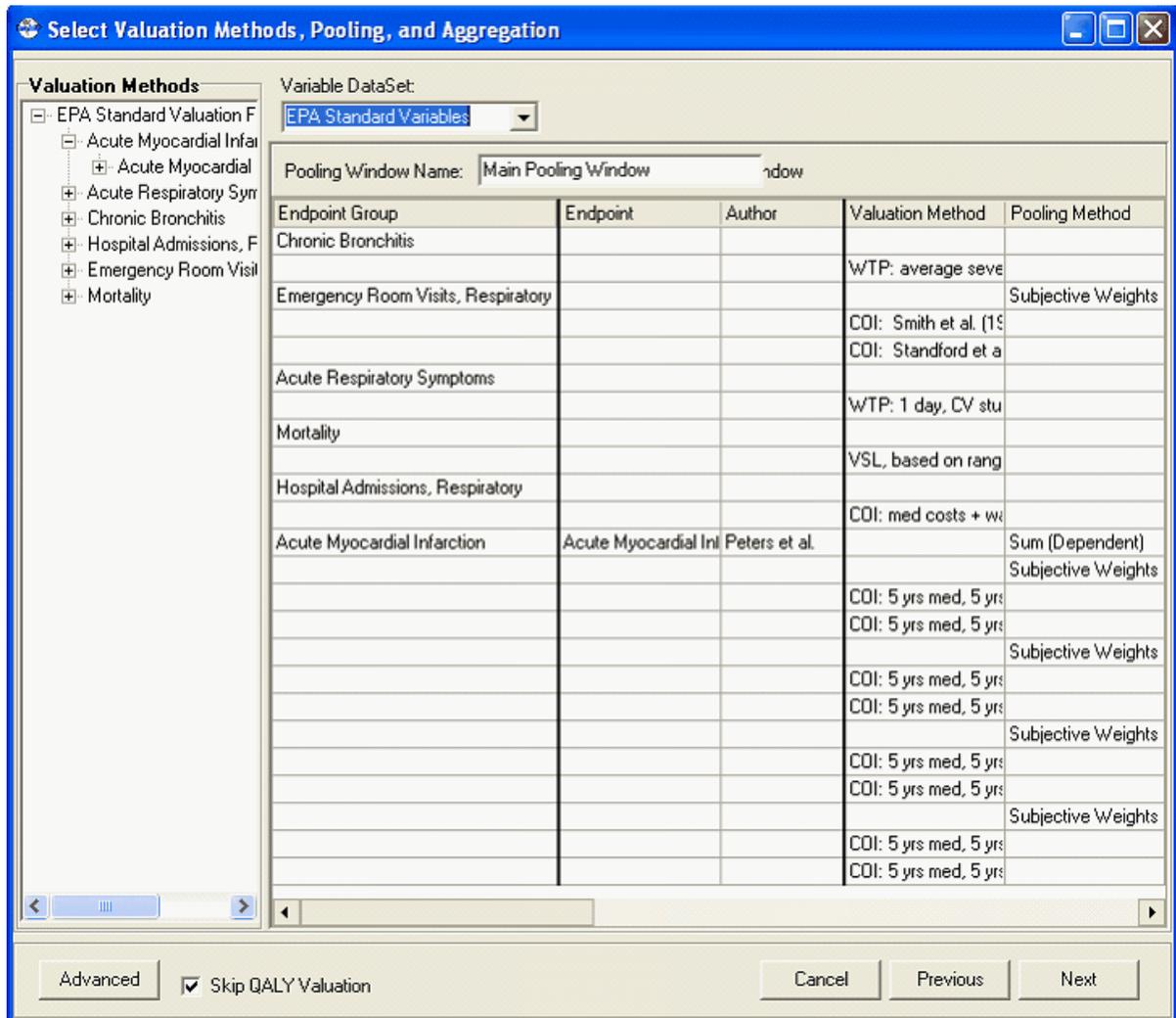
Figure 6-21. Setting the variable dataset

Background: The variable dataset includes variables that are used in calculating the valuation—for example, median income and average house size. Many of these variables are provided at the county level.

(ak) When you have set up all the valuation functions and pooling as instructed above, you are ready to move to Stage 4 of creating valuation results: saving the new configuration file (apv) and generating the valuation results file (apvr).

Click **Next**. This will bring up a *Save Aggregation, Pooling, and Valuation Configuration* window (Figure 6-22). We do want to save this new configuration (so that it can be re-used), so click **Save**.

This will bring up a *Save As* window. Under the "Configurations" folder, in the "File name" field, type in the new configuration file name, "PM25_RIA_2020_course_modified_state", and click **Save**.
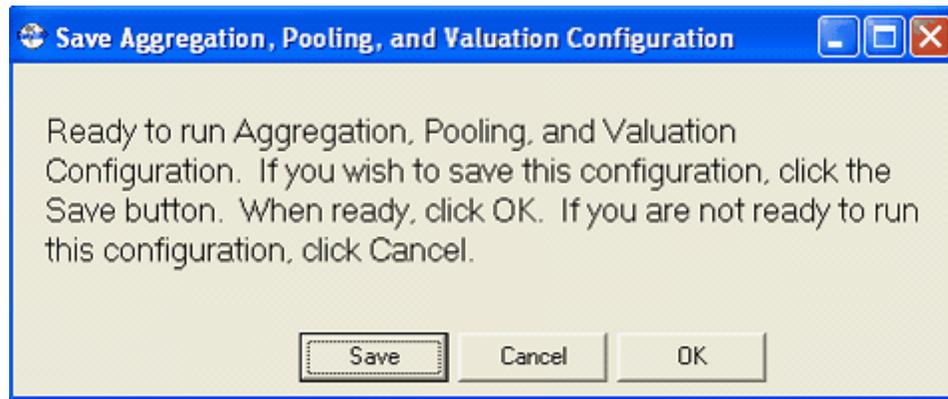
Figure 6-22. *Save Aggregation, Pooling, and Valuation Configuration* **window**

(al) After the APV configuration is saved, you will be returned to the *Save Aggregation, Pooling, and Valuation Configuration* window (Figure 6-22). Now run the configuration by clicking **OK**. This will bring up another *Save As* window. Here we will save the valuation results (apvr). Under the "Configuration Results" folder, save the results as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state".

The calculation of the results will begin and a *Progress* window will appear. The calculation of the results may take a few minutes. When the calculations are finished, you will be returned to the main BenMAP window.

Tip: If you are not ready to run this configuration, click **Cancel**. If you generated a configuration, at a later time you can open the configuration and run it to generate valuation results at a later time.

Background: The process of creating an aggregation, pooling, and valuation configuration (apv) is more restrictive than the process of creating a health incidence configuration (cfg) (discussed in Section 5). If you change the number or type of health incidence results used in an APV process, you have to re-create the health incidence pooling. If you change the health incidence pooling, you have to re-create the valuation functions and pooling.

Analysis:

The rest of the exercises in Section 6.2 focus on analyzing the results of our BenMAP run.

Specifically, we will look at the newly created aggregation, pooling, and valuation configuration file (apv) and results file (apvr). Recall from Section 5 that quality-assuring both your configuration and your results is a good idea.

(a) First, we will use the audit trail to look at the newly created configuration file (apv). From the main BenMAP window, click on the "Report" graphic in the right-hand panel. Select the "Audit Trail Reports" in the Select Report Type window and click OK. Under the "Configurations" folder, select the new configuration, "*PM25_RIA_2020_course_modified_state.apv*" and click **Open**.

(b) Exercise (6.3): What is the income growth adjustment year? For the incidence pooling, what are the pooling methods for the mortality and HA endpoint groups? What is the pooling method and advanced pooling method for the "HA, Chronic Lung Disease" endpoint? For the valuation pooling, what is the pooling method for the ER and AMI endpoint groups? What is the population year? When you are done with this exercise, click **OK** to close the Audit Trail Report window.

Answer:

(c) Using the "Tools" menu in the main BenMAP window, open a *BenMAP GIS* window. Open the "APV Configuration Results (*.apvr)" and select the "Pooled Incidence Results". Under the "Configuration Results" folder, open our newly created file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state.apvr".

(d) In the *Edit GIS Field Names* window, provide more meaningful names for your health incidence results then click **OK**. Here are some suggested names: ChronBronc, ER, AcutResp, Mortality, HA, AMI18, AMI45, AMI55, AMI65.

Note: As you would expect, the individual results have been pooled together (for example, the Pope and Woodruff results have been pooled into one mortality result).

(e) **Exercise (6.4):** What are the numbers of avoided mortalities in California, Pennsylvania, and Illinois? How many acute respiratory symptoms were avoided in the same states?

Answer:

(f) **Exercise (6.5):** Compare the AMI (heart attacks) for various age groups. What are the maximum values for the AMI18 and AMI65 functions?

Answer:

(g) Next we will overlay a pooled valuation map for the same apvr. From the same GIS window, open "APV Configuration Results (*.apvr)" and select "Pooled Valuation Results". Under the "Configuration Results" folder, open our newly created file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state.apvr ".

(h) In the *Edit GIS Field Names* window, provide more meaningful names for your valuation results. Some suggested names are ChronBronc, ER, AcutResp, Mortality, HA, AMI.

(i) In the *Valuation Sums Layer* window, add a sum for mortality and morbidity. For morbidity, check the following endpoint groups: chronic bronchitis, ER, acute respiratory symptoms, HA, AMI. Use "Dependent" as the summation type.

Remember to edit the "GIS Field Name" to "morbidity" and "mortality",

respectively. Make the "Control_PM25_RIA…(Pooled Valuation Results 0)" layer the active layer; simply right-click on that layer and select "Move Up".

(j) **Exercise (6.6):** What are the monetized benefits for the avoided acute respiratory symptoms in California, Pennsylvania, and Illinois? Now make the "Pooled Valuation Results Sums" layer active. What are the monetized benefits for the avoided morbidity events in the same states? What are the monetized benefits for the avoided mortalities in the same states? In comparing your answers to those for Exercise 6.5, what conclusion can you draw about the mortality valuation function versus the acute respiratory symptoms valuation function? When you are done with this exercise, close the GIS window.

```
Answer:



```

(k) Create a new report for our valuation results. From the main BenMAP window, click on "Report" in the right-hand panel and then select the "Incidence and Valuation Results: Raw; Aggregated and Pooled" type. Under the "Configuration Results" folder, open our newly created file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state.apvr ". For the result type, select "Pooled Valuation Results".

In the *APV Configuration Results Report* window (Figure 6-23), select the "Endpoint Group" within the "Pooled Valuation Method Fields" panel. In the "Results Fields" panel, uncheck the "Variance". In the "Display Options" panel, reduce the "Digits After Decimal Point" to 0. The "Preview" panel in the bottom half of this window will reflect your choices.
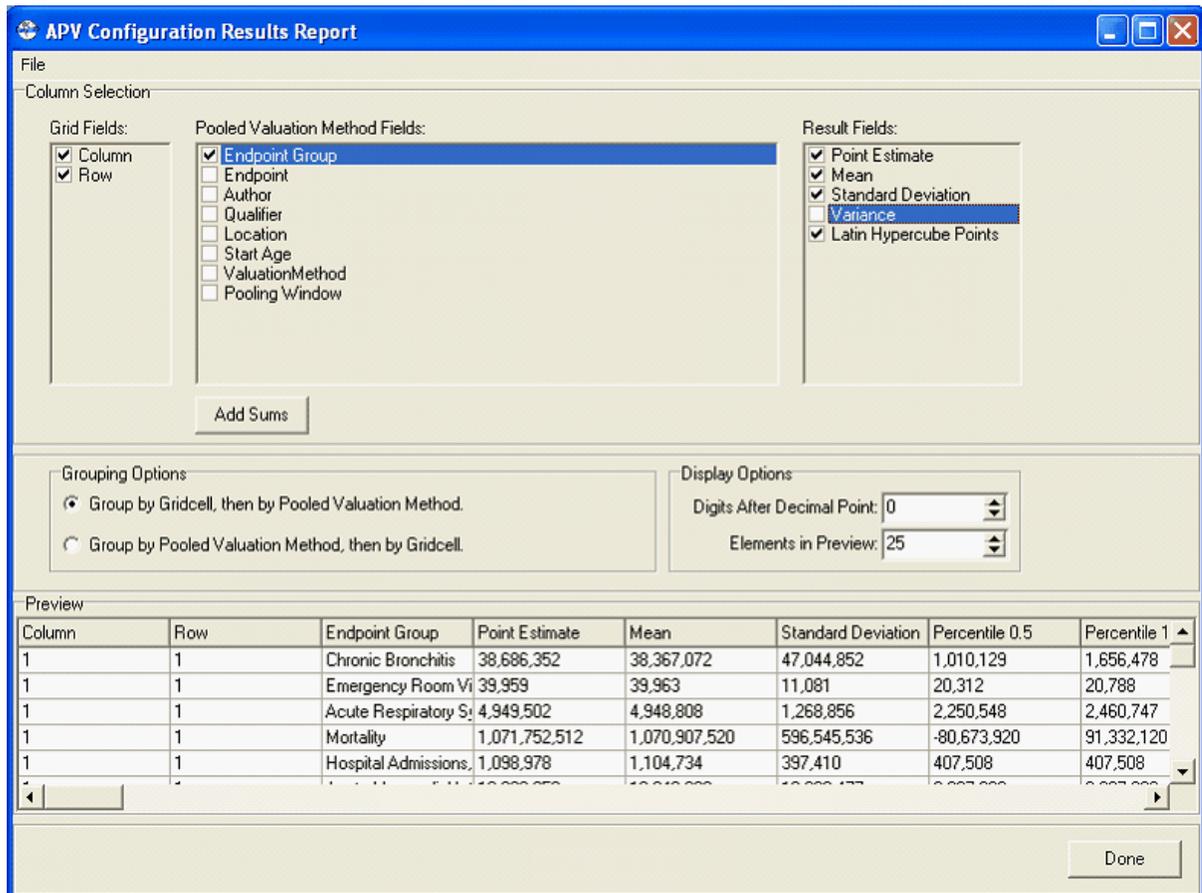
Figure 6-23. Pooled valuation report.

(l)   Now that you have chosen all the columns to include in your valuation results report, you can save the report. From the *APV Configuration Results Reports* window, type **Ctrl-S**. Under the "Reports" folder, save the file as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state_value".

After the file has been saved, close the window by clicking **Done**. Use Windows Explorer to navigate to the "Reports" folder. Double-click on the newly created csv file and it should open in Excel.

(m) **Exercise (6.7):** All of the following questions refer to California, FIPS code 6 (i.e., column = 6, row = 1). What is the point estimate of the monetized benefit for the number of avoided AMIs? What is the estimate of the monetized benefit for AMIs at the 0.5 percentile? What is the estimate at the 99.5 percentile? When you are finished with this exercise, close the Excel window.

Answer:

A.1.6.2 Example: PM2.5 Control 2020 14/35 Adjusted

The goal of this exercise is to re-use our newly created aggregation, pooling, and valuation configuration (apv) (Section 6.2) and produce valuation results for the adjusted control scenario RIA 2020 14 annual, 35 daily µg/m3.

Procedures:

(a) In the main BenMAP window, click on the graphic titled "Pooling, Aggregation, and Valuation" in the right-hand panel. This will open the APV Configuration Creation Method window

(a) In this window, select "Open Existing Configuration file for Aggregation, Pooling, and Valuation". Under the "Configurations" folder, select the newly created configuration "PM25_RIA_2020_course_modified_state.apv" and click **Open**.

(a) In the *Incidence Pooling and Aggregation* window that opens, we will change the input health incidence file (cfgr). Click the **Browse** button next to the "Configuration Results File Name(s)" field. Under the "Configuration Results" folder, select the health incidence file "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_adjusted.cfgr".

(a) Review your health incidence pooling window, "Main Pooling Window". The pooling configuration should be identical to our previous example (see Figure 6-8).

(a) Click **Next**. The *Select Valuation Methods, Pooling, and Aggregation* window will appear. Review your valuation functions and pooling. The configuration should be identical to our previous example (see Figure 6-17). Click on **Advanced** to confirm that these settings are the same as before (see Figure 6-19 and Figure 6-20). When you are done reviewing the configuration, click **Next**.

(a)**Do *not* save the configuration**. There is no need to save since the only change was to the health incidence results file (cfgr). Do create the valuation results (apvr) by clicking **OK**. Under the "Configuration Results" folder, save the results as "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state_adjusted" .

The calculation of the results will begin and a *Progress* window will appear. The calculation may take a few minutes.

Analysis: Now we will look at the newly created aggregation, pooling, and valuation results file (apvr).

(a)Open an audit trail for the new valuation results file, "Control_PM25_RIA_2020_modified_cmaq_grid_14_annual_35_daily_state_adjusted.apvr".

(a)**Exercise (6.8):** What is the population year used in this apvr? What are the weights used for pooling the two ER valuation functions?

Answer:

(a)Open a *BenMAP GIS* window, then open the same valuation results file (apvr). Map the "Pooled Incidence Results".

(a)**Exercise (6.9):** What are the numbers of avoided mortalities in California, Pennsylvania, and Illinois? How many acute respiratory symptoms were avoided in the same states? How do these results compare to the nonadjusted results (see Exercise 6.4)?

> Answer:
>
>

(a) Overlay the "Pooled Valuation Results" for the same apvr. Edit the GIS field names. In the *Valuation Sums Layer* window, create sums for morbidity and mortality the same way you did above Exercise 6.6. Make the "Control_PM25_RIA…(Pooled Valuation Results 0)" layer the active layer.

(a) **Exercise (6.10):** What are the monetized benefits for the avoided acute respiratory symptoms in California, Pennsylvania, and Illinois? Now make the "Pooled Valuation Results Sums" layer active. What are the monetized benefits for the avoided morbidity events in the same states? What are the monetized benefits for the avoided mortalities in the same states? How do these results compare to the nonadjusted results (see Exercise 6.6)? When you are done with this exercise, close the GIS window.

> Answer:
>
>

(a) Create a new report for the valuation results file. After selecting the report type and opening the file, you should select the results type, "Pooled Valuation Results".

> In the *APV Configuration Results Report* window, select the "Endpoint Group" within the "Pooled Valuation Method Fields" panel. In the "Results Fields" panel, uncheck the "Variance". In the "Display Options" panel, reduce the "Digits After Decimal Point" to 0.

> Under the "Reports" folder, save the file as "Control_PM25_RIA_2020_modified_

cmaq_grid_14_annual_35_daily_state_adjusted_value". Open the csv file in Excel.

(a) **Exercise (6.11):** All of the following questions refer to California, FIPS code 6 (i.e.,. column = 6, row = 1). What is the point estimate of the monetized benefit for the number of avoided AMIs? What is the estimate of the monetized benefit for AMIs at the 0.5 percentile? What is the estimate at the 99.5 percentile? How do these results compare to the nonadjusted results (see Exercise 6.7)? When you are finished with this exercise, close the Excel window.

Answer:

A.1.6.3 Example: Modifying One-Step Analysis Parameters

The goal of this exercise is to modify the One-Step Analysis to use our newly created health incidence and aggregation, pooling, and valuation configurations.

Procedures:

(a) From the main BenMAP window, select "One Step Analysis" from the "Parameters" menu (Figure 6-24).
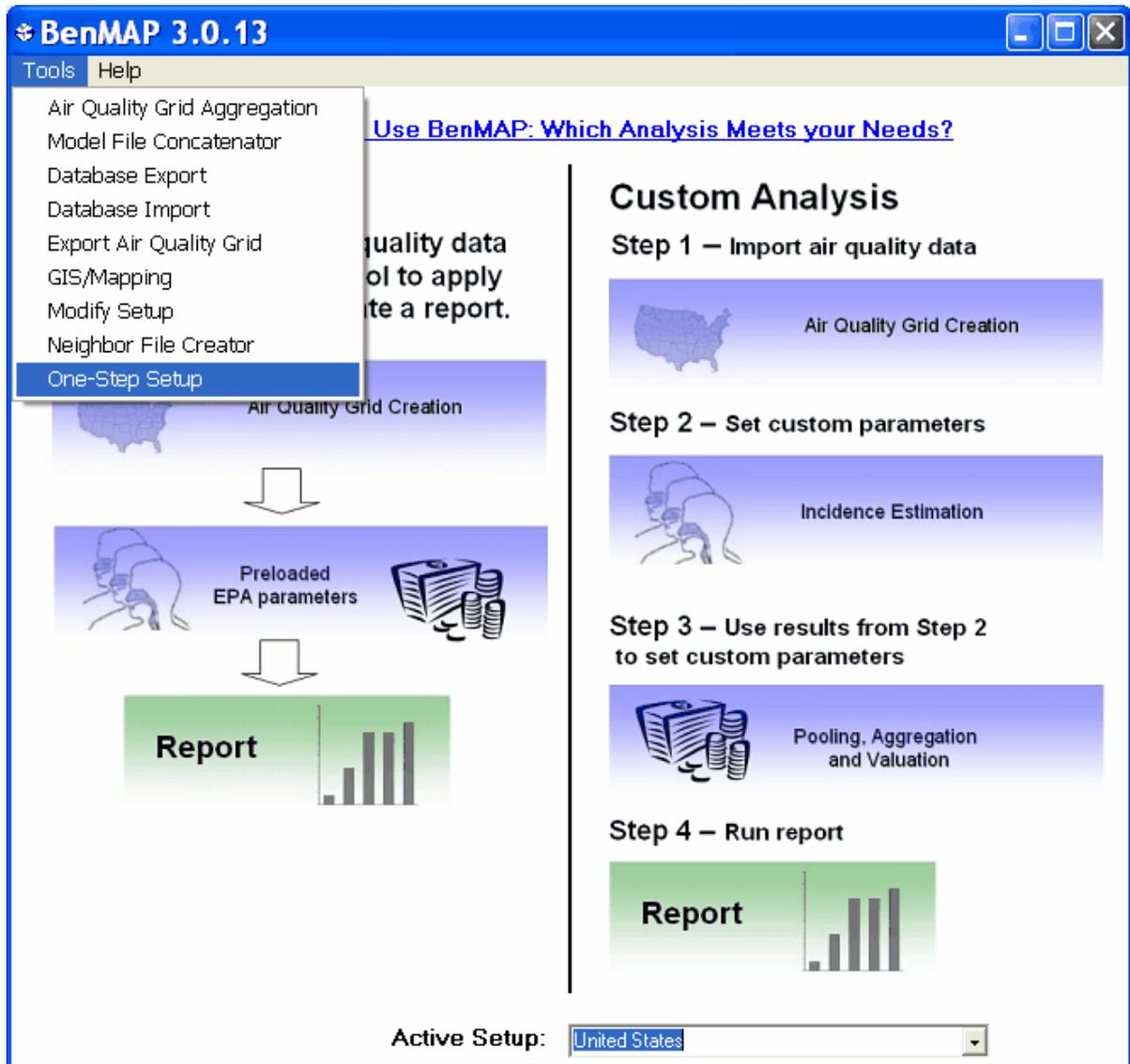
Figure 6-24. Parameters menu in the main BenMAP window.

(a) This will open the *One Step Setup Parameters* window (Figure 6-25). Here you can select the health incidence (cfg) and valuation configuration (apv) files that will be used in the One-Step Analysis approach.
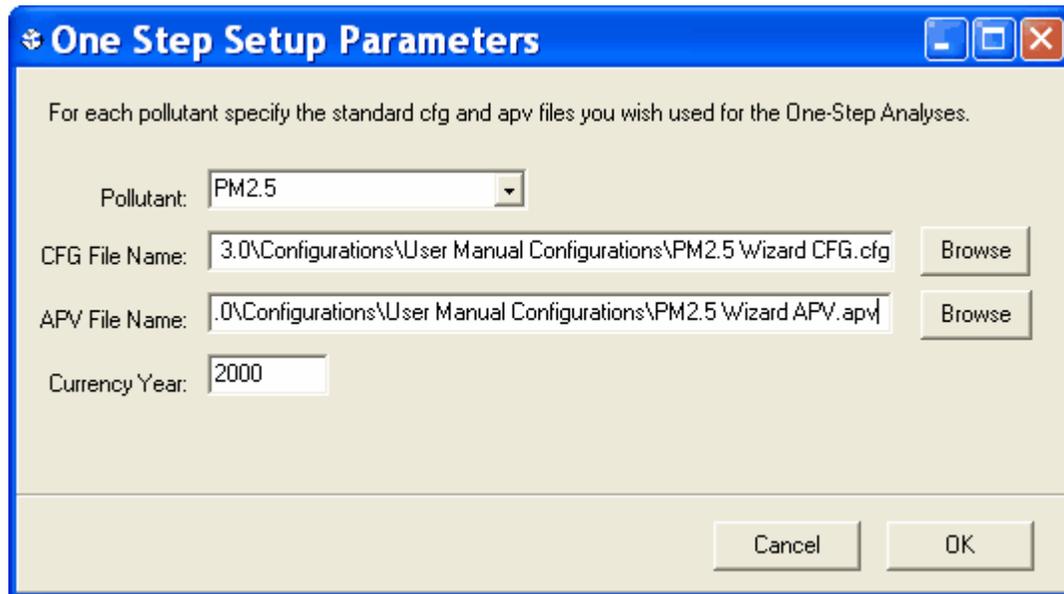
Figure 6-25. Modifying configurations for One-Step Analysis

(a) Select "PM 2.5" from the Pollutant drop-down menu. Click the **Browse** button to the right of the CFG File Name box. In the *Open* window, under the "Configurations" folder, select "PM25_RIA_2020_course_modified.cfg" and click **Open**. This will return you to *One Step Setup Parameters* window.

> Click the **Browse** button to the right of the APV File Name box, so that you can select the new aggregation, pooling, and valuation configuration. In the *Open* window, under the "Configurations" folder, select "PM25_RIA_2020_course_modified_state.apv" and click **Open**. This will again return us to *One Step Setup Parameters* window.

> We have the option of changing the currency year, but in our case leave the "Currency Year" as 2000. Finally, click **Save**. The next time you run One-Step Analysis, it will use these new configurations.

(a) This completes the "Aggregation, Pooling, and Valuation" lab. In the next lab, "Adding New Datasets & Independent Study", we will add the necessary data to do a metropolitan scale analysis of Detroit. After adding the new datasets, we will use our new configurations (from Sections 5 and 6) to create health incidence and valuation results for our new domain.

A.1.7   Section 7. Adding New Datasets & Independent Study

In this section, you will add new datasets to BenMAP and run a local-scale benefit analysis

in Detroit, Michigan.

BenMAP contains pre-loaded data necessary to perform a health impact assessment that will meet most users' analytical needs. However, you can also import your own datasets into BenMAP when the pre-loaded datasets are not adequate for your analysis. For example, you can add new population datasets, new grid definitions, new health impact and valuation functions, and new background incidence rates for specific health endpoints. If you decide to conduct either a local-scale analysis or a non-U.S. analysis, you will likely need to add new datasets to model the benefits and adequately reflect those local factors. In other words, the U.S. national datasets may not be the best available data or functions for your study.

In this section, you will add new datasets to conduct an entire benefit analysis for a change in air quality in the Detroit metropolitan area. Below are a few key aspects of the analysis:

Our model area is the greater Detroit metropolitan area, partially covering three counties: Wayne, Oakland, and Macomb.

The air quality grid cells are 1 km by 1 km.

The air quality model data are for 2020.

The control scenario models a 14.5 µg/m3 annual PM2.5 standard for the core of the metropolitan region.

Unlike our previous national studies, this lesson uses a finer-resolution grid, Detroit-specific population data, and Detroit-specific background incidence rates for hospital admissions due to asthma.

A.1.7.1 Example: Adding Datasets for Detroit

The goal of this exercise is to add the necessary datasets so that we can conduct a local-scale analysis in Detroit. Specifically, we will add a new 1-km grid definition, a new population dataset for Detroit, and new background incidence for asthma hospital admissions for Detroit.

Procedures:

(a)From the main BenMAP window, select "Modify Setup" from the Tools menu (Figure 7-1).
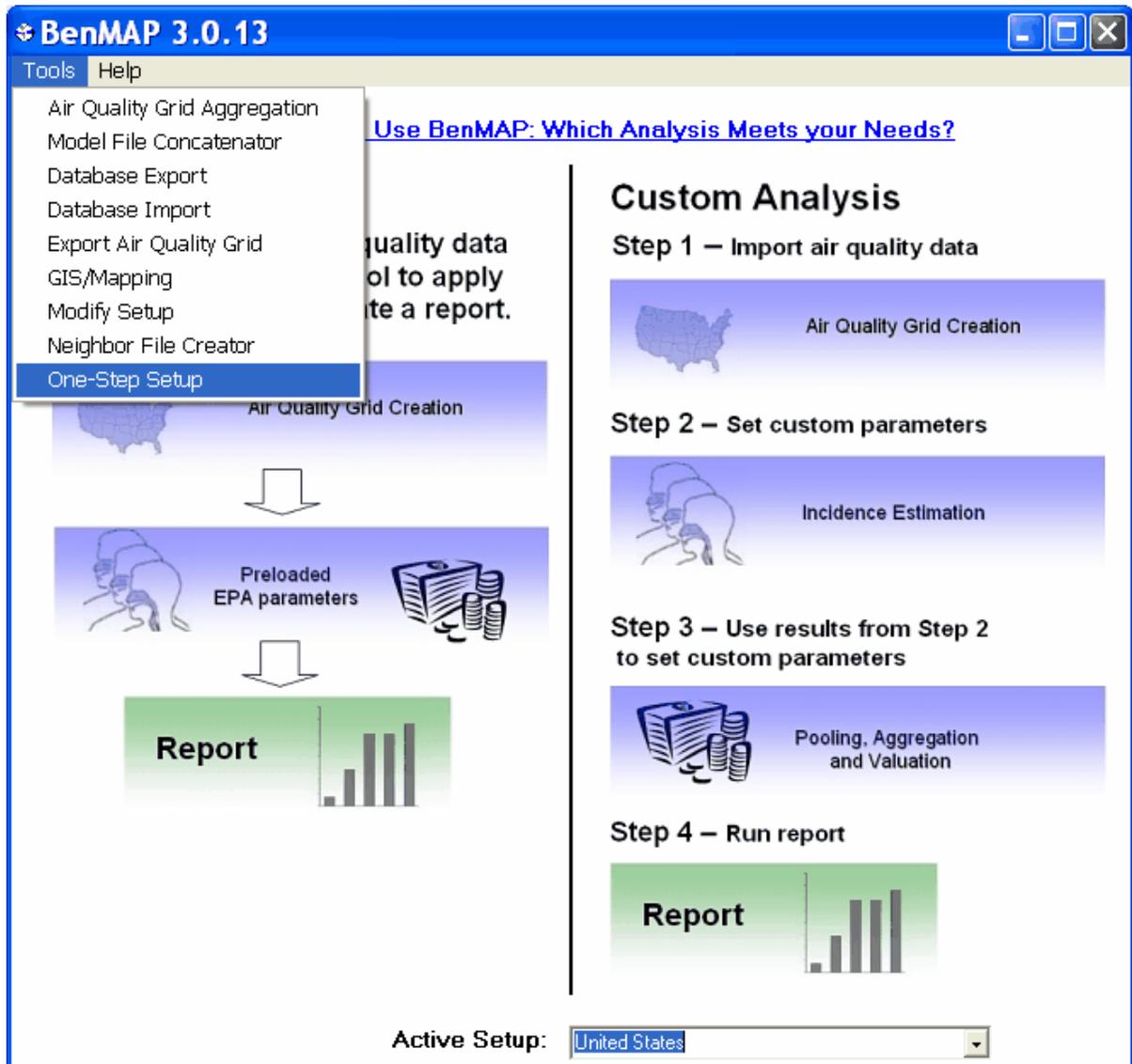
Figure 7-1. Choosing "Modify Setup"

(b)This will open the *Manage Setup* window (Figure 7-2). Through this window, you can modify many of the datasets, parameters, and functions used in BenMAP. Specifically, you can add to or change any of the following 11 categories of data: grid definitions, pollutants, monitor datasets, incidence/prevalence datasets, population datasets, C-R function datasets, variable datasets, inflation datasets, valuation datasets, income growth adjustments, and QALY distribution datasets. To modify any one of these, you can click the **Edit** button below the appropriate list.
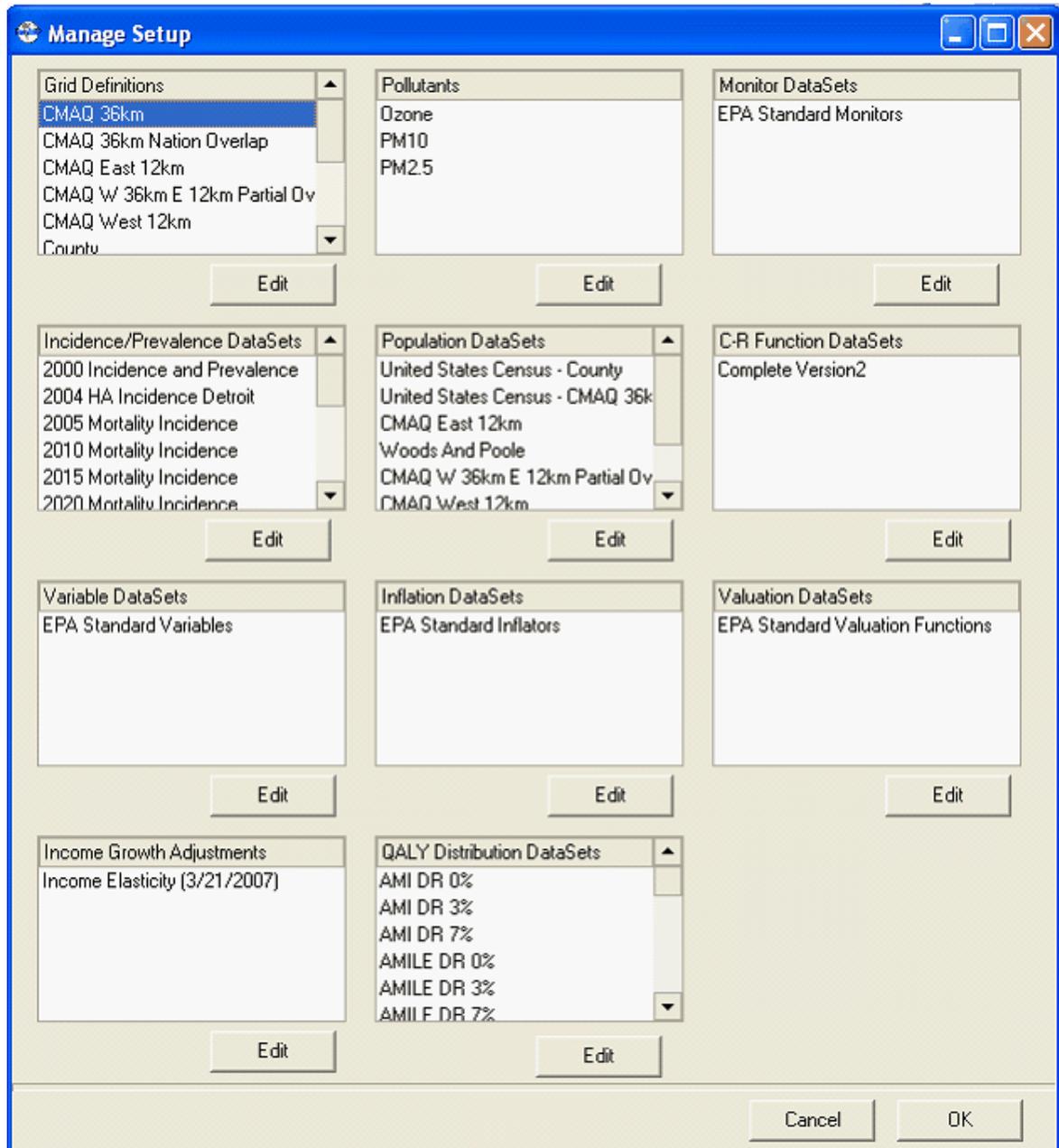
Figure 7-2. *Manage Setup* **window**

(c) We will start by adding a new grid definition. Click **Edit** below the *Grid Definitions* list. This will open the *Manage Grid Definitions* window (Figure 7-3). Here you can add new grid definitions or you can delete or edit existing grid definitions.
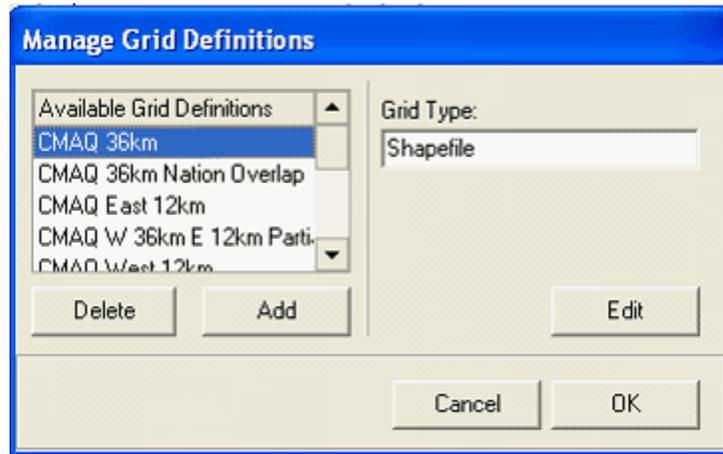
Figure 7-3. *Manage Grid Definitions* **window**

Tip: You can use the **Edit** button to view other grid definitions. Simply highlight the particular grid in the *Available Grid Definitions* list and click **Edit**. *Make sure you do not accidentally save any changes while you are viewing the definition.* You can use this same technique to view other datasets that have already been loaded (e.g., "Pollutants").

(d) We will add a new grid definition for Detroit. In the *Manage Grid Definitions* window, click the **Add** button. This will open a *Grid Definition* window (Figure 7-4). Here we can define our new grid.

> *Note*: When adding a series of new datasets, you should generally load the new grid definition first. Adding other datasets (e.g. adding population data) will typically depend on the new grid definition.
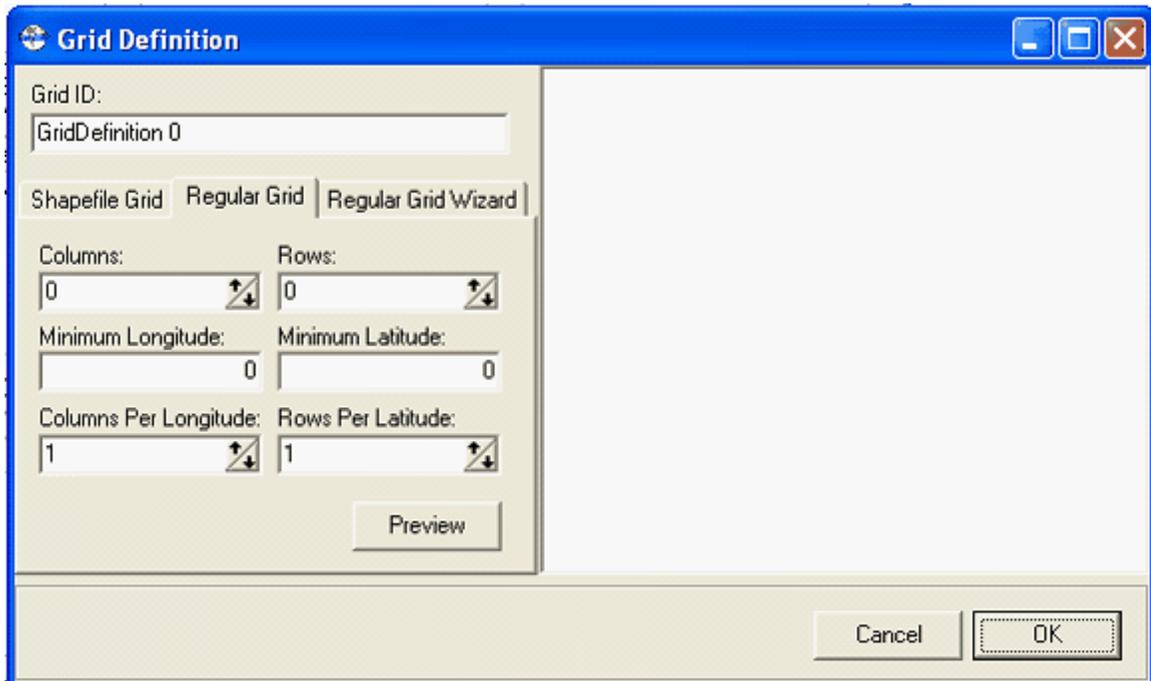
Figure 7-4. *Grid Definition* **window**

(e) We will define the new grid based on an ESRI shapefile. First, we will set the name of the new grid. Edit the *Grid ID* field by changing the text to "Detroit CMAQ 1km".

Next we load our shapefile. Click the "Shapefile Grid" tab. Next to the *Load Shapefile* field, click the browse button. Under the folder "Inputs" and the subfolder "Detroit", select the file "Detroit_grid.shp" and click **Open**. This will return you to the *Grid Definition* window.

Click **Preview** to view your new grid (Figure 7-5).
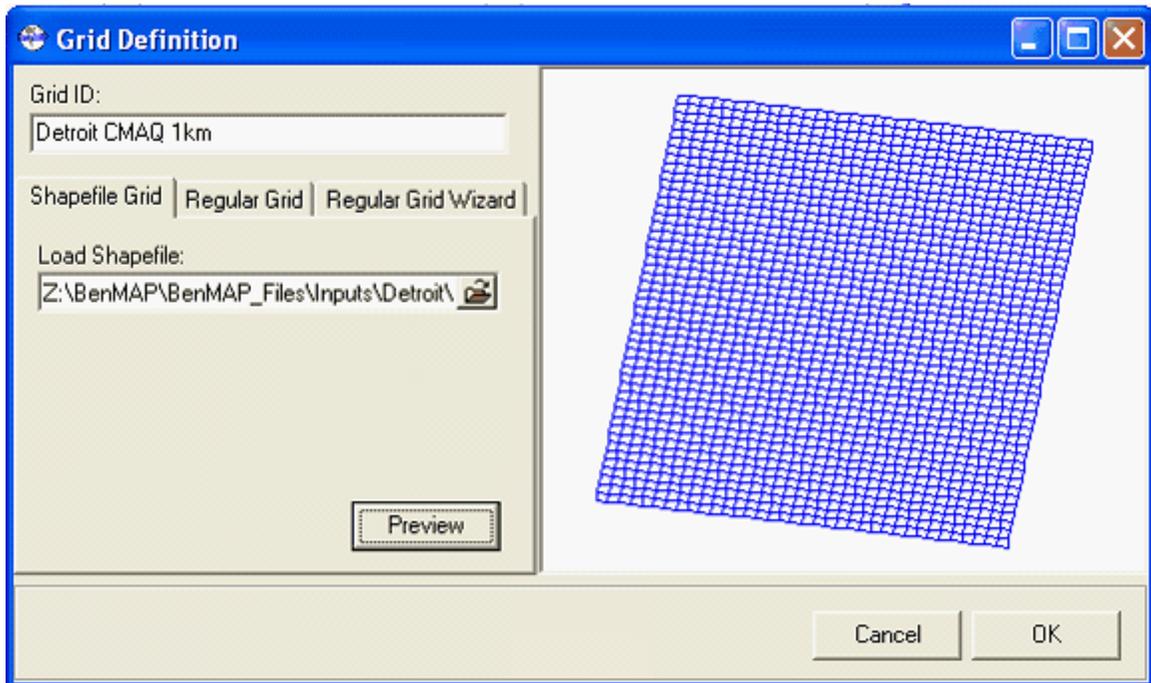
Figure 7-5. Loading and previewing the Detroit CMAQ 1km grid

(f) Click **OK**. This will return you to the *Manage Grid Definitions* window. Scroll through the list of available grid definitions to confirm that your new grid is included then click **OK**. This will return you to the *Manage Setup* window. *Note*: It may take a minute or two to complete the loading of the new grid (indicated by the *Manage Setup* window becoming active again).

Confirm that the new definition is in the *Manage Setup* window's grid definitions list. If it is, then you have successfully added a new grid definition.

(g) Next we will add a new background incidence rates dataset. This new dataset will have specific rates for Detroit instead of national averages. However, for the purposes of this lesson, the dataset will include only the "HA, Asthma" endpoint. Click **Edit** below the *Incidence/Prevalence DataSets* list. This will open the *Manage Incidence DataSets* window (Figure 7-6). Here you can add new background incidence rates datasets or you can delete or edit existing background incidence rates datasets.
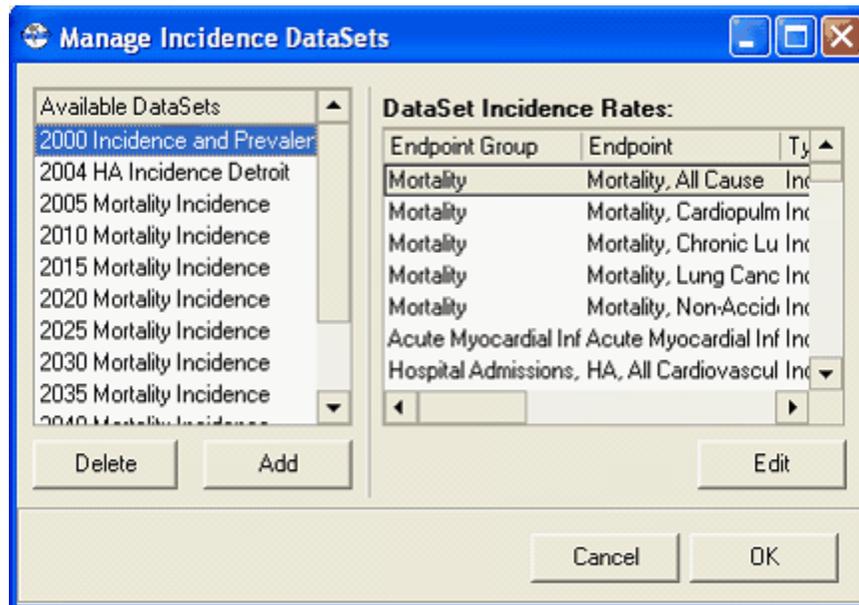
Figure 7-6. *Manage Incidence DataSets* **window**

(h)Click the **Add** button. This will open an *Incidence DataSet Definition* window (Figure 7-7) that we will use to define our new background incidence rates dataset.
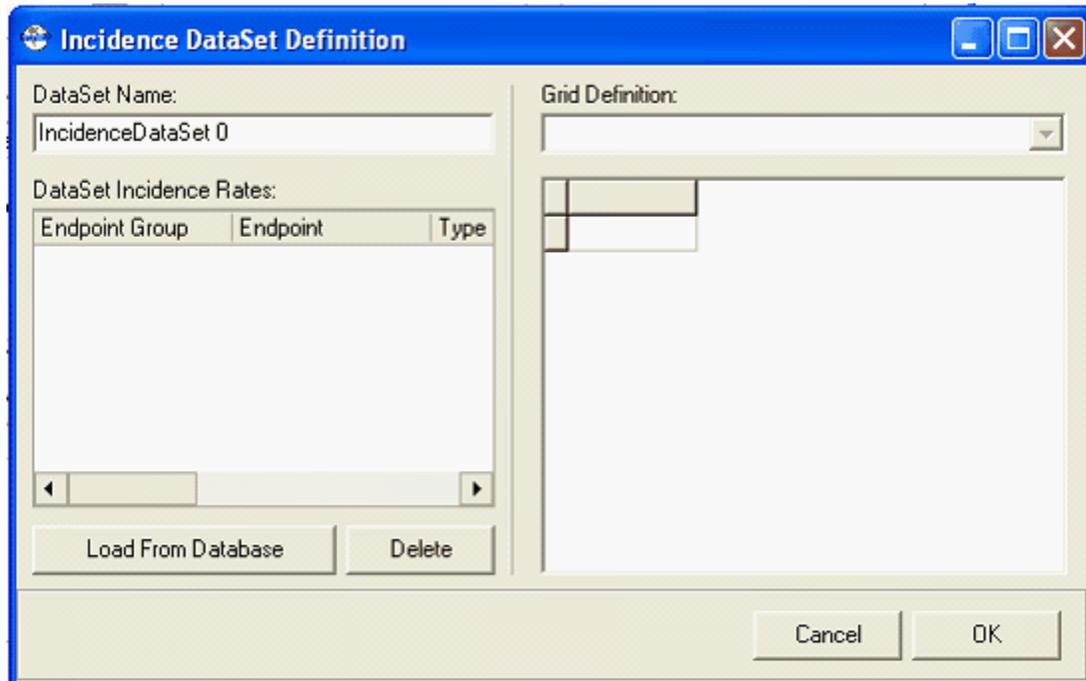
Figure 7-7. *Incidence DataSet Definition* **window**

(i) Edit the *DataSet Name* field by changing the text to "2004 HA Incidence Detroit". Click the **Load From Database** button. This will open a *Load Incidence/Prevalence DataSets* window. Using the *Grid Definition* drop-down menu, select the "Detroit CMAQ 1km" grid definition (Figure 7-8).
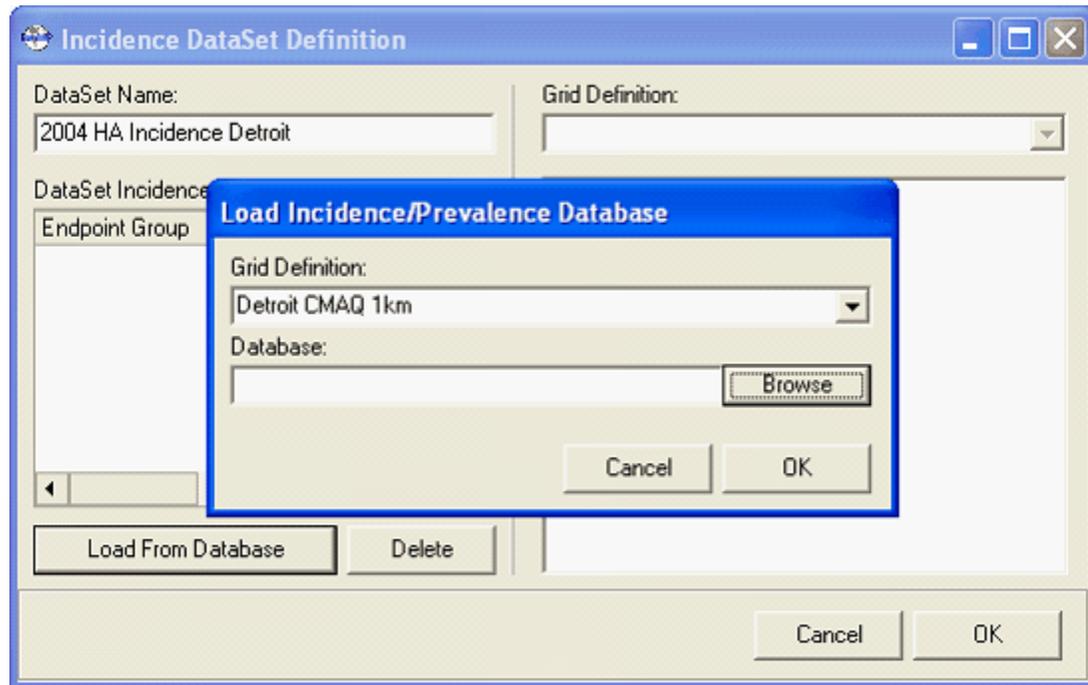
Figure 7-8. *Load Incidence/Prevalence Database* **window**

(j) Next to the *Database* field, click **Browse**. In the new *Open* window, set the *Files of type* to "Excel Files". Under the folder "Inputs" and the subfolder "Detroit", select the file "background_incidence_Detroit" and click **Open**.

(k)This will open a *Select a Table* window. Here you will select the particular Excel sheet (within the Excel file) that contains the new background incidences rates data Using the drop-down menu, select "Sheet1$" (Figure 7-9)
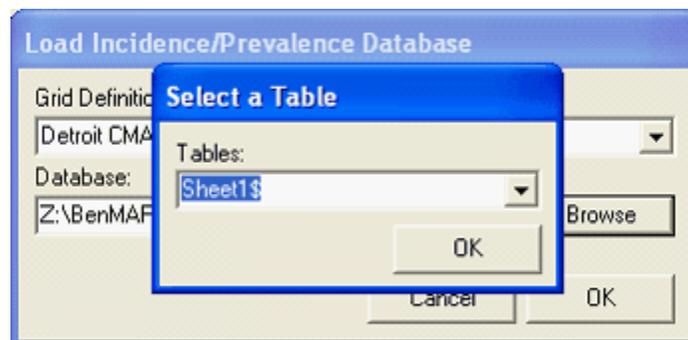
Figure 7-9. *Select a Table* **window**

(l) Click **OK**. This will return you to the *Load Incidence/Prevalence Database* window. Click **OK** in that window to return to the *Incidence Dataset Definitions* window, where fields have now been filled in from the dataset you loaded. Explore some of the background incidence rates by highlighting various age ranges in the "DataSets Incidence Rates" panel and using the scrollbar in the right-hand panel (Figure 7-10).
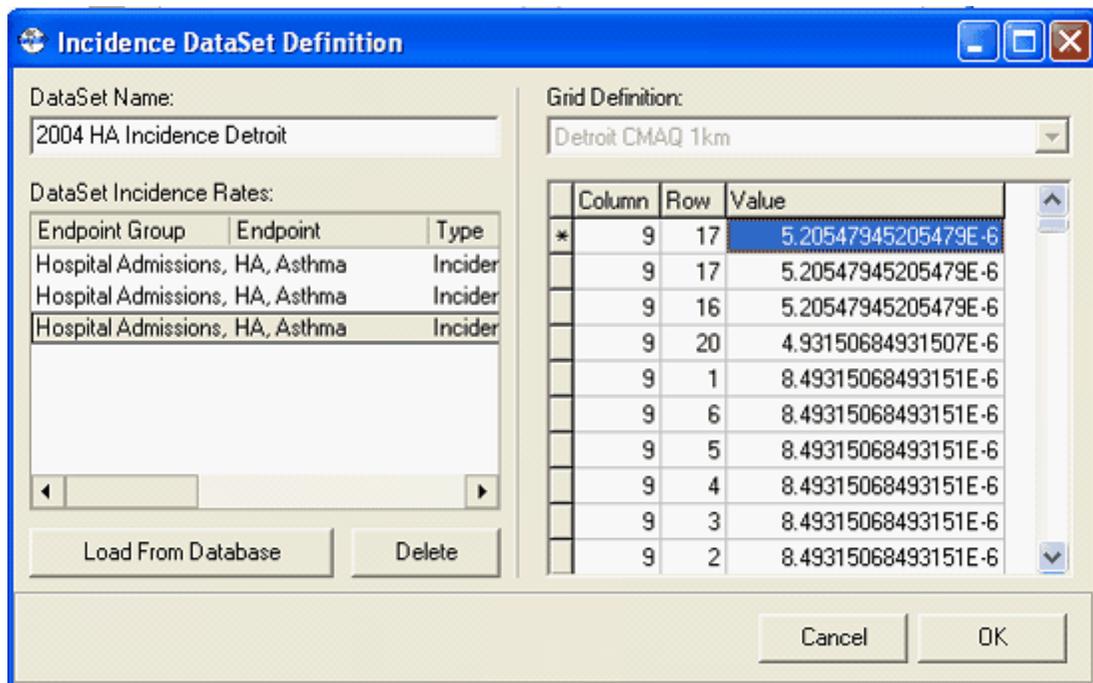
Figure 7-10. Loading and viewing the Detroit background incidence rates

(m) Click **OK**. This will return you to the *Manage Incidence DataSets* window. Scroll through the list of available datasets to confirm that your new background incidence rates dataset is included, then click **OK**. This will return you to the *Manage Setup* window. *Note*: It may take a minute or two to complete the loading of the new dataset (indicated by the *Manage Setup* window becoming active again).

Confirm that the new dataset is in the *Manage Setup* window's incidence/prevalence datasets list. If it is, then you have successfully added a new background incidence rates dataset.

(n) Next we will add a new population dataset. This new dataset will have more specific population information for the Detroit area than the national datasets. Click **Edit** below the *Population DataSets* list. This will open the *Manage Population DataSets* window (Figure 7-11). Here you can add new population datasets or you can delete (but not edit) existing population datasets.

Note: The population datasets are relatively large, so it may take a few minutes to display each of the windows discussed here.
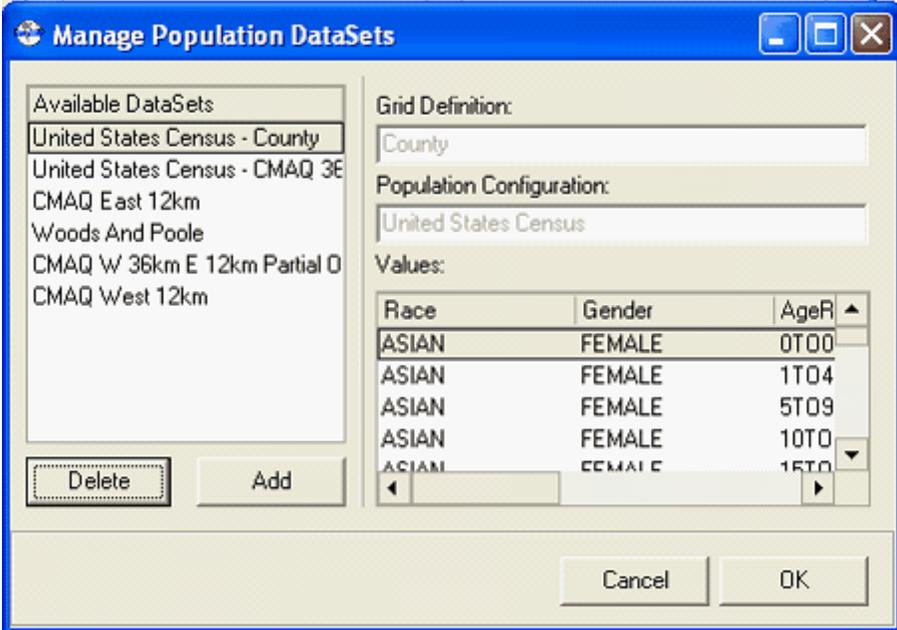
Figure 7-11. *Manage Population DataSets* **window**

(o) Click the **Add** button. This will open a *Load Population DataSet* window (Figure 712).

Figure 7-12. *Load Population DataSet* **window**

(p) Edit the *Population DataSet Name* field by changing the text to "Detroit CMAQ 1km". In the *Population Configuration* field, use the drop-down menu to select "United States Census". For the *Grid Definition* field, use the drop-down menu to select "Detroit CMAQ 1km".

> *Background*: The "Population Configuration" defines the age range, race, and gender variables in your population database. It is critical that the definitions in the population configuration match those used in the development of the database.

(q) Next to the *Database* field, click the **Browse** button. In the new *Open* window, set the *Files of type* to "Text Files". Under the folder "Inputs" and the subfolder "Detroit", select the file "Detroit_Pop" and click **Open**. This will return you to the *Load Population DataSet* window (Figure 7-13).
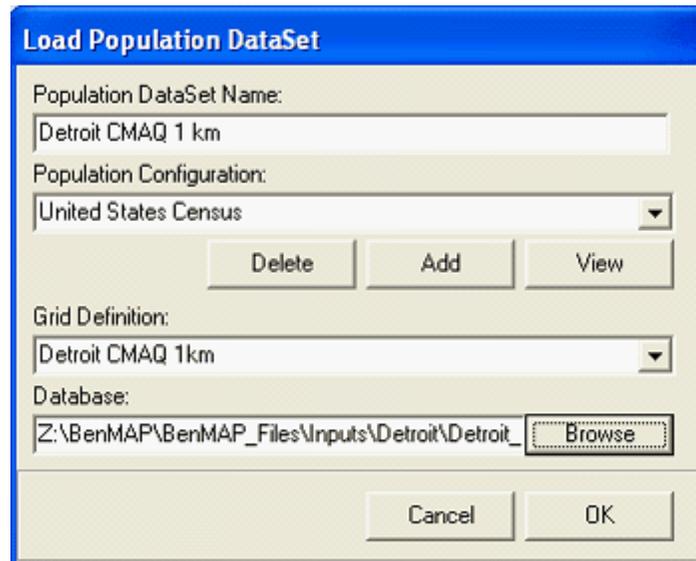
Figure 7-13. Setting the population configuration and the grid definition

(r) Click **OK**. A *Progress* window will appear. It will take a few minutes to load the population dataset. When it is loaded, you will be returned to the *Manage Population DataSets* window, where the fields have now been filled in from the dataset you loaded. Explore your new population dataset by highlighting "Detroit CMAQ 1km" in the *Available DataSets* list and using the horizontal and vertical scrollbars on the right-hand panel (Figure 7-14).
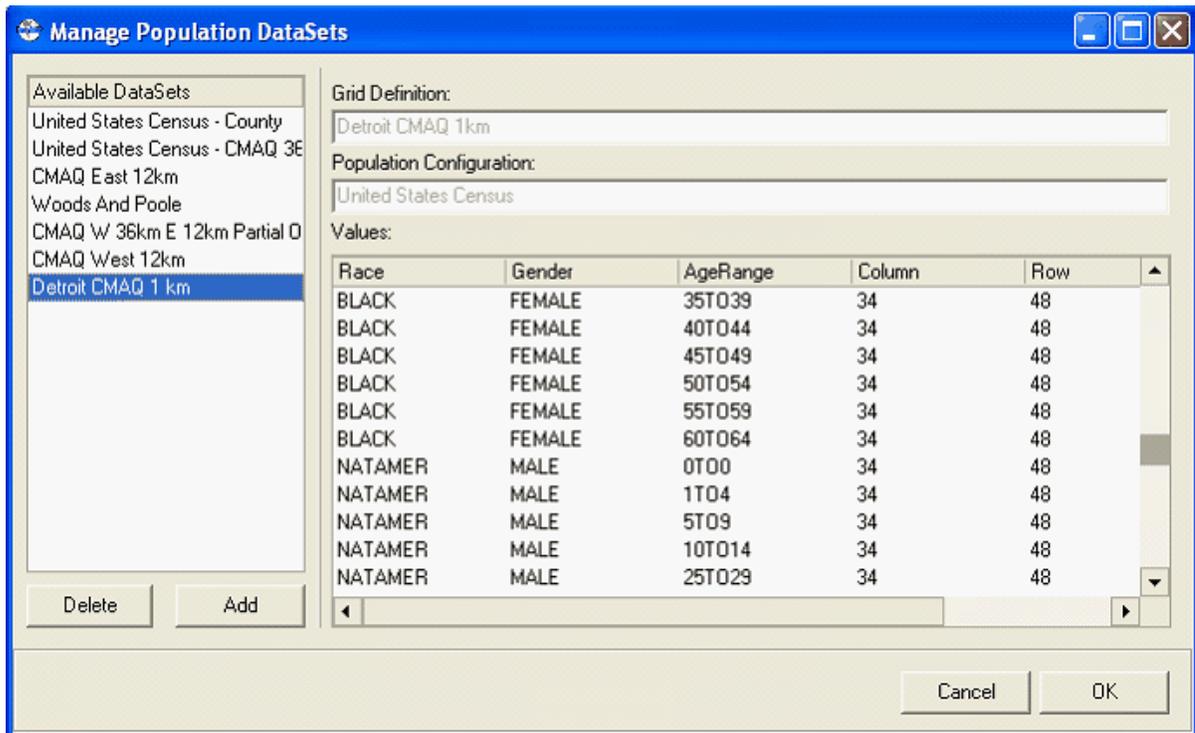
Figure 7-14. Loading and viewing the Detroit population dataset

(s) Click **OK**. This will return you to the *Manage Setup* window. *Note*: It may take a minute or two to complete the loading of the new dataset (indicated by the *Manage Setup* window becoming active again). Confirm that the new dataset is in the *Manage Setup* window's population datasets list. If it is, then you have successfully added a new population dataset.

(t) You have now finished adding the necessary datasets for our Detroit study. From the *Manage Setup* window, click **OK**. This will return you to the main BenMAP window.

Analysis:

The rest of Section 7.2 focuses on analyzing one of the new datasets for Detroit.

(a) We will focus on the new population data for Detroit. Using the "Tools" menu in the main BenMAP window, open a *BenMAP GIS* window. Open the "Population" dataset (Figure 7-15).
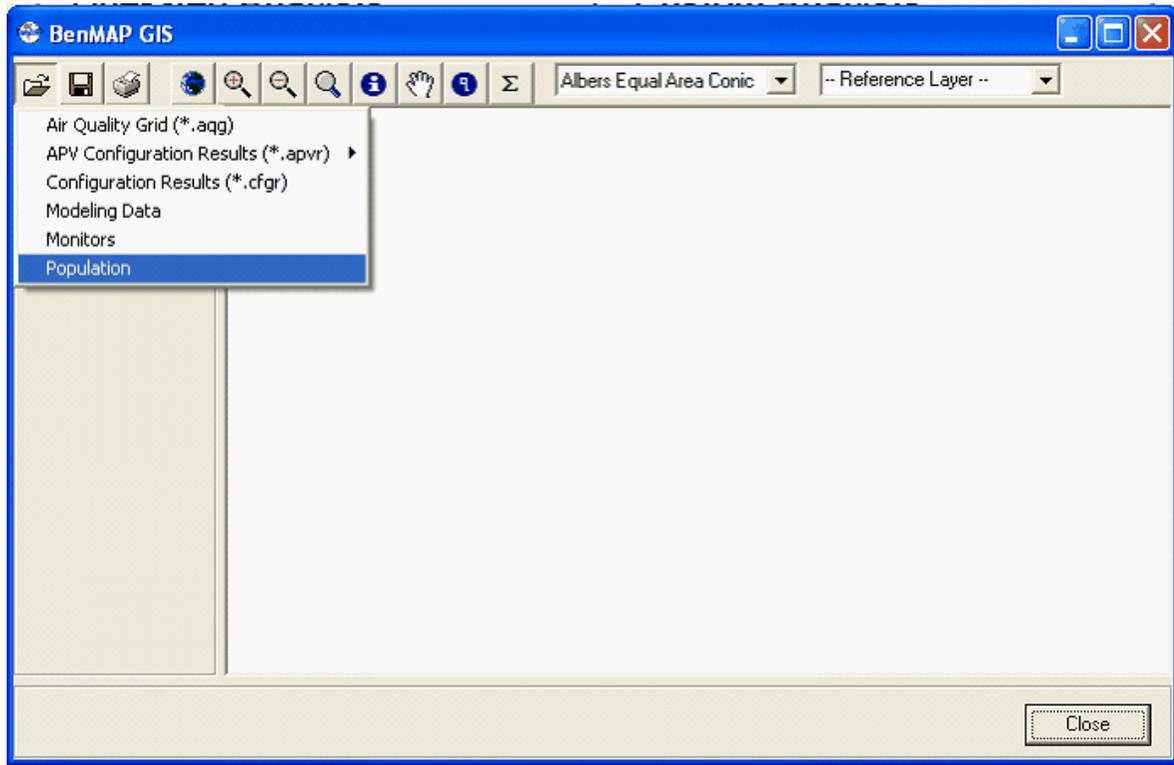
Figure 7-15. Mapping population data

(b) This will open a *Select Population Data* window. Using the drop-down menus, select "Detroit CMAQ 1km" as the population dataset and 2020 as the population year (Figure 7-16). Click **OK**. A *Progress* window will appear. It may take a few minutes for your population data layer to appear in the "Layers" panel.
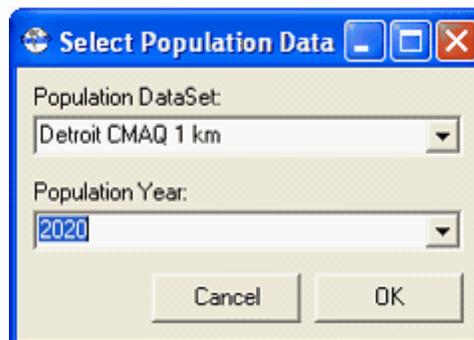
Figure 7-16. *Select Population Data* **window**

(c) Display the population data for African-American females, ages 30 to 34 (B_F_30TO34).

Select the "Detroit CMAQ 1km" as the reference layer. Now, change the reference layer to the county grid. This will automatically zoom out to the national domain. Zoom back in to the Detroit area (Figure 7-17).
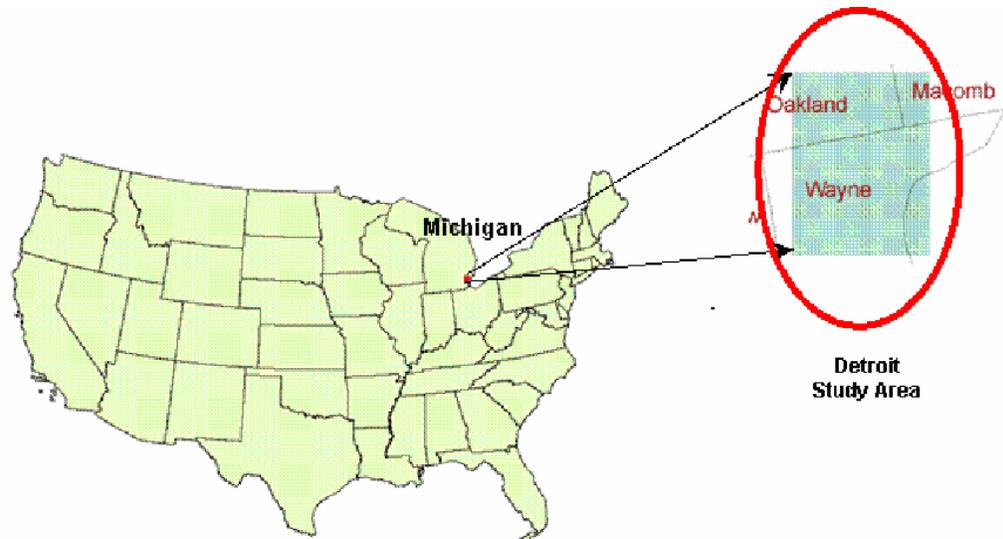
Figure 7-17. Detroit study area

Background: The following are example demographic codes in the Detroit dataset: N_M_40TO44 = Native American male 40 to 44 years; B_F_50TO54 = African-American female 50 to 54 years; A_M_60TO64 = Asian-American male 60 to 64 years; W_F_50TO54 = white female 50 to 54 years; O_M_1TO4 = other male 1 to 4 years.

(d) **Exercise (7.1):** Compare the different demographics from the Detroit population dataset. How does the spatial pattern of African-American females 30 to 34 differ from the spatial pattern of white females 30 to 34?

Answer:

A.1.7.2 Independent Study: Detroit Benefits Analysis

The goal of this exercise is to model the health effects and the subsequent monetized benefits of a change in air pollution (PM2.5 concentrations) for the Detroit region.

Procedures:

The following steps should be thought of as general guidance on how to complete the Detroit benefit analysis. Unlike the earlier labs, this lab does not give detailed instructions on how to complete each stage of the model run and analysis for the Detroit study. If needed, you may refer to previous labs for more detailed instructions on specific steps.

(a) Create a baseline and control aqg from the corresponding model data. You will find two model datasets in the "Detroit" subfolder under the "Inputs" folder. Both datasets are for the 2020 model year and are on the "Detroit CMAQ 1km" grid. The control scenario has some areas reduced to 14.5 $\mu g/m^3$ annual average $PM_{2.5}$ concentrations. Be sure to clearly name your files so that you know which file is the control and which is the baseline.

(b) Open and modify the health incidence configuration created in Section 5. Change the population dataset to your new Detroit population dataset. Change the baseline and control to your Detroit baseline and control.

Add a new function for the "HA, Asthma" endpoint (no threshold). Set the new "HA, Asthma" function's incidence rate to the Detroit-specific background incidence rates dataset. Save a new health incidence configuration and create a results file.

(c) Re-create the apv configuration (from Section 6) using your new Detroit health incidence results. Because we are adding a new incidence endpoint, you will have to start from scratch (see Figures 6-8, 6-19, 6-20, and 6-21 for reference). Include the new "HA, Asthma" incidence result in your "Hospital Admission, Respiratory" pooling.

In selecting the valuation function for "Hospital Admissions, Respiratory", take into account your new age range for your pooled results (i.e., look at the age ranges of the health impact functions that make up the pooled result). Aggregate your health incidence and valuation results to the county level. Save the apv

configuration and create a results file. This completes the modeling portion of the Detroit study.

Analysis:

We provide a series of exercises to guide your analysis of the results from the Detroit study.

(a) **Exercise (7.2):** What pooling type did you use to combine "HA, Pneumonia", "HA, Asthma", and "HA, Chronic Lung Disease" together? Why?

Answer:

(b) **Exercise (7.3):** Compare the total adult population (30-99) from the mortality incidence to the demographic population (Exercise 7.1). How does the "total" population's spatial pattern differ from the patterns for the two demographics (African-American females ages 30 to 34 and white females ages 30 to 34)?

Answer:

(c) **Exercise (7.4):** Why are the aggregated incidence results a significant underestimate of the total change in incidence for the three counties? *Hint*: look at the extent of the nonaggregated domain.

> Answer:

(d) **Exercise (7.5):** What is the total (over the three counties) monetized value for avoided premature mortalities? What is the total monetized value for avoided acute respiratory symptoms? What are the total health incidence results for these two endpoints?

> Answer:

(e) **Exercise (7.6):** Answer the following questions using the totals for each county: What are number of "Hospital Admissions, Respiratory" avoided? What are the number of "HA, Asthma" avoided? What is the background incidence for "HA, Asthma"?

Answer:

(f) **Exercise (7.7):** What is the total (summed over endpoints) monetized benefit for each county? Give the mean and confidence intervals (0.5th and 99.5th percentiles).

Answer:

Synthesis Questions:

The following questions are meant to help you synthesize what you have learned as you have worked through the entire body of BenMAP course material. They draw from multiple labs and course slides.

(a) **Exercise (7.8):** Would you expect the benefits to increase or decrease if you used a later population year? Why?

> Answer:
>
>

(b) **Exercise (7.9):** Would you expect the monetized benefits to be higher or lower if you used a later currency year? Why?

> Answer:
>
>

(c) **Exercise (7.10):** Why might your results differ if you used a national incidence baseline instead of a local incidence baseline for the "HA, Asthma" endpoint?

> Answer:
>
>

(d) **Exercise (7.11):** If some of the health incidence results extended beyond the analysis year (in our case, 2020), should we discount these monetized benefits?

Answer:

(e) **Exercise (7.12):** Unlike in our configuration, EPA generally uses more than one study to model adult mortality. They also do not tend to pool their mortality results. Why might you not want to pool adult mortality and instead report a range for your incidence results?

Answer:

(f) **Exercise (7.13):** Is BenMAP better suited to perform national (large-scale) analyses or local (urban-scale) analyses? Why?

Answer:

(g) **Exercise (7.14):** What are some of the benefits of reducing air pollution that BenMAP does not currently quantify?

Answer:

(h) **Exercise (7.15):** Do you think that valuation estimates would be higher if we used willingness-to-pay (WTP) studies or cost-of-illness (COI) studies?

Answer:

(i) **Exercise (7.16):** What are some of the sources of uncertainty in a BenMAP analysis?

Answer:

A.1.8   Answers to Training Exercises

Section 2 Answers

(2.1)   D24HourMean = 10.56 µg/m³; QuarterlyMean =10.44 µg/m³; lat/lon = (47.35, -68.32).

(2.2)   States of Alabama, California, Georgia, Illinois, Michigan, Pennsylvania, Ohio, Tennessee

(2.3)   QuarterlyMean maximum = 28.32 µg/m³; minimum = 2.76 µg/m³

(2.4)   D8HourMax maximum = 67.68 ppb; minimum = 17.19 ppb
States of Arizona, California, South Carolina, Tennessee, Utah.

(2.5)   No, because the CMAQ grid is a regular grid, whereas the state and other political grids have irregular borders.

(2.6)   States of California, Montana, Oregon.

(2.7)   The age range is 30 to 99 years.

(2.8)   States of California, Oregon, Pennsylvania, Washington.

(2.9)   The number of acute respiratory symptoms avoided is much greater than the number of mortalities avoided.

(2.10)   Because the health incidence values are a function of both delta and population. Therefore, a high delta in a low population area will still have a small health incidence change. In contrast, a relatively low delta in a high population area may have a large health incidence change—in other words, significantly fewer people having that health incidence.

(2.11)   Incidence of mortality: 213; incidence of acute respiratory symptoms: 188,303.
Valuation of mortality: $1.4 billion; valuation of morbidity $66 million.
This means that even though there are far fewer mortalities, they are valued at a much higher rate than the morbidity incidences (as one would expect).

(2.12)   Population year 2020.
CMAQ 36 km

(2.13)   0-17 years.
Norris, G., et al.

Section 3 Answers

(3.1)   States of Arizona, California, Idaho, Maryland, Montana, Nevada, Ohio, Oregon,

Pennsylvania, Utah, Washington, West Virginia, Wyoming.

(3.2)   No, because the C-R functions are based on deltas and population. The human population in the ocean is 0, so both the incidence and valuation over the ocean will be 0.

(3.3)   Initially: County, State, Report_regions, and Nation grids.
After setting Incidence to State, grids at the same scale or coarser are then available.

(3.4)   endpoint = Mortality, All Cause
author = Pope *et al.*
incidence aggregation = valuation aggregation = County

(3.5)   The CMAQ 36km Nation Overlap grid.
Because aggregation occurs in the valuation step, the aggregated health incidence data is in the apvr file, not the cfgr file.

(3.6)   States of California, Washington, Oregon, Utah, Michigan, Ohio, Pennsylvania, New Jersey, New York, Maryland, and Virginia.

(3.7)   The County grid.

(3.8)   State code = 6, FIPS code = 37 (Los Angeles County); avoided mortalities = 132.76

(3.9)   health incidence maximum = 66.68; sum = 1707.67
pooled incidence maximum = 132.76; sum = 1707.67

(3.10) State code = 6, FIPS code = 37 (Los Angeles County); mortality = $876 million; morbidity = $38 million
National mortality: $11.27 billion; National morbidity: $403 million.

(3.11) For col 1, row 9, mortality has the greatest standard deviation = 247,031
Chronic bronchitis has the greatest coefficient of variation = 16.587/13,520 = 1.23

(3.12) The delta for the 14/35 analysis is significantly larger. For example, a larger portion of the states have more than 0.2 $\mu g/m^3$ difference.

(3.13) Illinois: 227 mortality for 14/35 scenario, 3.75 mortality for 15/35 scenario.
California: 559 mortality for 14/35 scenario, 573 mortality for 15/35 scenario.
Overall, the 14/35 scenario has more states showing significant numbers of avoided mortalities, especially in the East, whereas the 15/35 scenario has some western states with slightly greater reductions (e.g., California, Oregon, Washington). Total number of mortalities avoided for the lower 48 states is 4,787 for the 14/35 scenario and 1,707 for the 15/35 scenario.

(3.14) California: $3.69 billion saved in prevented mortality, $147 million saved in prevented morbidity.

Illinois: $1.5 billion saved in prevented mortality, $50 million saved in prevented morbidity.

$31.6 billion mortality and $1.1 billion morbidity for the 14/35 scenario versus $11.2 billion mortality and $403 million morbidity for the 15/35 scenario. In other words, our intuition was correct that the 14/35 scenario had higher benefits than the 15/35 scenario.

Section 4 Answers

(4.1)   The adjusted aqg's have no values greater than 15 μg/m^3. In comparison, the non-adjusted aqg's do have regions with annual values greater than 15 μg/m^3.

(4.2)   The VNA AQG is much smoother, because it uses distance weighting to smooth the AQG between monitor locations.

(4.3)   Neighbor scaling type = inverse distance;
Rollback region = California (6), Oregon (41), Pennsylvania (42), and Washington (53);
Rollback method = percentage, the percentage of rollback = 0.1 (i.e. 10%)

(4.4)   California, Ohio, West Virginia, Maryland, Georgia, and Alabama.

(4.5)   Region 1: using peak shaving for inter and intraday rollback. Rolling back to a standard (attainment test) of 35 μg/m^3 on the D24HourMean metric and an ordinality of 1.
Region2: using incremental rollback, reducing all monitors by an increment of 4 μg/m^3
.

(4.6)   The Western states have a generally constant delta. This makes sense because we applied an incremental change (a constant reduction) to all the monitors in the region. In contrast, the East Coast has been reduced to a standard. Therefore, only areas that were over the standard (35 μg/m^3 daily mean) will be reduced. We see that the most significant changes are in Pennsylvania and to a lesser degree in Georgia.
If we look at the eastern edge of the West coast states, we notice that the delta is not constant across each state. Initially, we might expect the deltas to be constant across the state, because we applied an incremental change to all the monitors. However, the AQG is the result of interpolating from the monitors to the grid cells. Therefore, on the eastern edge we are interpolating between monitors that did have a rollback and those areas that had no change.

Section 5 Answers

(5.1)   QuarterlyMean maximum = 1.67 μg/m^3.
QuarterlyMean mean = 0.11 μg/m^3 (misleading because includes model domain over the ocean).

Washington, California, Oregon, Georgia, Pennsylvania, West Virginia.

(5.2) Two functions.
Both functions are by Norris.
Differences include a qualifier (10 µg/m³ threshold), Beta and P1Beta (standard deviation), and C (C=10 for the threshold function).

(5.3) Endpoint: "HA, Chronic Lung Disease (less Asthma)". "HA, Chronic Lung Disease", HA, Pneumonia", "HA, Asthma".
Four functions: two by Ito and two by Moolgavkar.

(5.4) Pope.
Three functions are by Pope.
The three functions have different thresholds (0, 7.5, and 10 µg/m³).
The current configuration uses the 0 µg/m³ threshold Pope function.

(5.5) Pope age range 30-99 years, Woodruff age range 0-0 years.

Pope function = $\left(1 - \dfrac{1}{e^{\beta \Delta Q}}\right) \times Inc \times Pop$

Woodruff function = $\left(1 - \dfrac{1}{(1 - Inc)e^{\beta \Delta Q} + Inc}\right) \times Inc \times Pop$

(5.6) Ito and Moolgavkar have the same functional form.
Ito's $\beta = 0.001169$.
Moolgavkar's $\beta = 0.00183$.
Bonus: Moolgavkar is more sensitive to changes in AQ. An equal change in AQ will result in a larger change in the health incidence from Moolgavkar's function compared to Ito's function. Mathematically, the larger β, the smaller the value of $\dfrac{1}{e^{\beta \Delta Q}}$, and hence the larger the value of the whole function.

(5.7) Population year = 2020.
12 Concentration-Response (C-R) functions.
Woodruff location is 86 cities and incidence dataset is 2020 Mortality Incidence Rates

(5.8) Pope maximum = 66.78 premature deaths (ages 30-99) avoided.
Woodruff maximum = 0.15 premature deaths (infant) avoided.
Similar but not the exact same spatial patterns.

(5.9) Ito maximum = 3.01.
Moolgavkar maximum = 4.73.
Moolgavkar population is the same as the Ito population.
They have identical inputs. The only difference between the two functions is the β; therefore, the Moolgavkar study is more sensitive to changes in AQ.

(5.10) AMI18 maximum = 3.63, population maximum = 2,434,707.
AMI65 maximum = 65.36, population maximum = 988,819.
Although the population for the 18-44 group is much greater than the population for the 65-99 group and the functions are the same, the incidence results are much smaller for the 18-44 age group. This implies that the incidence rate for younger group is much smaller than the incidence rate for the older group. This result reflects our intuition.

(5.11) baseline = 5,452.946.
point estimate = 4.7457.
5th percentile = 4.0085.
95th percentile = 5.4827.

(5.12) QuarterlyMean maximum = 4.56 $\mu g/m^3$.
QuarterlyMean mean = 0.32 $\mu g/m^3$.
California, Washington, Oregon, Idaho, Wyoming, Montana, Utah, Pennsylvania, West Virginia.
Compared to Exercise 5.1, the adjusted control has much larger AQ delta values and larger geographic areas of significant delta, especially in the West.
The adjusted scenario should have a significantly larger change in adverse health effects because it has a significantly larger change in AQ.

(5.13) Population year = 2020.
12 CR functions.
Number of Latin Hypercube Points = 10.

(5.14) Pope maximum = 196.13.
Woodruff maximum = 0.52.
The adjusted results are greater than the nonadjusted results.

(5.15) AMI18 maximum = 11.13, population maximum = 2,434,707.
AMI65 maximum = 189.75, population maximum = 988,819.
The adjusted results are greater than the nonadjusted results. The adjusted population is equal to the nonadjusted population. There was also a change in spatial pattern. For the adjusted results, most of the significant results are in the West.

(5.16) baseline = 5452.946.
point estimate = 11.302.
5th percentile = 9.5474.
95th percentile = 13.0561.
The adjusted baseline equals the nonadjusted baseline. The adjusted mean is greater than the nonadjusted mean. The adjusted statistical spread and values are greater than the nonadjusted statistical spread and values.

Section 6 Answers

(6.1) Sum (Dependent) because the health incidence populations are nonoverlapping (distinct age groups).

(6.2) We used Sum (Dependent) because the four AMI valuation results are for nonoverlapping populations (distinct age groups).

(6.3) Income growth adjustment year = 2020.
HA pooling and mortality pooling both use Sum (Dependent).
HA Chronic Lung Disease pooling uses Random / Fixed Effects, with advanced pooling using round weight to two digits.
ER endpoint group valuation pooling = Subjective Weights
AMI endpoint group valuation pooling = Sum (Dependent)
Population year = 2020.

(6.4) California mortalities = 560, acute respiratory symptoms = 503,139.
Pennsylvania mortalities = 342, acute respiratory symptoms = 184,340.
Illinois mortalities = 227, acute respiratory symptoms = 163,566.

(6.5) AMI18 maximum = 30.
AMI65 maximum = 555.

6.6) California acute respiratory = $25.4 million, morbidity = $289.6 million, mortality = $3.7 billion.
Pennsylvania acute respiratory = $9.3 million, morbidity = $141.6 million, mortality = $2.3 billion.
Illinois acute respiratory = $8.3 million, morbidity = $106.1 million, mortality = $1.5 billion.
The mortality incidence numbers are far smaller than the acute respiratory symptom numbers, but the mortality valuation is far greater than the acute respiratory valuation. Therefore, the mortality valuation function (VSL) must be much larger than the acute respiratory symptoms function.

(6.7) AMI point estimate = $79.3 million.
0.5^{th} percentile = $14.6 million.
99.5^{th} percentile = $193.5 million.

(6.8) Population year = 2020.
Smith subjective weight = 0.60, Standford subjective weight = 0.40.

(6.9) California mortalities = 2,295, acute respiratory symptoms = 2.1 million.
Pennsylvania mortalities = 365, acute respiratory symptoms = 196,460.
Illinois mortalities = 231, acute respiratory symptoms = 166,252.
The adjusted health incidence results are greater than the nonadjusted results, especially in the West.

6.10) California acute respiratory = $107.2 million, morbidity = $1.19 billion, mortality = $15.15 billion.
Pennsylvania acute respiratory. = $9.9 million, morbidity = $151 million, mortality = $2.41 billion.
Illinois acute respiratory = $8.4 million, morbidity = $108 million, mortality = $1.53 billion.
The adjusted valuation results are greater than the nonadjusted results, especially in the West.

(6.11) AMI point estimate = $320 million.
0.5^{th} percentile = $59.7 million.
99.5^{th} percentile = $772 million.
The adjusted mean is greater than the nonadjusted mean. The adjusted statistical spread and values are greater than the nonadjusted statistical spread and values.

Section 7 Answers

(7.1) African-American females age 30-34 are concentrated in the central and eastern band of our domain (Northern Wayne Co.). White females age 30-34 are more concentrated in the northern (southern Oakland and Macomb Co.) and southwestern (central Wayne Co.) parts of our domain.

(7.2) We used sum (dependent) to combine together the various "HA" endpoints. We used random/fixed effects to combine together the Ito and Moolgavkar "HA, Chronic Lung Disease" results into one result. The sum (dependent) pooling makes sense for the three endpoints because the populations are distinct, nonoverlapping. On the other hand, the two "HA, Chronic Lung Disease" results are for an overlapping population; therefore, the random/fixed effects pooling is appropriate.

(7.3) The spatial pattern of total adult population is more spatially homogeneous than the pattern of the demographic data. It appears to be more similar to the combination of the African-American female and white female spatial patterns than to either of the individual demographic patterns.

(7.4) The aggregated incidence results are a significant underestimate of the total change in incidence for the three counties because the "Detroit CMAQ 1km" grid does not cover the entire spatial extent of the three counties. In other words, we are characterizing the whole county's incidence change based on a subset of the county, i.e., based on a subset of the population.

(7.5) Total monetized mortality benefit = $1.75 billion
Total monetized acute respiratory symptoms benefit = $8.64 million
Total mortality incidence results = 264
Total acute respiratory symptoms results = 170,868

(7.6) Macomb (FIPS 99):"HA, Respiratory " = 9.4, "HA, Asthma" = 0.85, "HA, Asthma"

baseline = 201.3
Oakland (FIPS 125): "HA, Respiratory " = 35.6, "HA, Asthma" = 3.77, "HA, Asthma"
baseline = 608.3
Wayne (FIPS 163): "HA, Respiratory " = 115.0, "HA, Asthma" = 36.5, "HA, Asthma"
baseline = 3,954.3

(7.7) Macomb (FIPS 99) total monetized benefit: mean = $119 million, 0.5^{th} = -$7.68 million, 99.5^{th} = $347 million
Oakland (FIPS 125) total monetized benefit: mean = $464 million, 0.5^{th} = -$29.2 million, 99.5^{th} = $1.37 billion
Wayne (FIPS 163) total monetized benefit: mean = $1.28 billion, 0.5^{th} = -$81.1 million, 99.5^{th} = $3.75 billion

(7.8) For most areas, a later population year would mean a greater population. Because most of the health impact functions are proportional to population, a greater population would mean a greater benefit.

(7.9) A later currency year would generally mean that the monetized benefits would be higher. The later the currency year, generally the greater the inflation, and hence the less the purchasing power of an individual dollar. Therefore, the same benefit in a later currency year would equal a greater number of dollars.

(7.10) A national background incidence rate would not reflect the local incidence rates for specific health endpoints. A local background incidence rate would likely more closely reflect the local population characteristics than the national average. The background incidence rates are an important variable in the underlying health impact function.

(7.11) The AMI functions calculate a change in benefits beyond 2020. The AMI results are for five years of medical costs and five years of opportunity costs. Because these costs span five years (2020-2025), the benefits should be discounted back to 2020. We use a discount rate to reflect our tendency to value future costs less than present costs. In other words, if the combined costs were incurred only in 2020, we would value these costs more than if these same costs were spread over five later years.

(7.12) By not pooling their adult mortality results, EPA is reporting a range of reasonable results. You can think of the different mortality results as spanning the minimum to maximum estimations of the health incidence results. In other words, a range provides a window between the worst- and best-case scenarios.

(7.13) It depends. If you use only the national and regional-level health data pre-loaded in BenMAP, it is better suited for national-scale analyses. Most of the input datasets have been developed with a national or at least regional perspective. For example, many of the health incidence functions were studied using large populations spread across multiple cities and states. As one goes to a more local scale, the default health incidence functions, valuation functions, and background incidence rates may become less and less representative of the local population. In addition, if the scale becomes

very small, then the populations become very small. If you have a small population, then the sample size may become problematic, which undermines the statistical functions used in BenMAP.

In contrast, imagine that you are able to find high quality local health incidence functions, valuation functions, and background incidence rates for your particular study area. You would expect that BenMAP would give more representative results from these specific functions and datasets that have been "tuned" to your particular study area than it would for a national analysis.

(7.14) BenMAP does not quantify the following benefits of reducing air pollution: improved ecosystem health, climatological benefits, improved visibility and consequentially improved aesthetics, and the reductions of pollution damages to infrastructure and buildings. However, with the addition of appropriate impact functions, it would be possible to use BenMAP to quantify these endpoints.

(7.15) The valuation estimates would probably be higher if we used WTP functions instead of COI functions. COI functions do not include the cost of pain and suffering in the estimate of monetized value. WTP functions attempt to capture both COI and the cost of pain and suffering.

(7.16) There are many sources of uncertainty in a human health benefit analysis. EPA has attempted to quantify some sources of uncertainty in BenMAP. For example, the uncertainty in the regression coefficients for the health impact functions and the underlying distribution are included in the valuation functions. Other uncertainties have not been quantified. For example, there is significant uncertainty in the baseline and control AQG, in the geographic variability of functions (i.e., which functions are really regional or local and do not translate to other areas), in the differences between personal exposure and outdoor pollution concentrations, and in the background incidence rates.

A.2   CityOne

Below is a very simple tutorial using the CityOne setup available at the BenMAP website (http://www.epa.gov/air/benmap/). The tutorial is based on a hypothetical scenario where ambient $PM_{2.5}$ concentrations are reduced by 25 percent in 2003. The steps in this analysis are as follows:

Step 1. Data Files Needed for Training

Step 2. Create Air Quality Grids for the Baseline & Control Scenario

Step 3. Specify Configuration Settings

Step 4. Select Health Impact Functions

Step 5. Specify Aggregation, Pooling and Valuation

Step 6. Generate Reports

Step 7. View Your Reports

Step 8. Map Your Results

Each step is explained in detail below.

A.2.1 Step 1. Data Files Needed for Training

To do this training you need to download the CityOne files from the BenMAP website: http://www.epa.gov/air/benmap/. The data are most easily accessible if loaded in using the Data Import tool, which is discussed in the here in the chapter on loading data.

A.2.2 Step 2. Create Air Quality Grids for the Baseline and Control Scenarios

Click on the **Create Air Quality Grids** button to begin inputting the air quality data needed by BenMAP. This will open up the window where you will input the air quality data. In general, you need two air quality grids to conduct a benefit analysis, one for a baseline scenario and one for the policy you are evaluating (the control scenario). We will be creating our baseline and control scenarios together, through the **Monitor Rollback** air quality grid creation method.
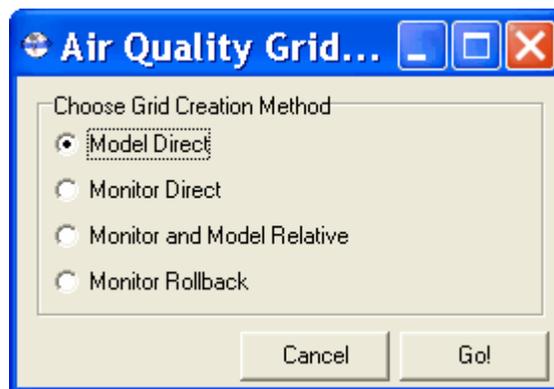
Select *Monitor Rollback* from the list and click on **Go!**

This will take you to the **Monitor Rolback Settings: (1) Select Monitors** screen where you will enter the information about the air quality monitoring data you want to use.

Choose *PM2.5* from the **Pollutant** drop-down menu. On the **Library** tab, choose *CityOne Monitors* from the **Monitor DataSet** drop-down menu, and choose *2003* from the **Monitor Library Year** drop-down list. Finally, in the **Rollback Grid Type** choose *Metropolitan Area*.

When your window looks like the window above, click **Next**.

This will take you to the **Monitor Rolback Settings: (2) Select Rollback Regions and Settings** window where you will choose the type of rollback for the *CityOne* metropolitan area.
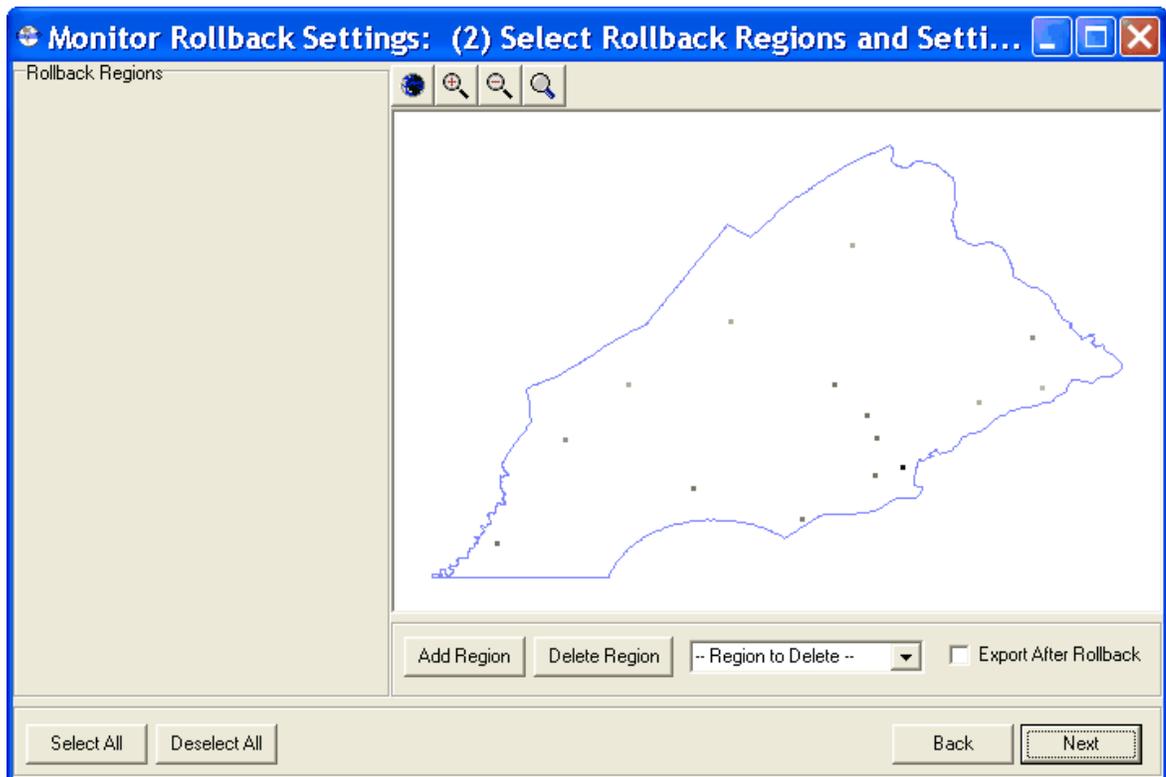
Click **Add Region**. This will bring up the **Select Region Rollback Type** window.
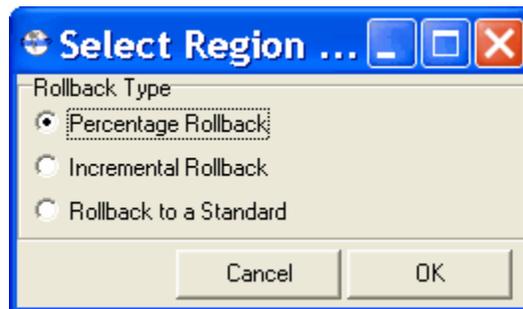
In the **Select Region Rollback Type** window you may select from three rollback options.
Select the **Percentage Rollback** option as shown above. Click **OK**.
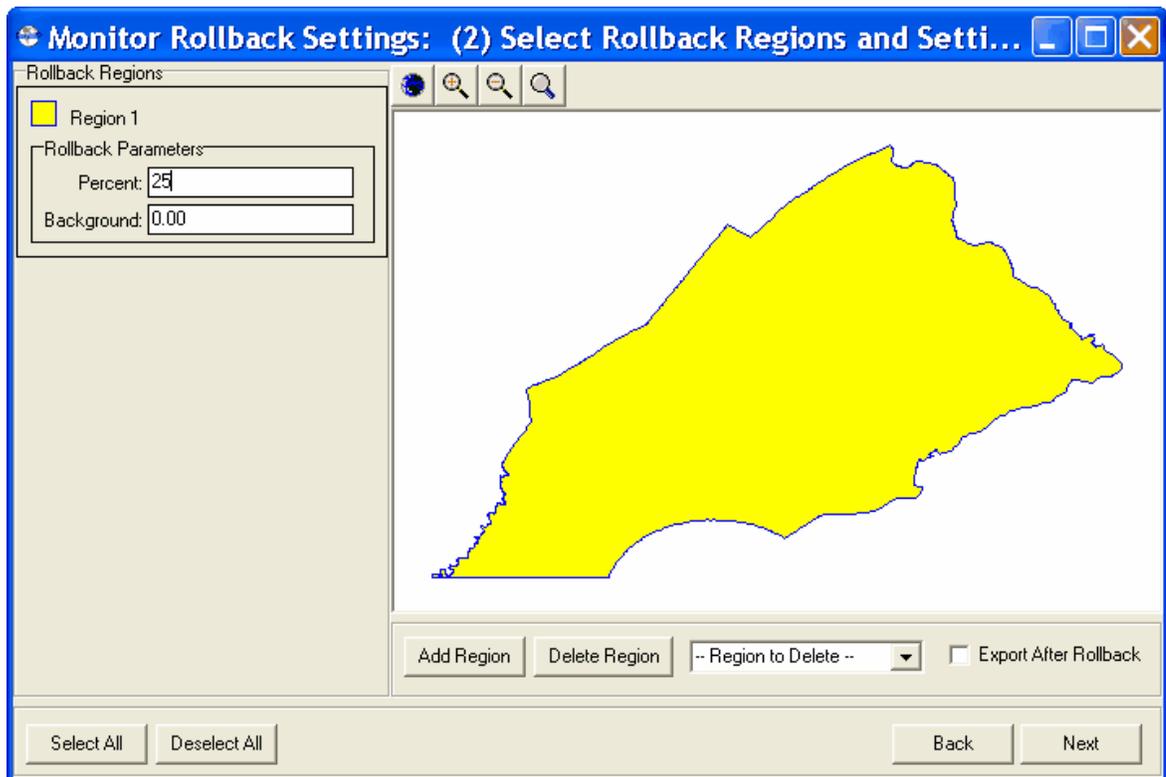
In the box of **Rollback Parameters** for **Region 1**, type 25 in the **Percent** box. (This will reduce each of the monitors in the CityOne area by 25 percent.) Then click on the **Select All** box. When your window looks like the window above, click **Next**.

This will take you to the **Monitor Rolback Settings: (3) Additional Grid Settings** window, the final step in creating rollback grids.

Choose the **Voronoi Neighborhood Averaging** interpolation method. Leave the scaling method as **None**. From the **Grid Type** drop-down list choose *County*. Leave the box checked next to **Make Baseline Grid (in addition to Control Grid)**. This option will cause BenMAP to create a baseline scenario air quality grid using the monitors selected in the previous step, but without rolling their values back. BenMap will create a second grid with the rolled back monitors, which will serve as our control scenario air quality grid.
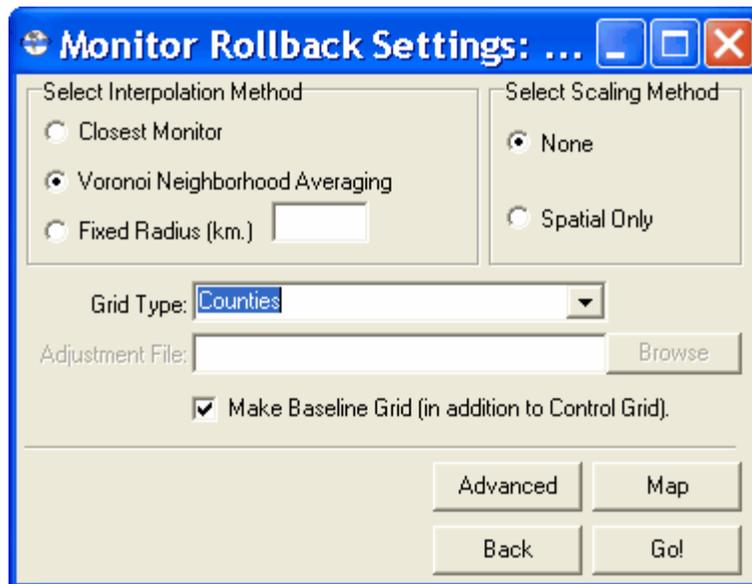
When your window looks like the window above, click **Go!**.

BenMAP will now prompt you to save the baseline air quality grid. Make sure you are in the *Air Quality Grids* subfolder in the *BenMAP* directory and then save the file as: *PM2.5 CityOne County Baseline 2003 VNA.aqg* (you do not have to enter the ".aqg" extension).
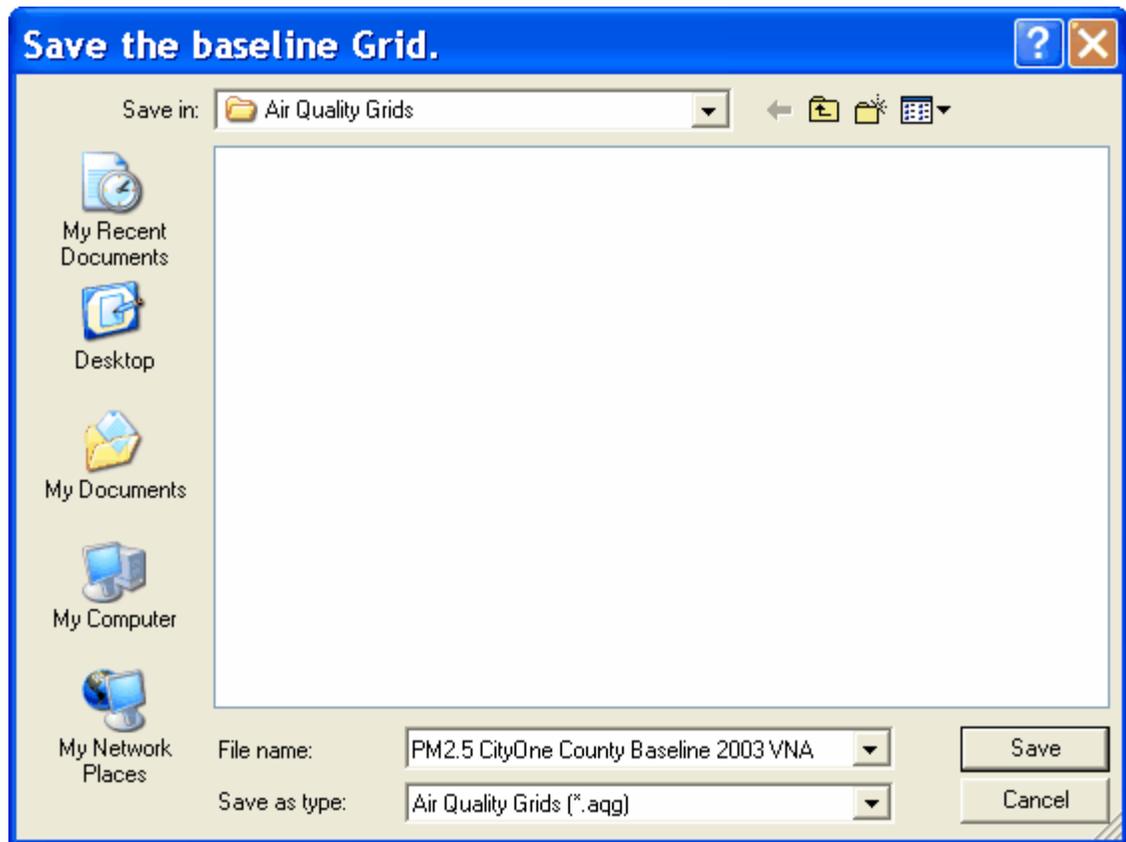
BenMAP will now prompt you to save the control air quality grid. Again, make sure you are in the *Air Quality Grids* subfolder in the *BenMAP* directory and then save the file as: *PM2.5 CityOne County 25 Pct Rollback 2003 VNA.aqg* (you do not have to enter the ".aqg" extension).
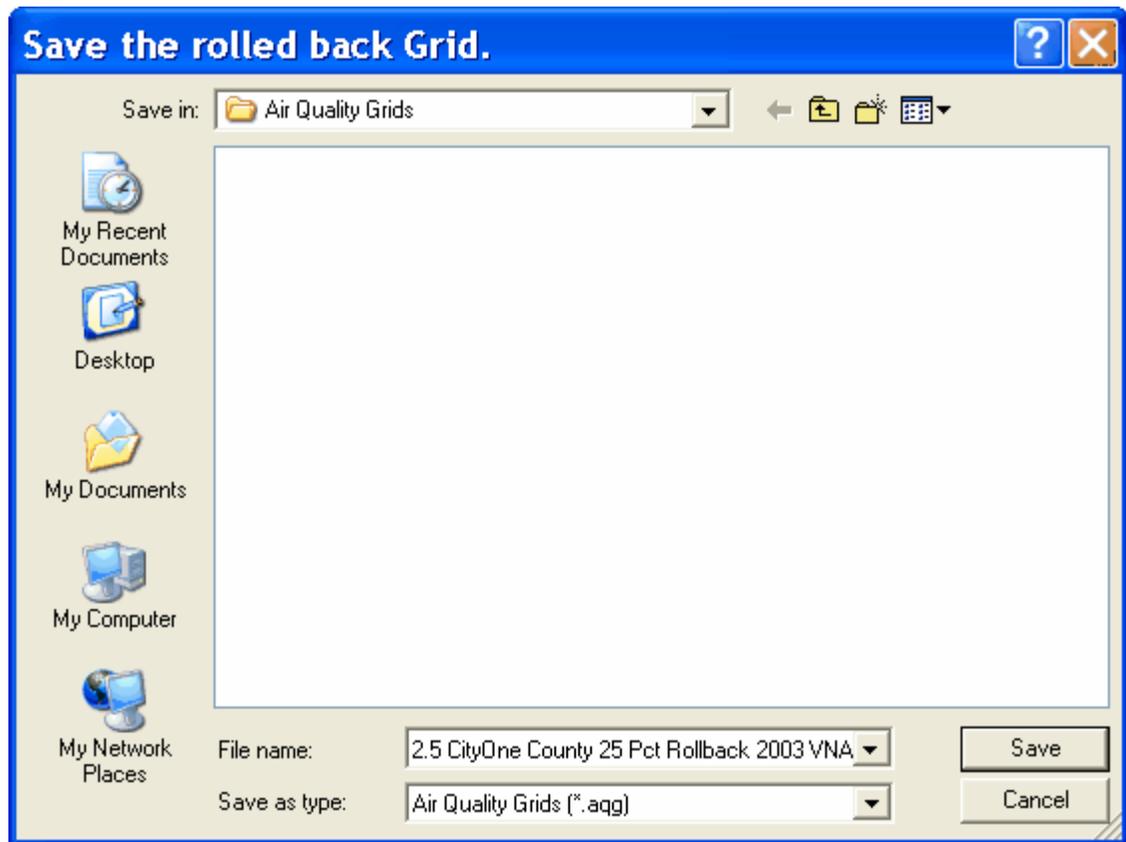
BenMAP will now create baseline and control air quality grids that you can use in your benefit analysis. When the progress bar is complete, BenMAP will return to the main BenMAP screen.

A.2.3   Step 3. Specify Configuration Settings

On the main BenMAP screen, click on the **Create and Run Configuration** button. In the following box, select *Create New Configuration* and click **Go!**.
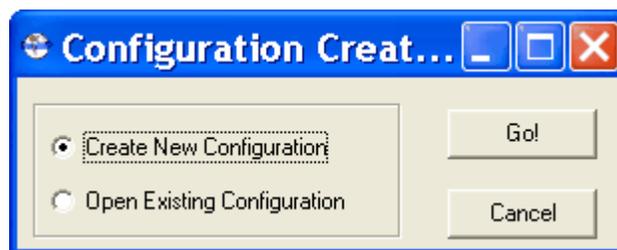
This will bring up the **Configuration Settings** form, where you will enter the basic information about your analysis before selecting the health effects you wish to estimate.

In the **Baseline File** field, you can either enter the path for your baseline air quality grid, or click **Open**. For this example, click **Open** and browse to the *Air Quality Grids* folder. Select *PM2.5 CityOne County Baseline 2003 VNA* and click **Open**.

Next, click on **Open** next to the **Control File** field and select *PM2.5 CityOne County 25 Pct Rollback 2003 VNA* and click **Open**.

This specifies that you want to conduct a benefit analysis of the difference between the baseline and control scenarios for which we created air quality grids in Step 2.

In the **Settings** section of this window, there are several fields which set the overall scope of the analysis.

In the **Population DataSet** field, select *CityOne Tract Population* from the drop down menu. This tells BenMAP that you want your analysis to use tract-level population data from this dataset when calculating health impacts.

In the **Population Year** field, enter *2005* or select *2005* from the drop down menu. This tells BenMAP that you want your analysis to use 2005 populations when calculating health impacts.

In the **Latin Hypercube Points** field, enter *10* or select *10* from the drop down menu. This tells BenMAP that you want to estimate the percentiles of the distribution of health endpoint incidence using Latin Hypercube Sampling with 10 percentiles of the distribution, representing the 5th, 15th, 25th, and so on up to the 95th percentile.

Leave the **Run in Point Mode** box unchecked.

Leave the **Threshold** field blank. This tells BenMAP that you want to estimate benefits associated with all changes in PM2.5, regardless of where those changes occur along the range of PM2.5 concentrations. Selecting a non-zero threshold means that you would only want to calculate benefits for changes occurring above the threshold.

When your window looks like the above, click **Next**.

This will bring up the next page of the **Configuration Settings** form, where you can select health impact functions from a set of available health impact functions.
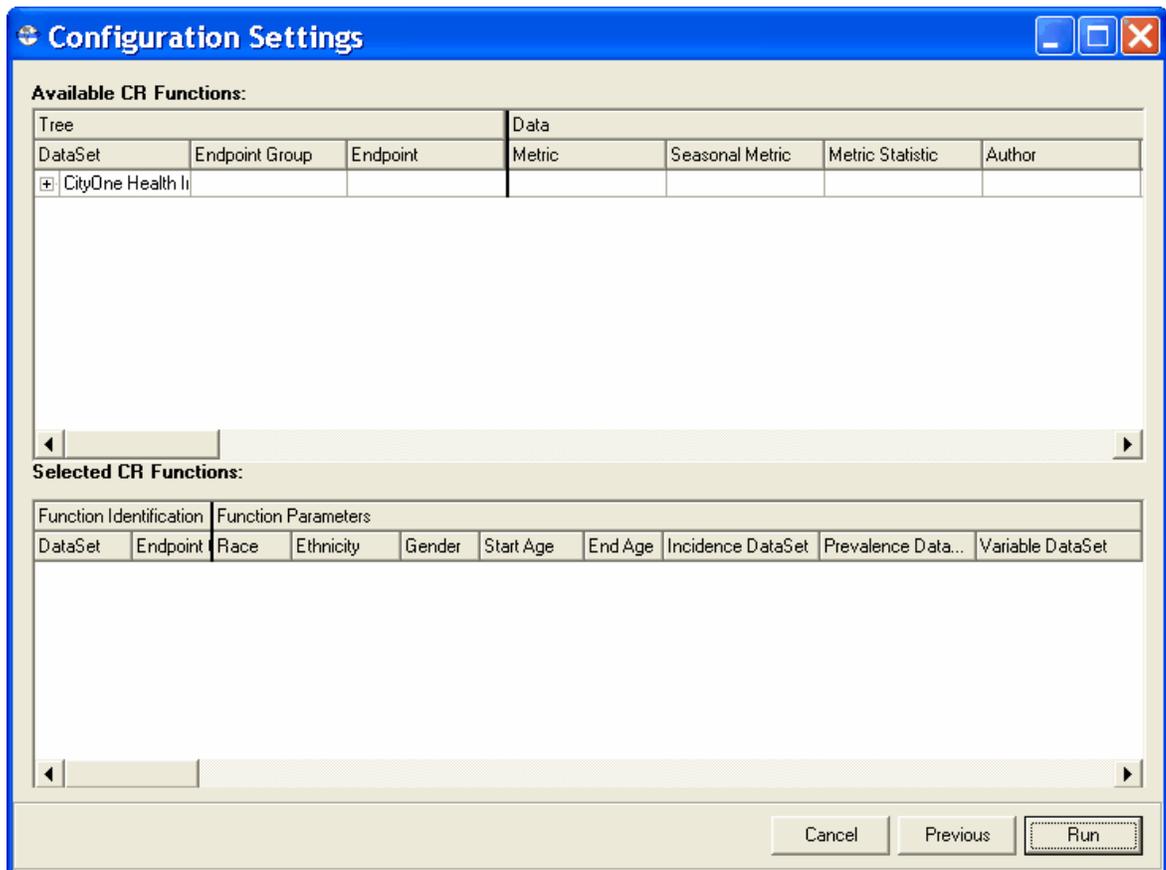
(Note: This screen can be resized if you are having trouble seeing all of the information. Individual columns can also be resized. Just click on the border of a column and drag to increase or decrease its width.)

A.2.4 Step 4. Select Health Impact Functions

In this screen, you can select health impact functions to use in your analysis. For this example, we are going to estimate the change in incidence of three health endpoints associated with PM2.5: acute bronchitis, acute myocardial infarctions (heart attacks), and emergency room visits for asthma. To select a health impact function, you must drag it from the upper box (**Available C-R Functions**) in the window to the lower box (**Selected C-R Functions**). You can drag groups of health impact functions over, or drill down and drag over individual functions.

For acute bronchitis, drill down until you see the function by the **Author** *Dockery et al*. Drag the function into the lower panel of the window. You should see a new row with the **Endpoint Group** *Acute Bronchitis*.
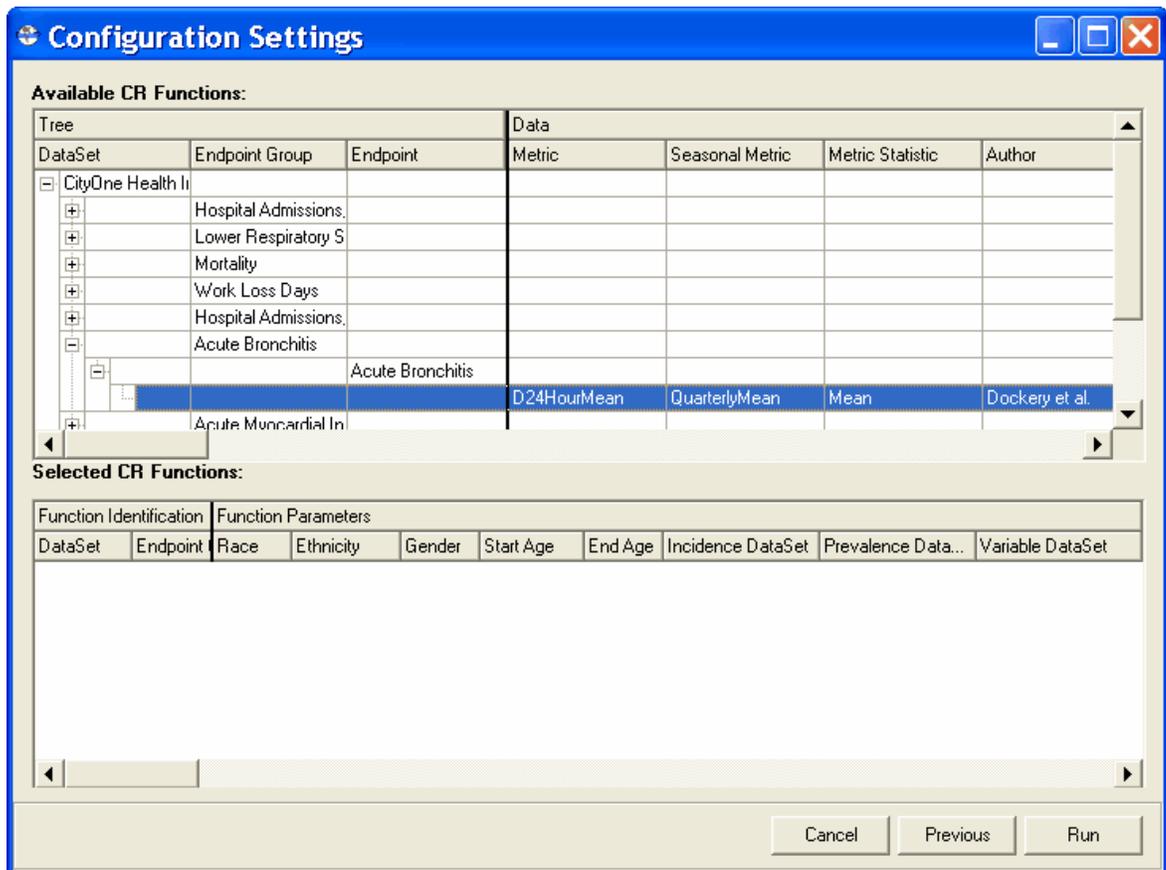
For acute myocardial infarctions (AMI), drag the entire **Endpoint Group** titled *Acute Myocardial Infarction* to the lower panel (do not drill down). This will include the full set of age-specific health impact functions for AMI.

For asthma emergency room visits, also drag over the entire the **Endpoint Group** titled *Emergency Room Visits, Respiratory*.

You should now have seven health impact functions listed in the lower panel: one acute bronchitis function, five AMI functions, and one ER visit function.
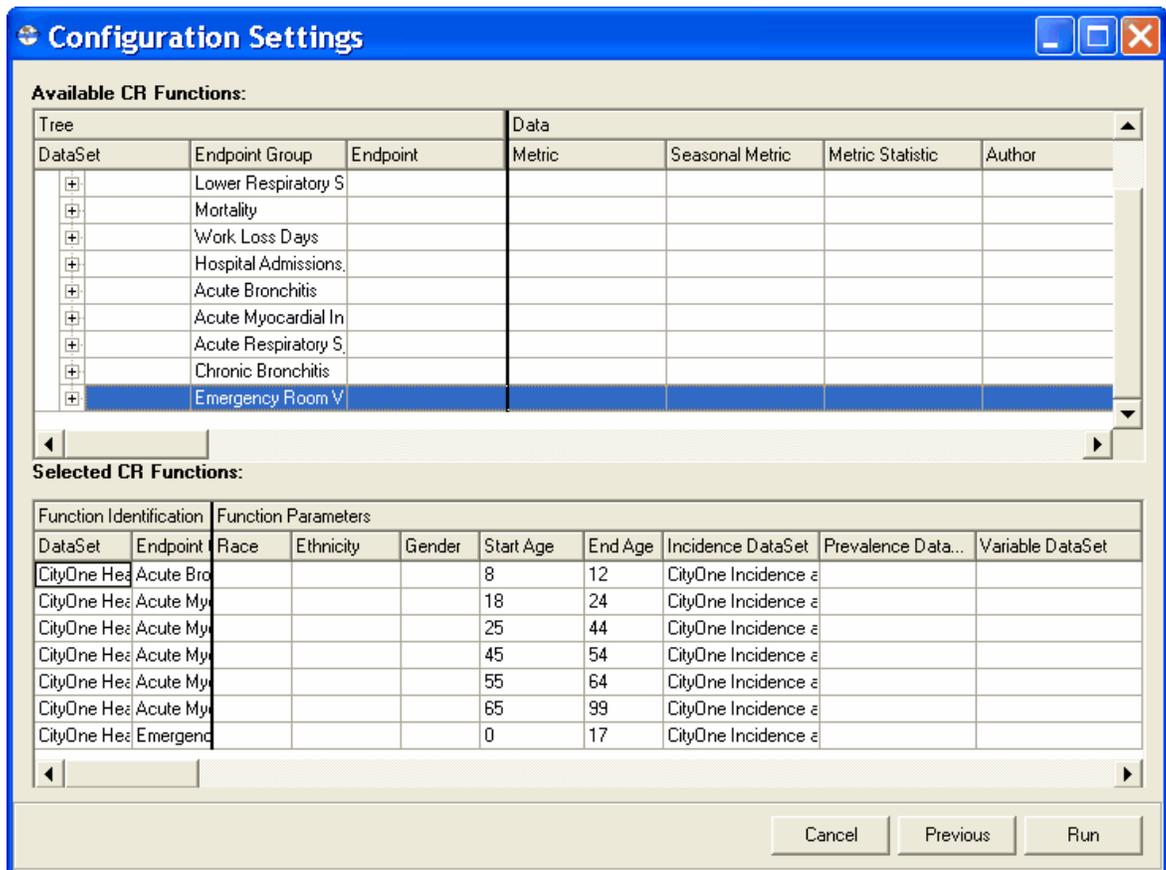
BenMAP will then prompt you to save your file. Click **Save**. Browse to the *Configurations* subfolder within the *BenMAP* directory and save the file as: *PM25 Example Configuration.cfg* (you do not need to include the ".cfg" extension).

When you have saved the configuration file, click **OK** to run the configuration.

BenMAP will prompt you to "Save Configuration Results to File". Browse to the *Configuration Results* subfolder within the BenMAP directory and save the file as: PM2.5 *CityOne County 25 Pct Rollback 2003 VNA Example.cfgr* (you do not need to include the ".cfgr" extension)

Once you have entered the filename, BenMAP will begin calculating the change in incidence for the set of health impact functions you have selected. The run may take a few minutes to finish; a progress bar will let you know how it is proceeding. When BenMAP is finished running your configuration, it will return to the main BenMAP screen.

A.2.5 Step 5. Specify Aggregation, Pooling and Valuation

This step allows you to take the incidence results that BenMAP just produced and place an economic valuation on them. Although not covered in this tutorial, this is also where you can select the geographic level of aggregation and combine individual incidence results

into pooling groups.

From the main screen, click on the **Aggregation, Pooling and Valuation** button. This will bring up a menu screen with two choices: *Create New Configuration for Aggregation, Pooling and Valuation*, or *Open Existing Configuration for Aggregation, Pooling and Valuation (*.apv file)*.
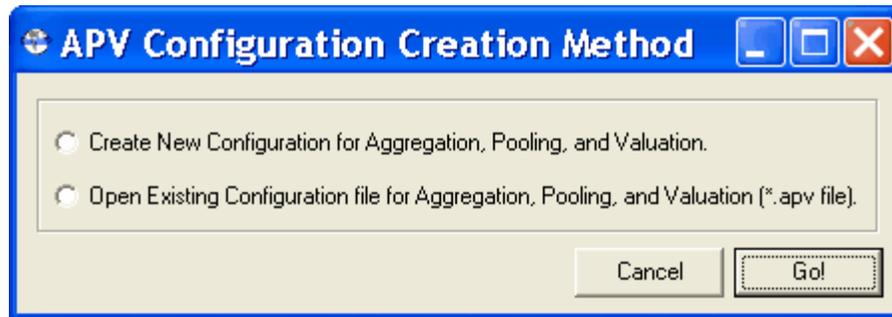
Select *Create New Configuration for Aggregation, Pooling and Valuation* and click on **Go!**.

BenMAP will prompt you to open a **Configuration Results File**. Browse to the *Configuration Results* subfolder and select *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example.cfgr*. Then click on **Open**.
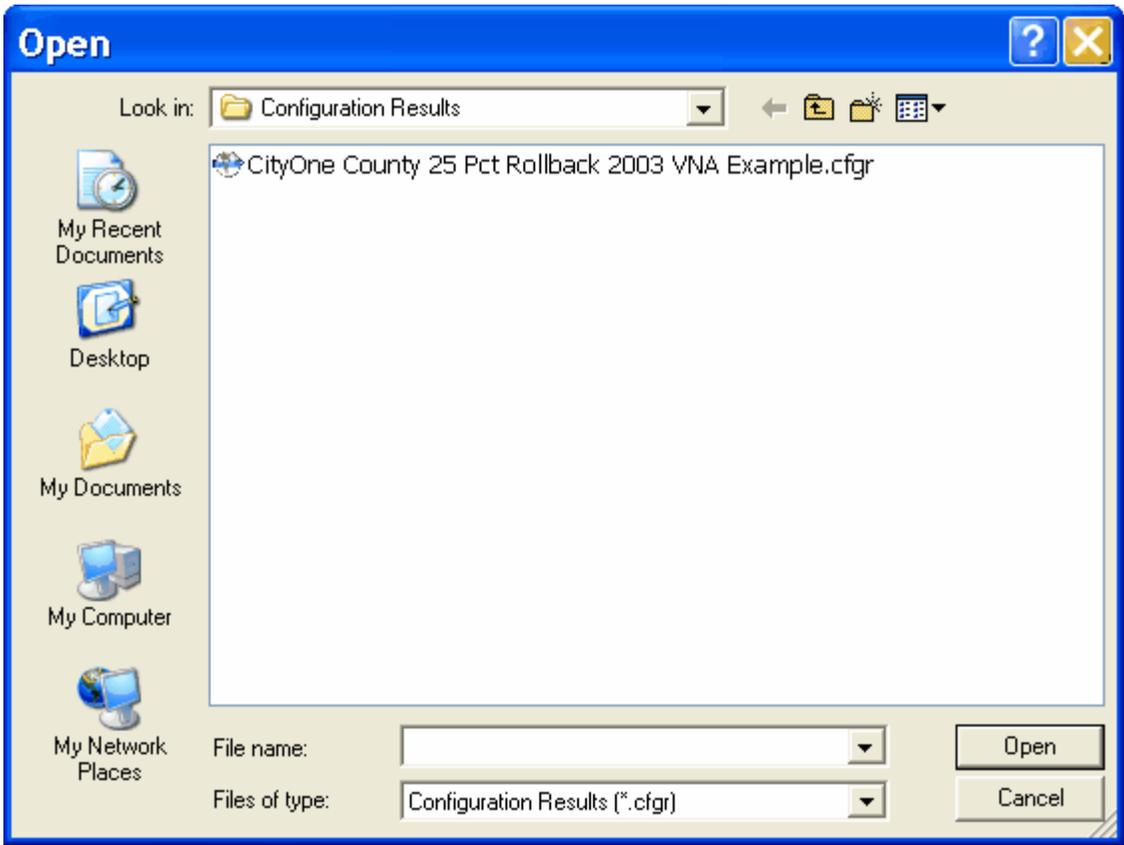
BenMAP will then open the **Incidence Pooling and Aggregation** window with the results from running your configuration. You should see a window that looks like the following:
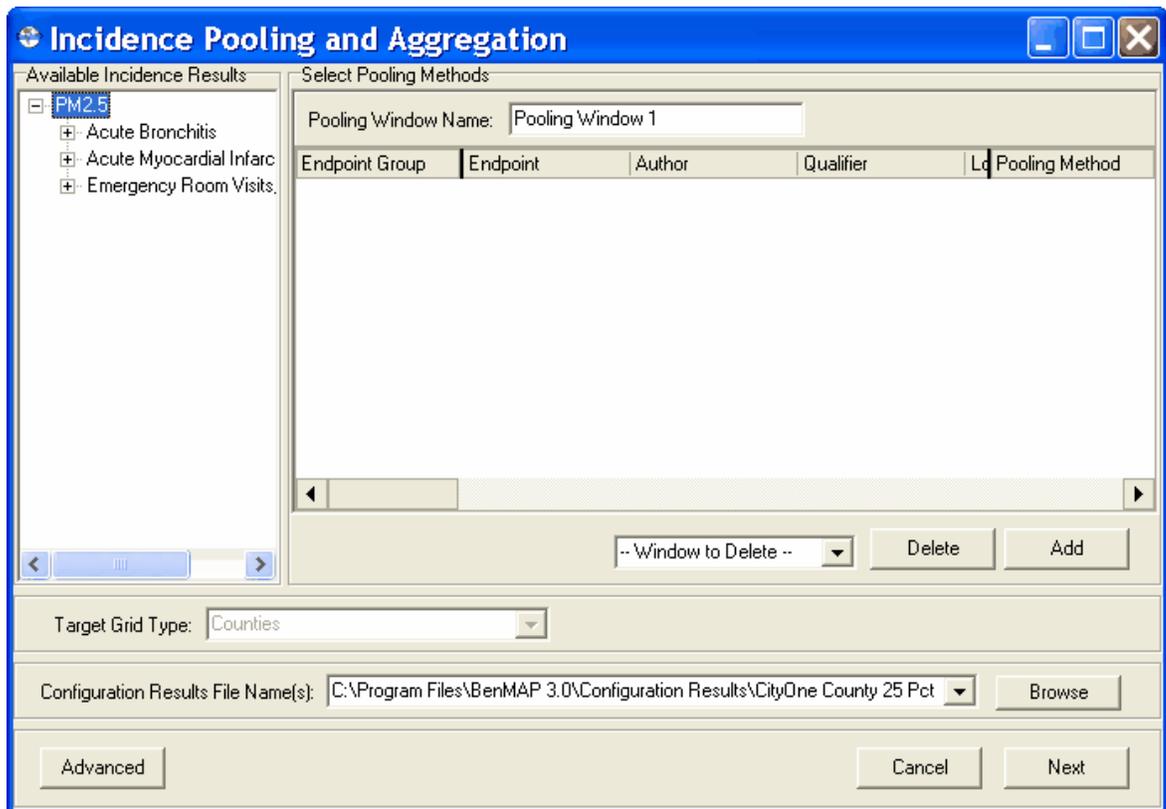
Click on each of the results groups (acute bronchitis, acute myocardial infarction, and emergency room visits) and drag them to the right panel.

For this example, we are not pooling any of the incidence results (although we will pool valuations in the next window), so just click on **Next** at the bottom of the window.

This will take you to the **Select Valuation Methods, Pooling, and Aggregation** window.
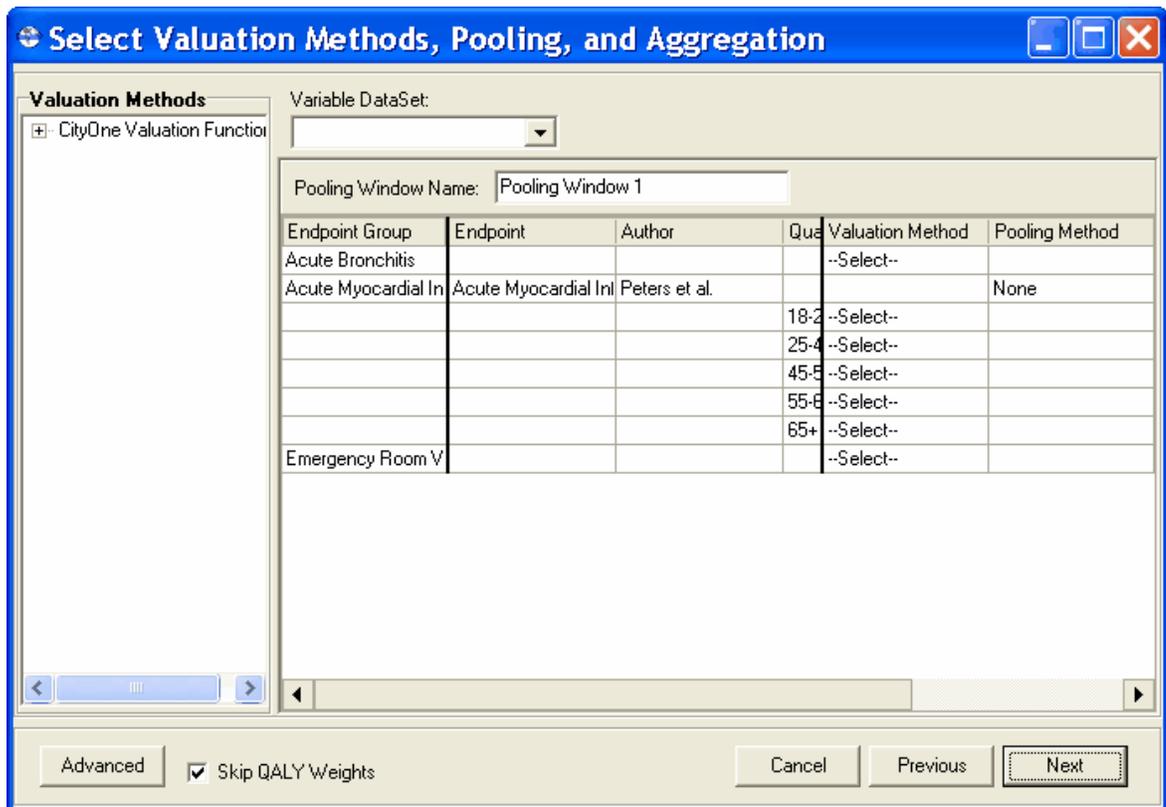
A) Select a value for acute bronchitis

To select a valuation method for acute bronchitis, drill down the *Acute Bronchitis* valuation group until you see individual valuation methods. Click on the *WTP: 6 day illness, CV studies | 0-17* method and drag it onto the *Acute Bronchitis* endpoint group in the right hand panel. You should see the method appear under acute bronchitis in the **Valuation Method** column in the right hand panel.

B) Select values for acute myocardial infarctions (heart attacks)

To select valuation methods for acute myocardial infarctions, drill down the AMI valuation group until you see a (long) list of individual valuation methods. You might find it easier to expand the column width of the Valuation Methods column (drag the right hand edge of the column to the right to make it wider). We will be working with the valuation estimates from two studies, *Wittels* and *Russell*. For each of the studies, there are a number of age specific valuations. There are also two different discount rates (the discount rate is the rate at which future medical costs are discounted to the present). Drag the age-specific valuation estimates from *Wittels* for the 3 percent discount rate (COI, 5 yrs med, 5 yrs wages, 3% DR, Wittels (1990) | age) to each matching age-specific line in the right hand panel (**Pooling Window 1**). You may have to scroll over in the right hand panel to see the **Low Age** column.

Note that you will need to drag some age-specific valuation estimates to multiple lines in

the pooling window, since there is not a perfect match between the available age-specific valuation estimates and the age groups for which the incidence of heart attacks was estimated. For example, you will have to drag the valuation estimate for the *25 to 44* age group to both the *25 to 35* age group and the *35 to 45* age group in the pooling window.

Now repeat this process using the *Russell* 3 percent discount rate valuation estimates. When you are finished, you should have two valuation estimates for each AMI age group, and your pooling window should look like the one below.
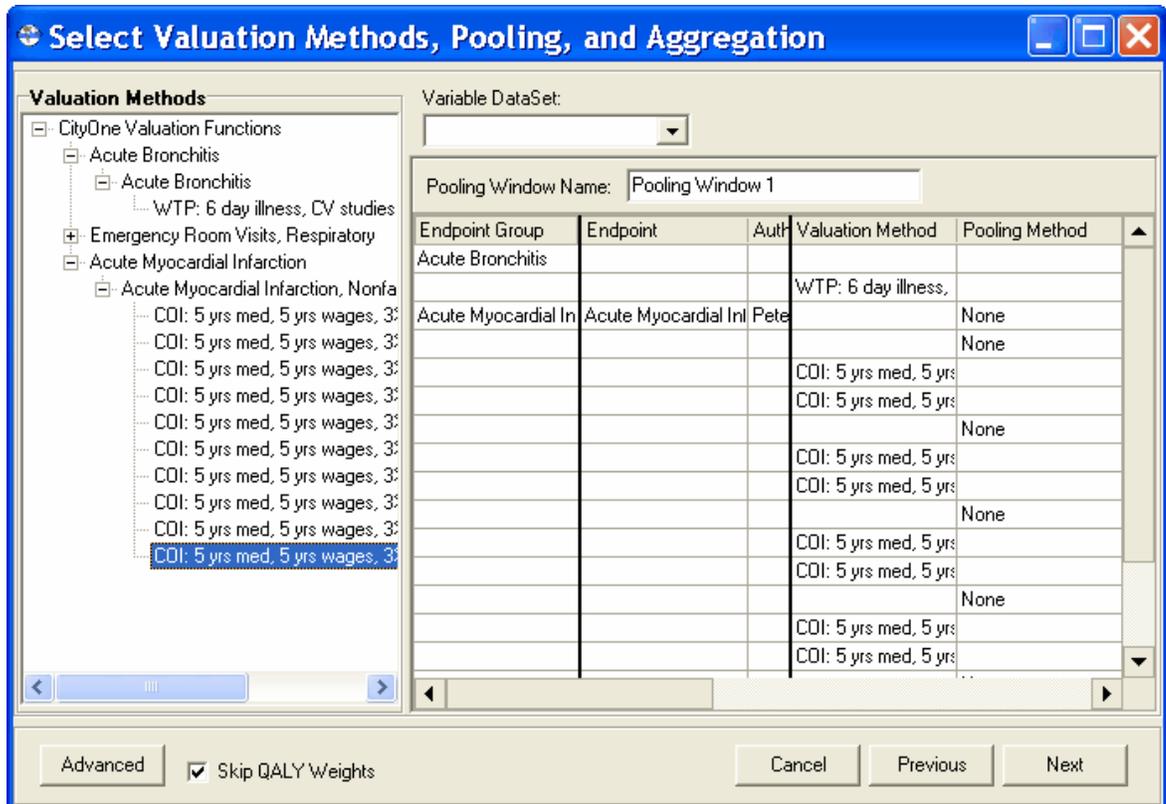
Now you can pool the valuation results for heart attacks in each age group using the unit values from both *Wittels* and *Rusell*. In order to do so, you must select a pooling method.

BenMAP lets you select from several different pooling methods. For this example, you will be using subjective weights. In other applications, you may wish to use fixed or random effects weights.

To set the pooling method for each age group result, click on the **Pooling Method** field in the row ABOVE each pair of valuation methods (where it says *None*) and use the drop down menu to select *Subjective Weights*. You must repeat this for EACH age group in order for pooling to take place over all age groups.

In addition to pooling the results over the two valuation methods, we also need to aggregate the results into a total estimate across age groups. In order to do so, in the row

with **Endpoint Group** (**Endpoint Group** = *Acute Myocardial Infarction*) click in the **Pooling Method** field and select Sum (*Dependent*) from the drop down menu.

Your screen should look like the following:
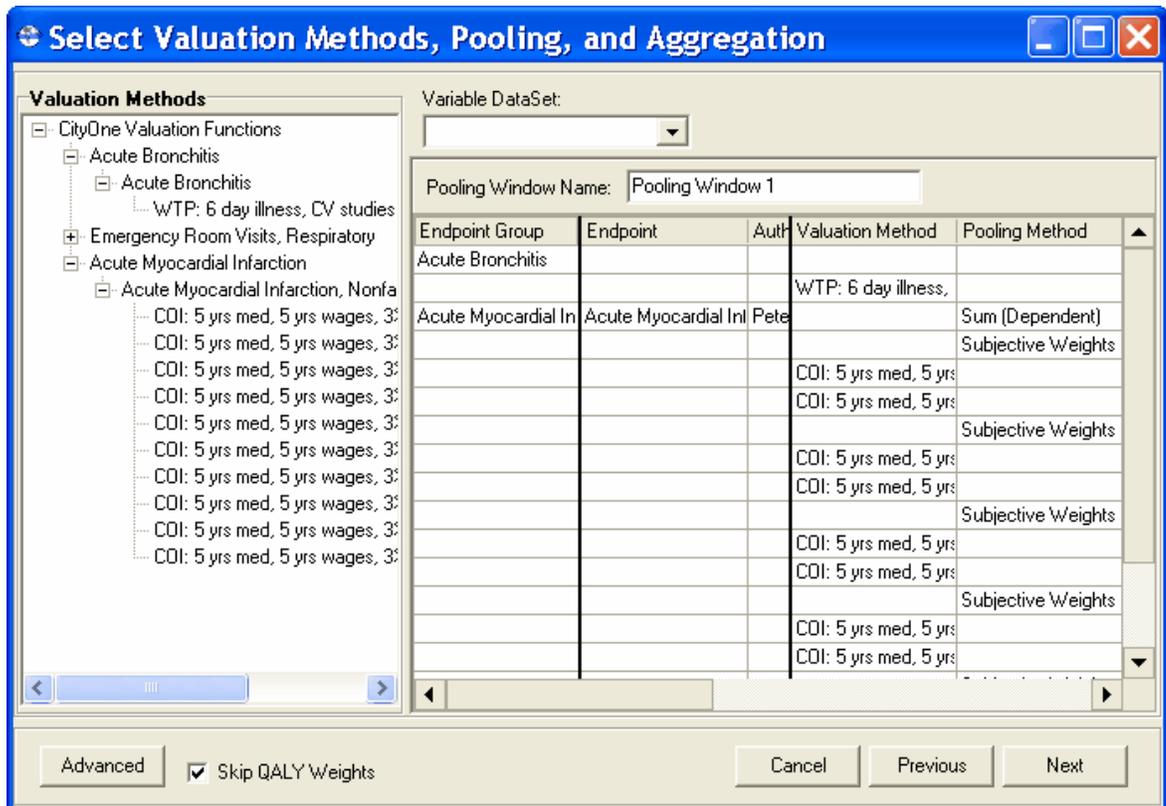
This pooling configuration for acute myocardial infarctions will assign a starting set of equal weights to each valuation method for the set of five age groups, and then create an overall estimate of acute myocardial infarctions by summing the age-specific pooled estimates, treating the distributions for each age group as dependent (i.e. a draw from the 5th percentile of the 45 to 54 age group will be added to the draw from the 5th percentile of the 55 to 64 age group and so on).

C) Select values for asthma emergency room visits

To select values for asthma ER visits, drill down the *Emergency Room Visits, Respiratory* heading to the *Emergency Room Visits, Asthma*, and then to the valuation approach Standford et al, 1999 | 0 - Max. Drag this to the *Emergency Room Visits* entry in the right hand panel.

D) Choose variable dataset

In the **Variable Dataset** drop-down menu, choose *CityOne Variables*. (BenMAP requires that there be a variable dataset be chosen before going to the next step.)

Your completed screen should look like the one below.
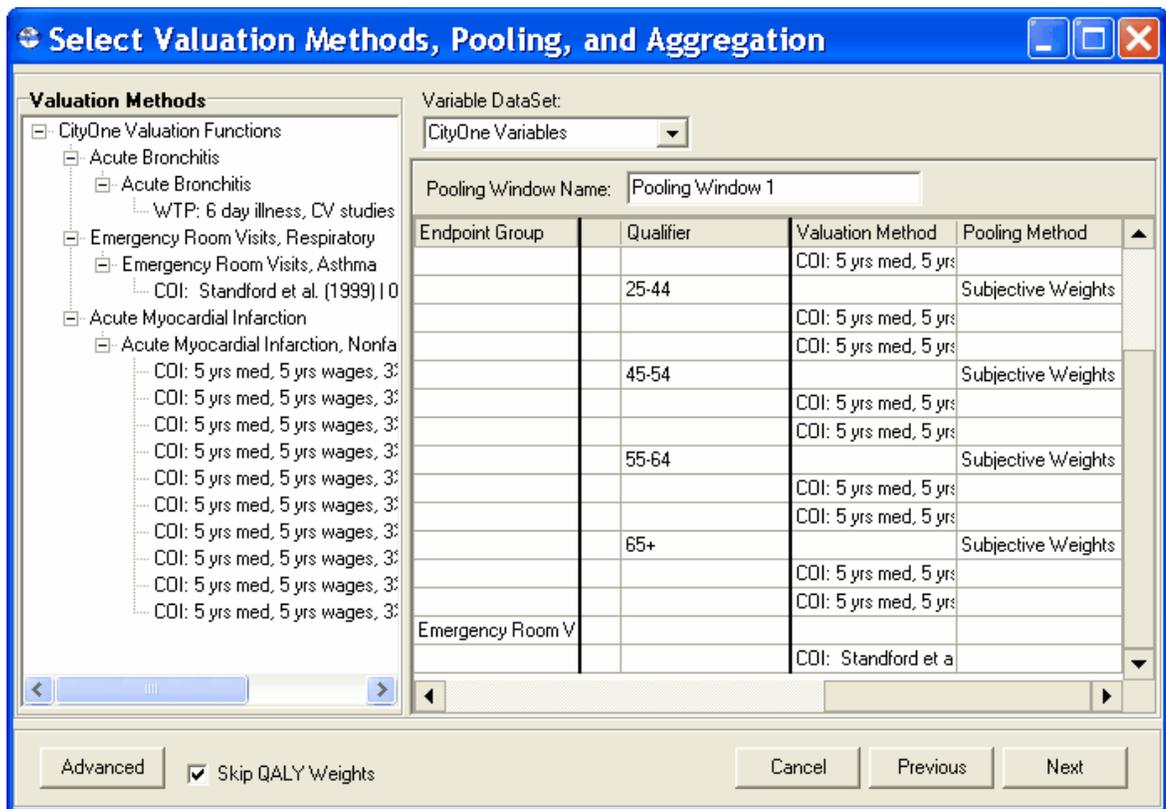
E) Entering subjective weights

Once you have completed this step, click on **Run**. BenMAP will now bring up a window to allow you to enter subjective weights.

BenMAP assigns a default equal weight to each selected valuation method. You can change these weights by clicking in the weight cells. However, for this exercise, you should leave them at 0.5 for each study. Click on **OK** at the bottom of the screen. You should see a save dialog box. Click on Save to save your APV configuration. Save the file as *PM25 Direct example APV*.
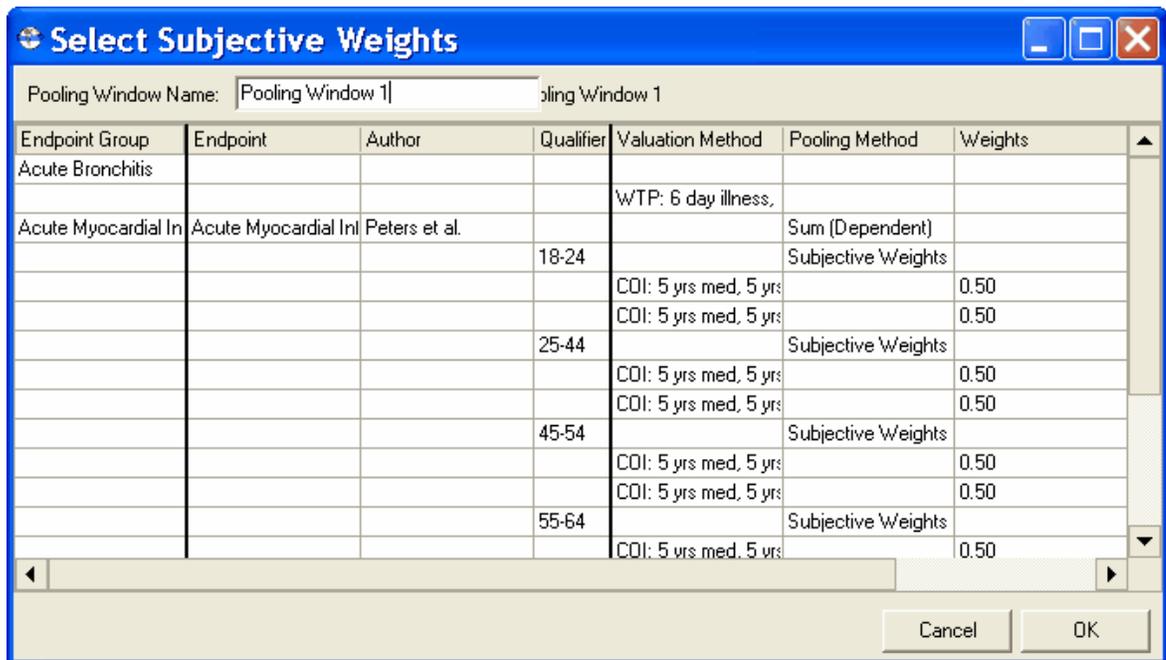
Click on **OK** to start the pooling and aggregation. First you will be prompted to enter a filename for the aggregation, pooling, and valuation configuration file that you just created. Enter *PM25 Example Configuration.apv* and click **Save**.

Then you will be prompted to enter a filename for the aggregation, pooling, and valuation results. Enter *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example.apvr* and click **Save**. When the progress bar disappears, you will be returned to the main BenMAP screen.

A.2.6   Step 6. Generate Reports

You may view your results within BenMAP, either in the preview window in the **Create Reports** button or through the mapping functions. Alternatively, you may export the results to a comma separated values file (*.csv), or a shapefile (*.shp) which can be viewed in a GIS program such as ESRI's ArcView product.

A) Generate a **Pooled Incidence Results** report.

A **Pooled Incidence Results** report contains the incidence results you previously generated, using the aggregation level and pooling that you specified in the **Incidence Pooling and Aggregation** window. Previously, in Step 5, we did not specify any pooling of incidence results (although valuations were pooled), so in this case the **Pooled Incidence Results** report will look just like the **Aggregated Incidence Results** report. If some incidence results had been pooled, the two reports would be different.

Click on the **Create Reports** button from the main BenMAP screen. This will bring up the **Select Result Type** window.
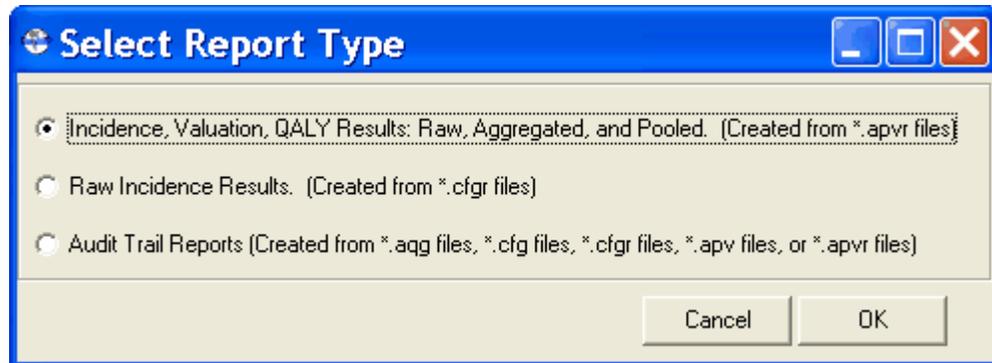
Select *Incidence and Valuation Results: Raw, Aggregated and Pooled*. Click **OK**. This will bring up a window where you can select a results file. Chose the *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example.apvr* file, and click **Open**. This will bring up the **Choose a Result Type** window.
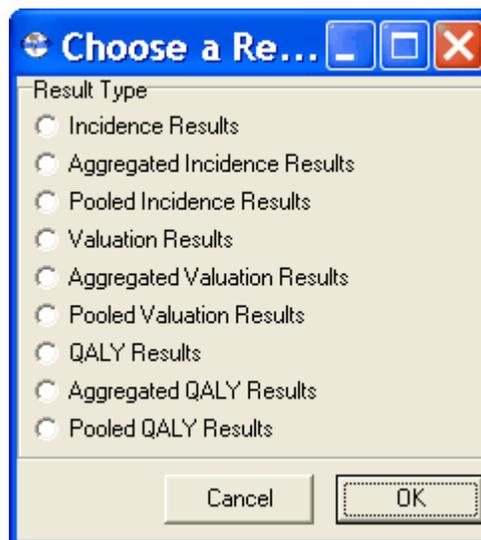
In the **Choose a Result Type** window, choose *Pooled Incidence Results*. Then click **OK**. This will bring up the **APV Configuration Results Report**, where you can customize your report display and select the fields you want to see in the report. In the **Pooled C-R Function Fields** box, check off **Endpoint Group** and **Qualifier**.
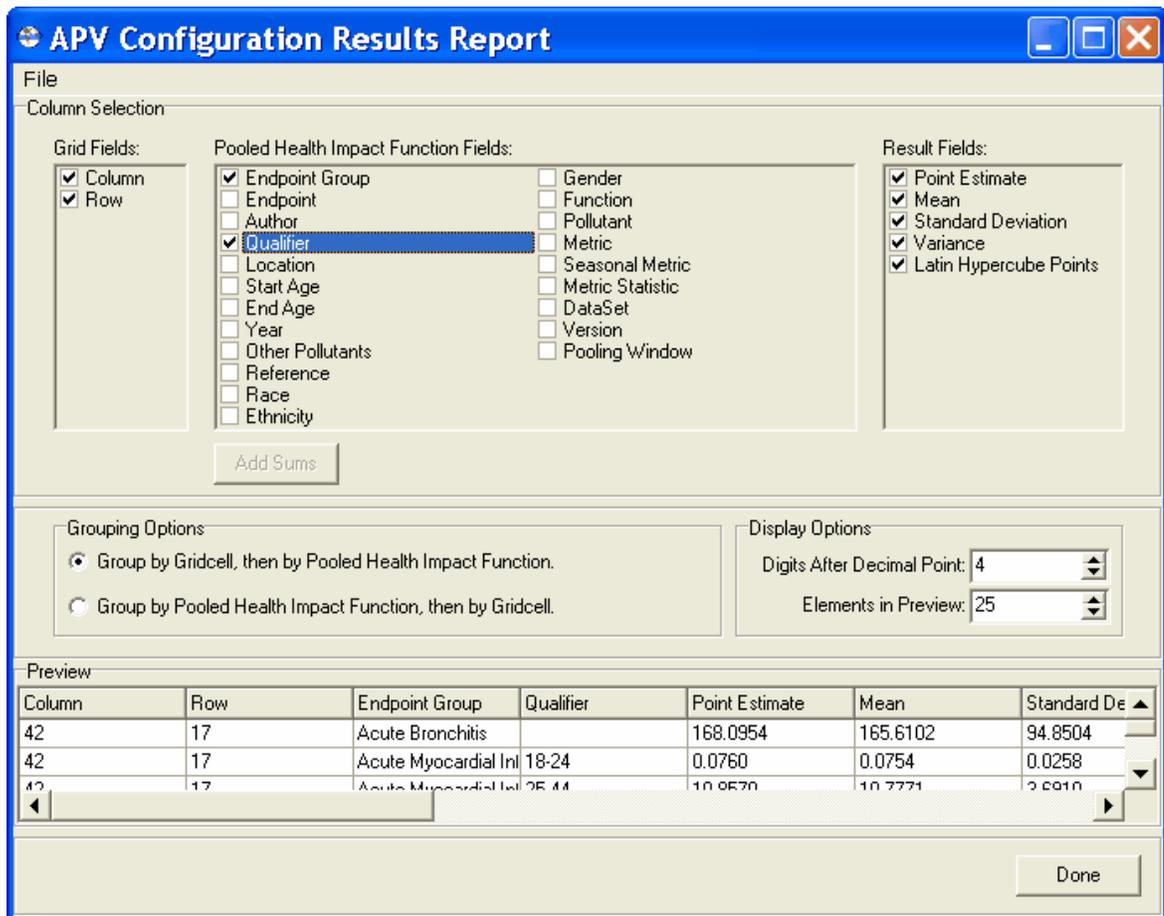
When your window looks like the window above, then go to the **File** menu and choose **Save**. In the **Save As** window that appears, type in the file name and browse to the location where you want to store the exported file. The **Reports** subfolder is a good location to keep exported reports. Type in the name, *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example Health.csv* in the box and click **Save**. You can now open the report in another application, such as a spreadsheet or database program.

B) Generate a **Pooled Valuation Results** report.

This report is similar to the **Pooled Incidence Results** report, and uses the valuation pooling you previously specified. Click on the **Create Reports** button from the main BenMAP screen. This will bring up the **Select Report Type** window. Select *Incidence and Valuation Results: Raw, Aggregated and Pooled*. Click **OK**. This will bring up a window where you can select a results file. Chose the *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example.apvr* file, and click **Open**. This will bring up the **Choose a Result Type** window.

In the **Choose a Result Type** window, choose *Pooled Valuation Results*. Then click **OK**. This will bring up the **APV Configuration Results Report** window, where you can customize your report display and select the fields you want to see in the report. In the **Pooled Valuation Methods Fields** box, check off **Endpoint Group**.
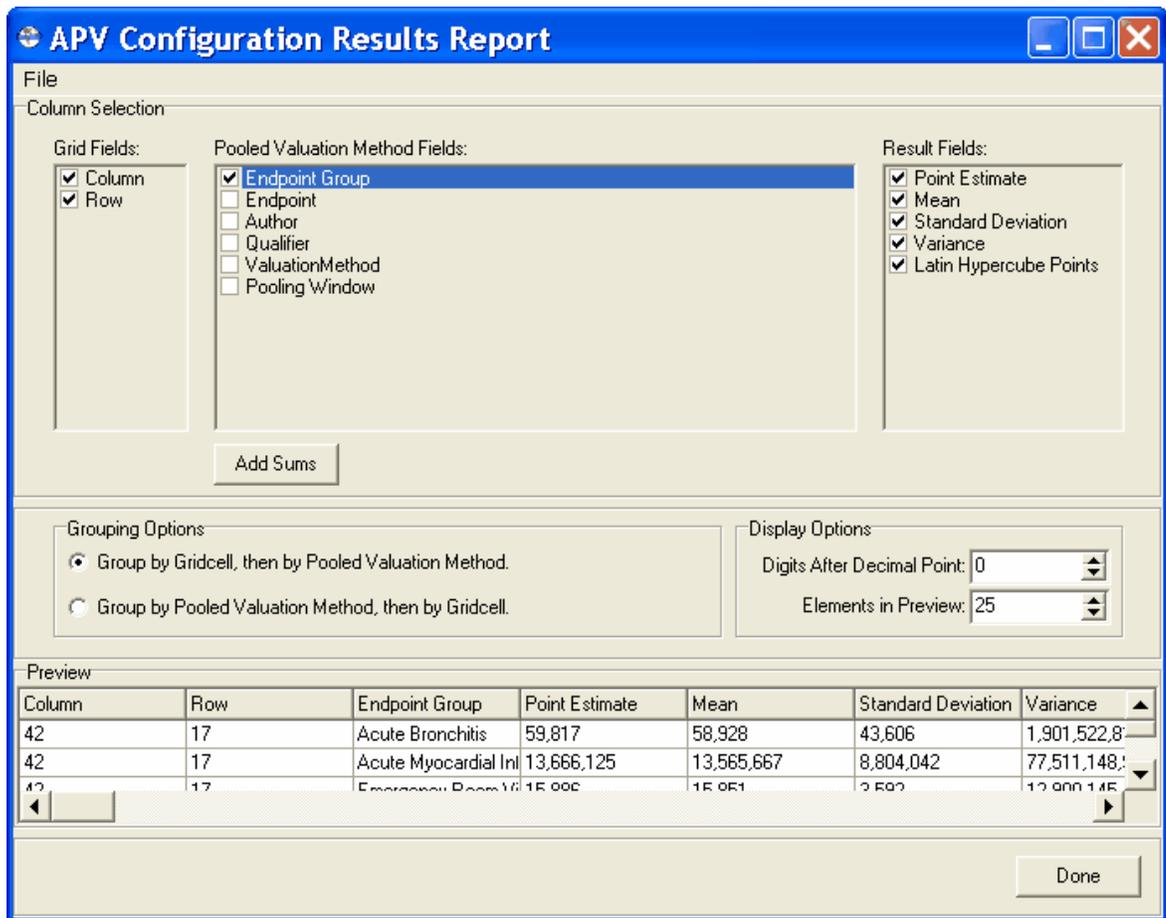
When your window looks like the window above, then go to the **File** menu and choose **Save**.

In the **Save As** window that appears, type in the file name and browse to the location where you want to store the exported file. The **Reports** subfolder is a good location to keep exported reports. Type in the name, *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example Valuation.csv* in the box and click **Save**. You can now open the report in another application, such as a spreadsheet or database program.

A.2.7 Step 7. View Your Reports

BenMAP generates comma separated values files (*.csv) that can be read by various spreadsheet and database applications, such as Microsoft Excel.

A.2.8 Step 8. Map Your Results

You can also map any of the results that you have generated so far. This includes the air quality grids, population data, incidence results, and valuation results. In this example, we will look at the air quality grid for the base scenario, and view our incidence results. For more information on these and other mapping functions, see Chapter 9.

To use the BenMAP mapping functionality, go to the **Tools** menu and choose *GIS / Mapping*. The BenMAP GIS window will appear, with buttons at the top for managing files and navigating the map.
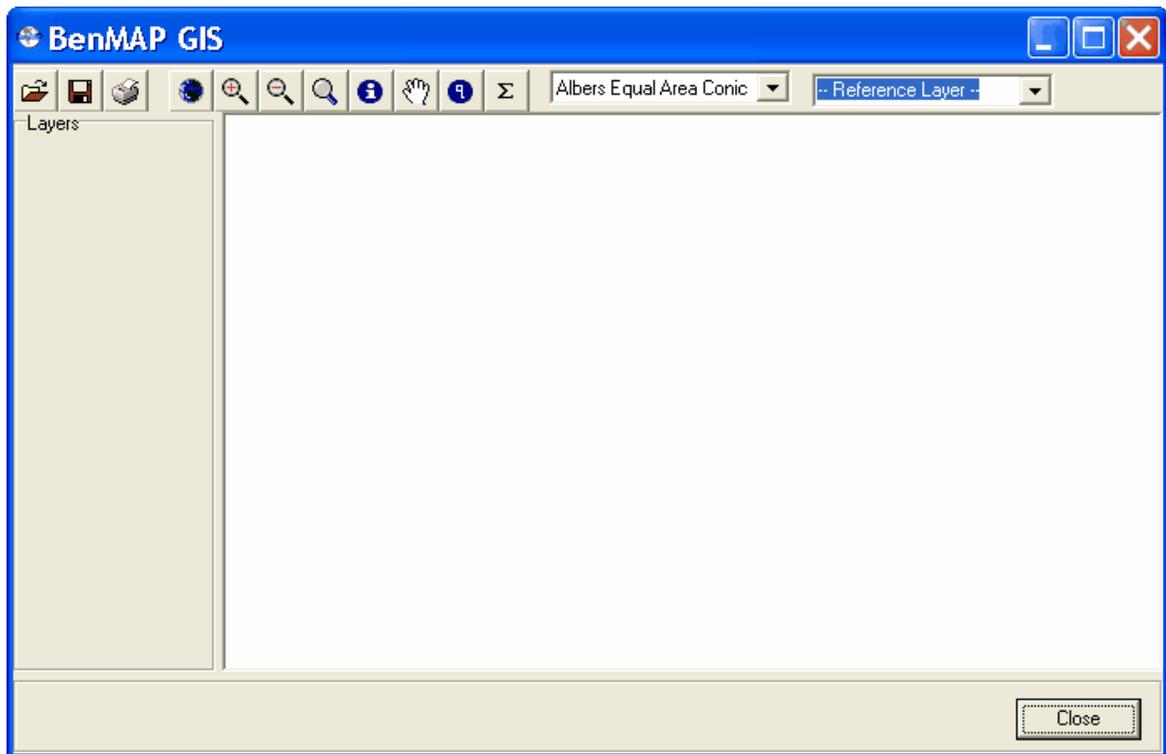
To see the name of each button, simply hold the cursor over it. Click on the **Open a file** button, and select *Air Quality Grid* from the drop-down menu. Browse to the file *PM2.5 CityOne County 25 Pct Rollback 2003 VNA.aqg* file and click **Open**.

The name of the file will appear in the left-hand panel under **Layers**. Double-click on the name and a small box will appear with **Display Options** for viewing this layer. Here you can select the variable contained in the layer (file) that you want to view. In the air quality grid, the variables that are available are the Quarterly Mean and the Daily Mean *(D24HourMean)*. Select *D24HourMean* for the annual mean of the Daily Mean in the **Variable**. In this box, you can also change the colors in the map display, and the maximum and minimum values displayed.
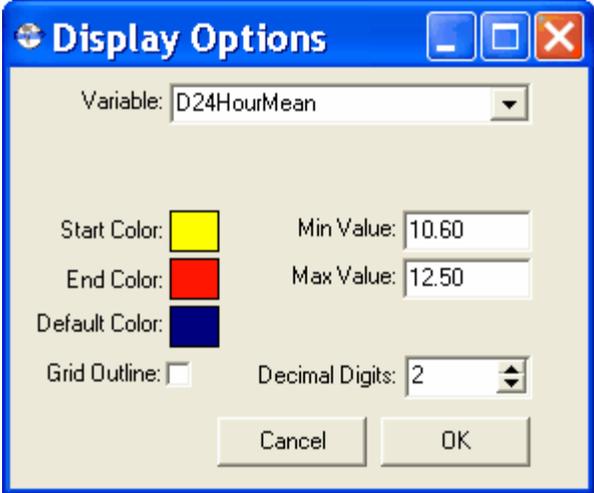
When done choosing your display options, click **OK**. You should see a map like the one below.
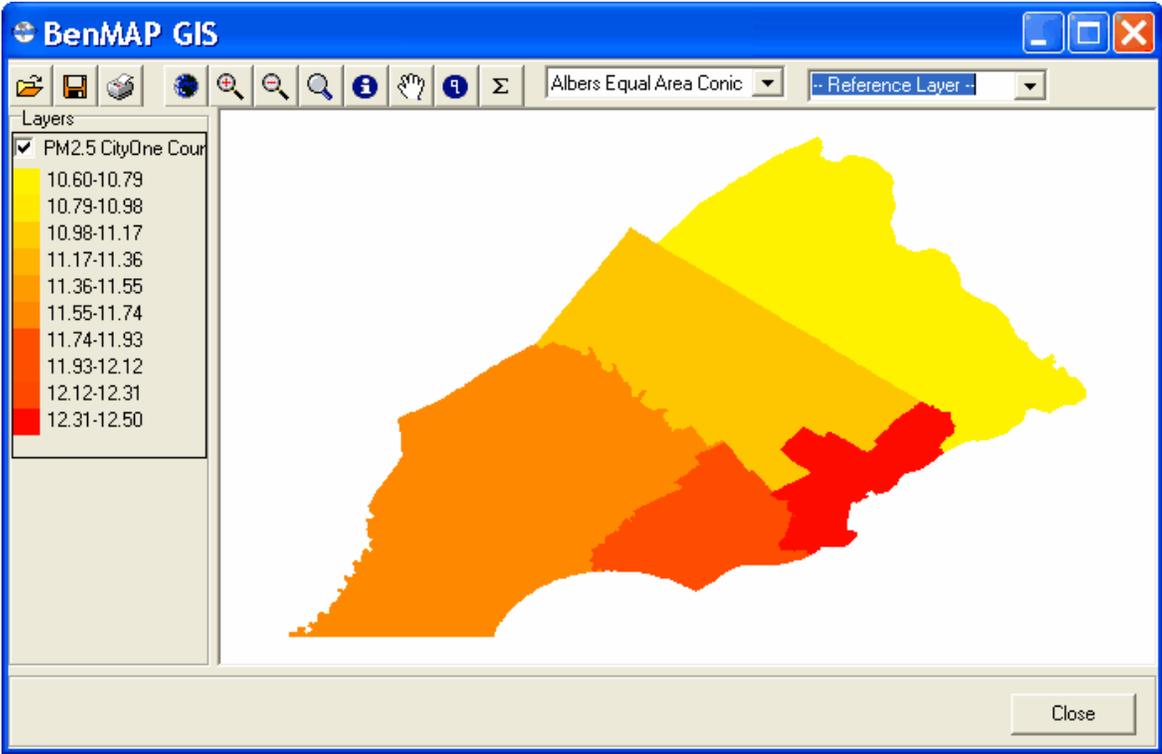
To see tract outlines, select *County* from the **Reference Layer** drop-down menu at the top of the screen. You may also use the other reference layers: *Metropolitan Area* and *Tract*. However, since the results have been calculated at the county level in this example, the county reference is generally most appropriate.
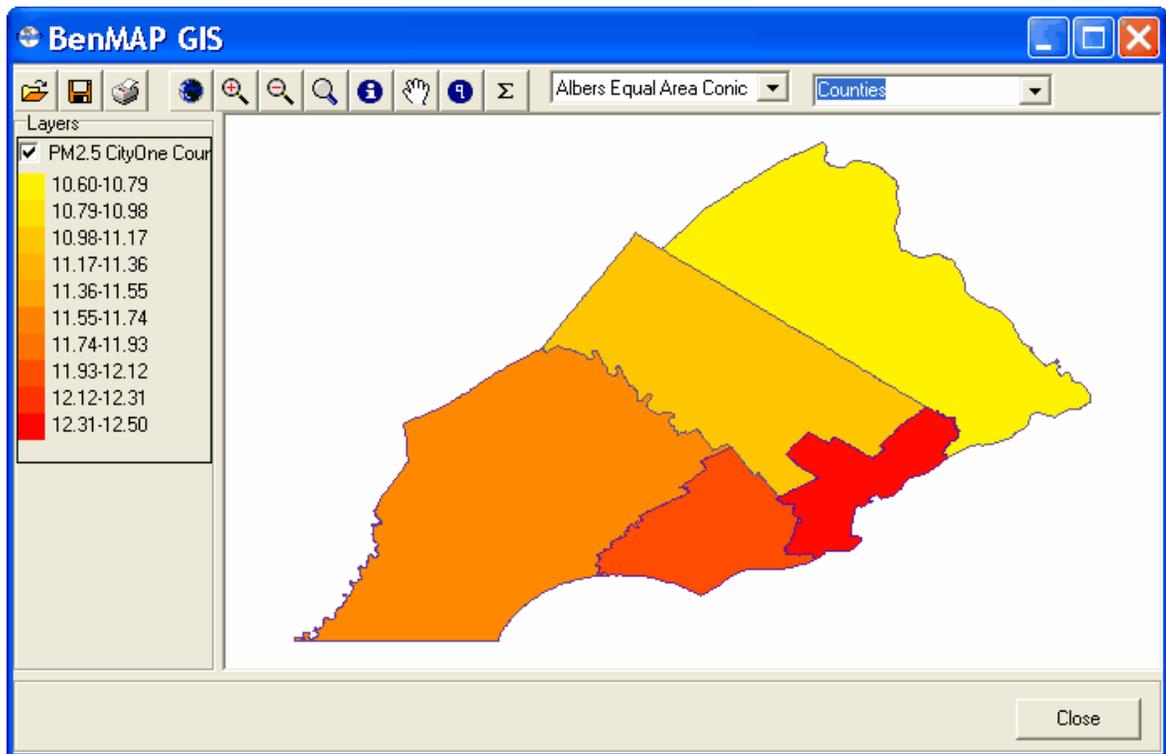
Now you can look at a geographical display of the incidence results you created for cases of bronchitis, acute myocardial infarctions, and emergency room visits. Click on the **Open a file** button at the top of the screen and select *APV Configuration Results*, then *Incidence Results*. In the next window, select *PM2.5 CityOne County 25 Pct Rollback 2003 VNA Example.apvr*, then click **Open**. BenMAP will load your incidence results and display them in a table. Because GIS programs can typically only accommodate field names that are 10 characters or less, there is a new column at the end of the table labeled **Gis Field Name**. Here you can name your variables, as shown in the table below.

When you are satisfied with the variable names, click **OK**. The new layer will show in the BenMAP GIS window on top of the first layer. If the previous layer is still checked, then it will appear, but underneath the new layer. Uncheck the box next to the bottom (previous) layer to hide it. Your screen should look like the one below.
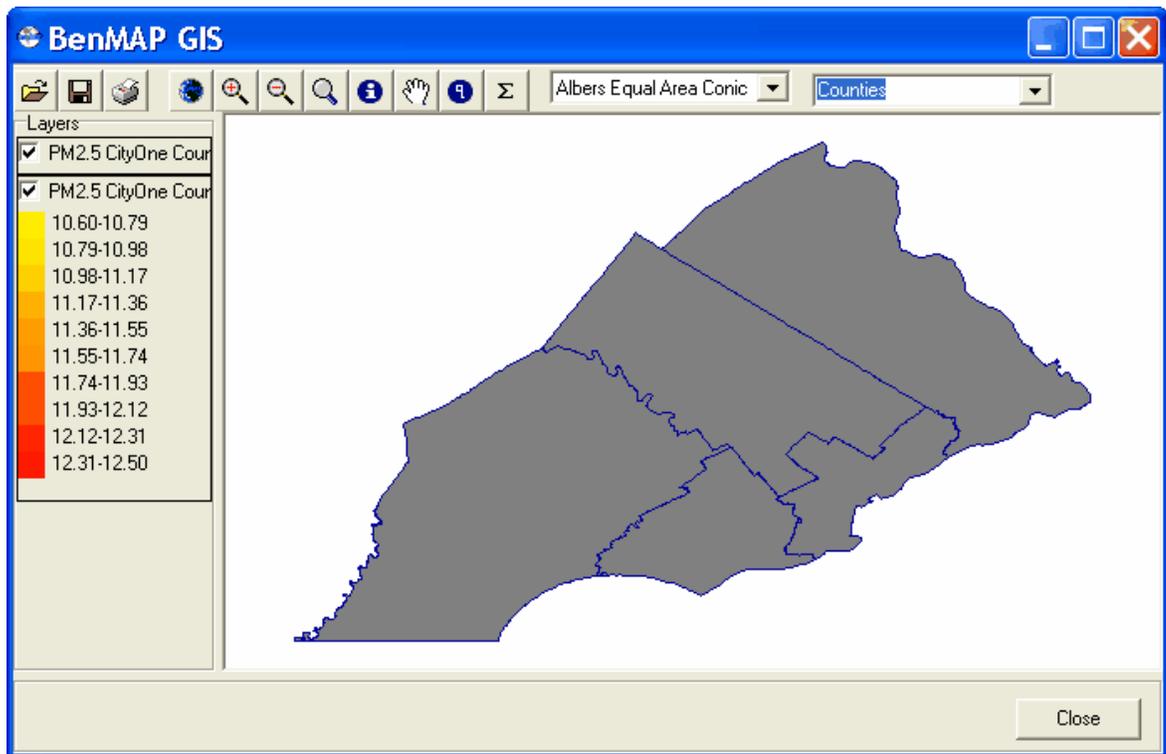
Like the previous layer, double click on the name to bring up the **Display Options** box.
Under Variable you will see a list of the variable names you defined in the previous step.
Select *Myo65up*, uncheck the **Grid Outline** box, and click **OK**. The viewer will now
display the annual increase in the number of acute myocardial infarctions for people ages
65 and up, as calculated between the base and control scenarios. You can use the Display
Options to select other variables to view or change how the values are displayed.
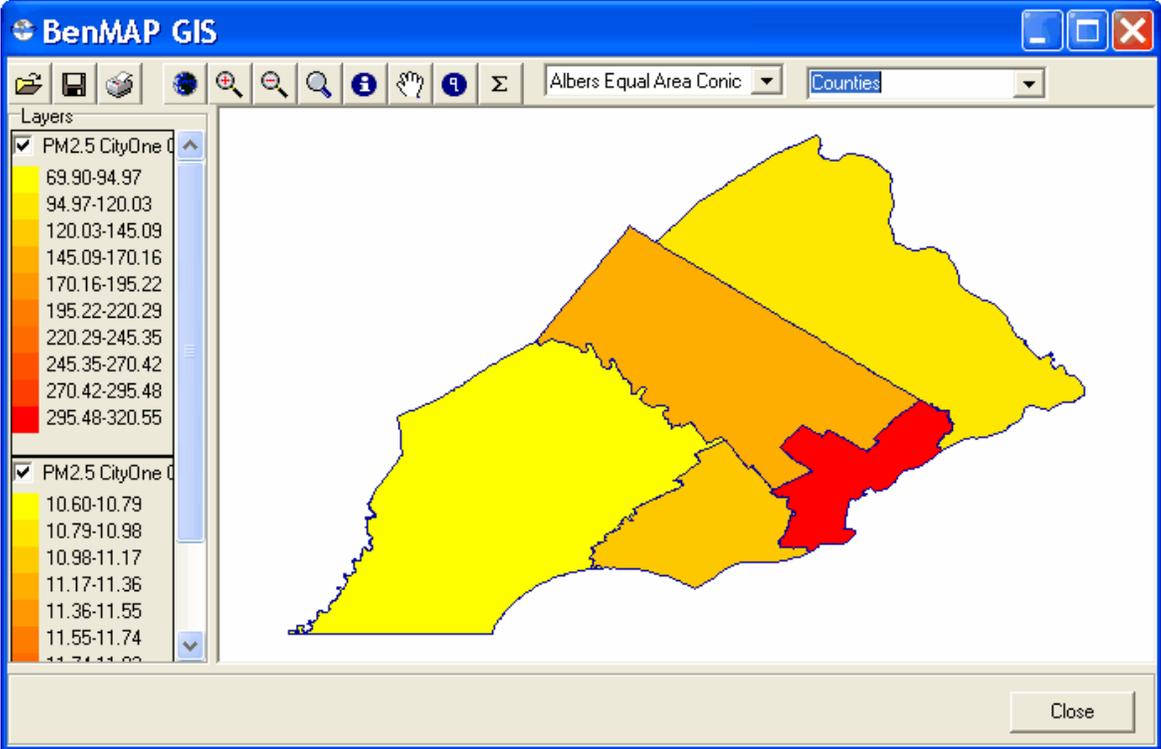
Appendix B: Monitor Rollback Algorithms

This Appendix details the rollback procedures that you can perform on monitor data. The rollback procedure is a quick way to determine the monitor levels that would exist under various kinds of changes that you can specify. This includes three basic types of rollbacks: Percentage, Increment, and Rollback to Standard.

Once a set of monitors has been selected, the user may define one or more non-overlapping rollback **regions**. A region is simply a set of states with an associated set of rollback parameter values. Three rollback **types** are available:

Percentage Rollback. Monitor values are reduced the same percentage.

Incremental Rollback. Monitor values are reduced by the same fixed increment.

Rollback to a Standard. Monitor values are reduced so that attainment of a specified standard is reached.

Each of these rollback types has different rollback parameters associated with it.

B.1   Percentage Rollback

Percentage Rollback involves setting only two parameters - a **percentage** and a **background level**. The rollback procedure is similarly straightforward - each observation at each monitor in the region has the portion of its value which is above **background level** reduced by **percentage**.

Example: Background Level: 35; Percentage: 25

Initial Observations at a monitor in rollback region:

20	20	25	59	35	51	83	35	30	67	87	79	63	35	35

If we select the background level of 35, we first calculate the portion of each observation that is above background level, that is, we subtract the background level from the initial observation level. Observations below background level are given a value of 0.

Observation portions above **background level**:

0	0	0	24	0	16	48	0	0	32	52	44	28	0	0

When we apply the rollback percentage, each observation portion gets reduced by 25%.

Reduced portions above background level:

| 0 | 0 | 0 | 18 | 0 | 12 | 36 | 0 | 0 | 24 | 39 | 33 | 21 | 0 | 0 |

Then, each reduced portion is added to the background level of 35. Zero values are replaced by the initial observations.

Reduced Observations:

| 20 | 20 | 25 | 53 | 35 | 47 | 71 | 35 | 30 | 59 | 74 | 68 | 56 | 35 | 35 |

B.2 Incremental Rollback

Incremental Rollback similarly involves setting only two parameters - an **increment** and a **background level**. The rollback procedure is quite similar to the percentage rollback procedure - each observation at each monitor in the region has the portion of its value which is above **background level** reduced by **increment**. The reduced values are not allowed to become negative, however - that is, they are truncated at zero.

Example: Background Level: 35; Increment: 25

Initial Observations:

| 20 | 20 | 25 | 59 | 35 | 51 | 83 | 35 | 30 | 67 | 87 | 79 | 63 | 35 | 35 |

Observation portions above background level:

| 0 | 0 | 0 | 24 | 0 | 16 | 48 | 0 | 0 | 32 | 52 | 44 | 28 | 0 | 0 |

Reduced portions above background level:

| 0 | 0 | 0 | 0 | 0 | 0 | 23 | 0 | 0 | 7 | 27 | 19 | 3 | 0 | 0 |

Reduced Observations:

| 20 | 20 | 25 | 35 | 35 | 35 | 58 | 35 | 30 | 42 | 62 | 54 | 38 | 35 | 35 |

B.3 Rollback to a Standard

Rollback to a Standard has two groups of parameters - those associated with the **Attainment Test**, which determines whether a monitor is in attainment (meets the standard), and those associated with the **Rollback Methods**, which are used to bring out of

attainment monitors into attainment.

The **Attainment Test** parameters are **Metric**, **Ordinality**, and **Standard**. A monitor is considered in attainment if the nth highest value of the metric specified by **Metric** is at or below the value specified by **Standard**, where n is the value specified by **Ordinality**. For example, if **Metric** is TwentyFourHourDailyAverage, **Ordinality** is two, and **Standard** is eighty five, a monitor will be considered in attainment if the second highest value of TwentyFourHourDailyAverage is at or below eighty five.

Supported metrics for pollutants with hourly observations (Ozone) include FiveHourDailyAverage, EightHourDailyAverage, TwelveHourDailyAverage, TwentyFourHourDailyAverage, OneHourDailyMax, and EightHourDailyMax. Supported metrics for pollutants with daily observations (PM10, PM2.5) include TwentyFourHourDailyAverage and AnnualAverage. For Annual Average, **Ordinality** does not apply, since there is only a single metric value to work with.

The **Rollback Method** parameters are **Interday Rollback Method**, **Interday Background Level**, **Intraday Rollback Method**, and **Interday Background Level**. These four parameters determine the rollback procedures used to bring out of attainment monitors into attainment. The **Interday Rollback Method** and **Background Level** are used to generate target values for the metric specified by the **Attainment Test**. The **Intraday Rollback Method** and **Background Level** are used to adjust hourly observations to meet the target metric values generated in the previous step. As such, the **Intraday Rollback Method** and **Background Level** are used only for pollutants with hourly observations (ozone).

B.3.1 Interday Rollback - Generating Target Metric Values

Because standards are defined on metrics, not directly on observations, the first step in rolling back out of attainment monitors is generating target metric values. There are four supported rollback methods for Interday Rollbacks - Percentage, Incremental, Peak Shaving, and Quadratic. Each of these rollback methods requires some preprocessing of the initial monitor metric values. We will discuss this preprocessing first, and then go through Percentage, Incremental, and Peak Shaving rollbacks in turn. Quadratic rollback is more complicated than these first three, and has its own section.

The Interday Background Level specifies the portion of each metric value which cannot be affected by human intervention - we call this portion the non-anthropogenic portion. Whatever portion is left over after subtracting out the background level is referred to as the anthropogenic portion. The anthropogenic portion of the initial monitor metric values is the only part which will be affected by the Interday Rollback Method.

BenMAP calculates an out of attainment value by determining the particular monitor metric value which caused the monitor to be out of attainment - this value is the nth highest value of the metric specified by the Attainment Test metric, where n is the Attainment Test ordinality. BenMAP then calculates an anthropogenic out of attainment value by subtracting the Interday Background Level from the out of attainment value.

BenMAP also calculates an anthropogenic standard by subtracting the Interday Background Level from the Attainment Test standard. Finally, BenMAP calculates a set of anthropogenic metric values and a set of non-anthropogenic metric values using the following procedure on each initial monitor metric value:

IF the metric value is less than or equal to the Interday Background Level,

non-anthropogenic metric value = metric value

anthropogenic metric value = 0

ELSE

non-anthropogenic metric value = Interday Background Level

anthropogenic metric value = metric value - Interday Background Level

B.3.1.1 Interday Rollback - Percentage

To generate target metric values using Percentage rollback, BenMAP calculates the percentage required to reduce the anthropogenic out of attainment value to exactly the anthropogenic standard. This percentage reduction is then applied to all of the anthropogenic metric values. Finally, these reduced anthropogenic metric values are added to the non-anthropogenic metric values to give the final target metric values.

Example:

Initial Metric Values:

30	35	50	100	80	44	67	88	90	70	50	30	55	90	80	85

Attainment Test: Highest value of metric <= 70

Interday Background Level: 40

Out of Attainment Value: 100

Anthropogenic Out of Attainment Value: 60 (= 100 - 40)

Anthropogenic Standard: 30 (= 70 - 40)

Percentage Reduction Required: 50% (=(60-30)/60)

Non-Anthropogenic Metric Values:

| 30 | 35 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 30 | 40 | 40 | 40 | 40 |

Anthropogenic Metric Values:

| 0 | 0 | 10 | 60 | 40 | 4 | 27 | 48 | 50 | 30 | 10 | 0 | 15 | 50 | 40 | 45 |

Reduced Anthropogenic Metric Values:

| 0 | 0 | 5 | 30 | 20 | 2 | 14 | 24 | 25 | 15 | 5 | 0 | 8 | 25 | 20 | 23 |

Target Metric Values:

| 30 | 35 | 45 | 70 | 60 | 42 | 54 | 64 | 65 | 55 | 45 | 30 | 48 | 65 | 60 | 63 |

B.3.1.2 Interday Rollback - Incremental

To generate target metric values using Incremental Rollback, BenMAP calculates the increment required to reduce the anthropogenic out of attainment value to exactly the anthropogenic standard. This incremental reduction is then applied to all of the anthropogenic metric values (but - they are not allowed to fall below zero). Finally, these reduced anthropogenic metric values are added to the non-anthropogenic metric values to give the final target metric values.

Example:

Initial Metric Values:

| 30 | 35 | 50 | 100 | 80 | 44 | 67 | 88 | 90 | 70 | 50 | 30 | 55 | 90 | 80 | 85 |

Attainment Test: Highest value of metric $<= 70$

Interday Background Level: 40

Interday Rollback Method: Incremental

Out of Attainment Value: 100

Anthropogenic Out of Attainment Value: 60

Anthropogenic Standard: 30 (=70 - 30)

Incremental Reduction Required: 30

Non-Anthropogenic Metric Values:

30	35	40	40	40	40	40	40	40	40	40	30	40	40	40	40

Anthropogenic Metric Values:

0	0	10	60	40	4	27	48	50	30	10	0	15	50	40	45

Reduced Anthropogenic Metric Values:

0	0	5	30	20	2	14	24	25	15	5	0	8	25	20	23

Target Metric Values:

30	35	45	70	60	42	54	64	65	55	45	30	48	65	60	63

B.3.1.3 Interday Rollback - Peak Shaving

To generate target metric values using Peak Shaving rollback, BenMAP simply truncates all anthropogenic metric values at the anthropogenic standard. These reduced anthropogenic metric values are added to the non-anthropogenic metric values to give the final target metric values.

Example:

Initial Metric Values:

30	35	50	100	80	44	67	88	90	70	50	30	55	90	80	85

Attainment Test: Highest value of metric <= 70

Interday Background Level: 40

Interday Rollback Method: Peak Shaving

Anthropogenic Standard: 30

Non-Anthropogenic Metric Values:

30	35	40	40	40	40	40	40	40	40	40	30	40	40	40	40

Anthropogenic Metric Values:

0	0	10	60	40	4	27	48	50	30	10	0	15	50	40	45

Reduced Anthropogenic Metric Values:

0	0	10	30	30	4	27	30	30	30	10	0	15	30	30	30

Target Metric Values:

30	35	50	70	70	44	67	70	70	70	50	30	55	70	70	70

B.3.2   Intraday Rollback - Adjusting Hourly Observations

Once a set of target metric values has been calculated for a pollutant with hourly observations (e.g., Ozone), BenMAP must adjust the hourly observations so that they produce the target metric values. There are three supported rollback methods for Intraday Rollback - Percentage, Incremental, and Quadratic. Each of these rollback methods requires some preprocessing of the initial monitor observations, and each can require multiple iterations to hit the target metric values. We will discuss this preprocessing and iteration first, and then go through Percentage and Incremental rollbacks in turn. Quadratic rollback is more complicated than these first two, and has its own section.

For various reasons, each of the Intraday Rollback methods can fail to hit the target metric values during a single pass through the rollback procedure (these will be discussed in detail below). As such, each of the rollback methods uses an iterative approach to get within a threshold of each of the target metric values - currently this threshold is 0.05. The iterative approach works as follows:

For each target metric value, BenMAP calculates the current value of the Attainment Test metric. For the first iteration, the metric value will be calculated using unadjusted hourly observations. For subsequent iterations, the metric value will be calculated using the current values of the adjusted hourly observations.

If the difference between the metric value and the target metric value is less than or equal to 0.05, the rollback procedure is finished. Otherwise, another iteration is required.

The Intraday Background Level specifies the portion of each observation which cannot be affected by human intervention - we call this portion the non-anthropogenic portion. Whatever portion is left over after subtracting out the background level is referred to as the anthropogenic portion. The anthropogenic portion of the initial monitor observations is the only part which will be affected by the Intraday Rollback Method.

In a way analogous to the Interday Rollback procedure, BenMAP calculates the twenty-four hourly anthropogenic observations and the twenty-four hourly non-anthropogenic observations using the following procedure for each hourly observation:

IF the current value of the observation is less than or equal to the Intraday Background Level,

non-anthropogenic observation = observation

anthropogenic observation = 0

ELSE

non-anthropogenic observation = Intraday Background Level

anthropogenic observation = observation - Intraday Background Level

Given (i) an Attainment Test Metric (e.g., EightHourDailyMax), (ii) an Intraday Background Level, and (iii) a target metric value for the day, BenMAP proceeds to adjust hourly observations in the following steps:

1. Calculate the Attainment Test metric (e.g., the 8-hour daily maximum);

2. Identify the "window" – i.e., the set of hours used to calculate the metric (e.g., if the 8-hour daily maximum is achieved in the first 8 hours, then the window is comprised of the first 8 hours);

3. Calculate the non-anthropogenic hourly observations (=min(hourly observation, Intraday Background Level));

4. Calculate the anthropogenic hourly observations (=hourly observation - Intraday Background Level);

5. Calculate the non-anthropogenic metric value (= the metric using the non-anthropogenic hourly observations in the "window");

6. Calculate the anthropogenic metric value (= the metric using the anthropogenic hourly observations in the "window");

7. Calculate the anthropogenic target metric value (= the target metric value minus the non-anthropogenic metric value);

8. Calculate the reduction required to get the anthropogenic metric value down to the anthropogenic target metric value;

9. Adjust all anthropogenic hourly observations by the reduction calculated on the previous step;

10. Calculate the adjusted hourly observations (= the adjusted anthropogenic hourly observation + the non-anthropogenic hourly observation).

B.3.2.1 Intraday Rollback - Percentage

Below, we present two examples of a percentage-based Intraday Rollback. In one example, a single iteration is needed, and in the second example, two iterations are

required because a number of the monitor values fall below the assumed background level.

B.3.2.1.1   Example: All Hourly Observations Exceed the Intraday Background (Single Iteration)

If all of the hourly observations in a day are greater than the Intraday Background Level, then the above procedure is straightforward and can be accomplished in a single iteration. We illustrate with the following example. Suppose that:

Metric = EightHourDailyMax,

Target metric value for a given day = 85

Intraday Background Level = 40.

And that the hourly observations on that day are:

530	45	50	60	45	45	45	60	70	100	100	100	100

100	100	100	100	60	45	50	45	45	47	47

Based on these observations, we see that the 8-hour daily maximum = 110.

Assuming a background level of 40, then the Anthropogenic hourly observations are:

490	5	10	20	5	5	5	20	30	60	60	60	60	60

60	60	60	20	5	10	5	5	7	7

Then, we know:

Anthropogenic metric value = 70.

Non-anthropogenic metric value = 40.

Anthropogenic target metric value = 45.

Percentage reduction required = ((70-45)/70) = 35.7%

All of the hourly anthropogenic observations are reduced by 35.7%. The average of the first 8 values (the window on which the Test metric is based) will be exactly 45, the anthropogenic target metric value. Finally, the adjusted hourly observations are calculated by adding the non-anthropogenic hourly observation to the adjusted hourly anthropogenic observations.

B.3.2.1.2   Example: Some Hourly Observations are Below the Intraday Background (Multiple Iterations Required)

In the above example, the anthropogenic target metric value was met on a single iteration

because all of the hourly observations were greater than the Intraday Background Level. In this case, a simple percent reduction of all hourly values will produce an average in the window that is equal to the anthropogenic target metric value. If some of the hourly observations in a day are less than or equal to the Intraday Background Level, however, then BenMAP uses an iterative procedure. On each iteration, it adjusts hourly observations using the 10-step method given above. It then compares the new metric value to the target metric value. If the difference is less than or equal to 0.05 ppb, the rollback procedure is finished. Otherwise, another iteration is required. The iterative procedure is illustrated in the following example.

Suppose that:

Metric = EightHourDailyMax,

Target metric value for a given day = 85

Intraday Background Level = 40.

Suppose also that the hourly observations on that day are:

| 530 | 20 | 25 | 60 | 35 | 35 | 40 | 60 | 70 | 100 | 100 | 100 | 100 |

| 100 | 100 | 100 | 100 | 60 | 33 | 40 | 30 | 30 | 25 | 20 |

Then, we know that the 8-hour daily maximum = 100.6.

Non-Anthropogenic Hourly Observations, Iteration One:

| 40 | 20 | 25 | 40 | 35 | 35 | 40 | 40 | 40 | 40 | 40 |

| 40 | 40 | 40 | 40 | 40 | 40 | 40 | 33 | 40 | 30 | 30 | 25 | 20 |

Anthropogenic Hourly Observations, Iteration One:

| 490 | 0 | 0 | 20 | 0 | 0 | 0 | 20 | 30 | 60 | 60 | 60 | 60 | 60 |

| 60 | 60 | 60 | 20 | 0 | 0 | 0 | 0 | 0 | 0 |

Non-Anthropogenic Metric Value: 34.4   (EightHourDailyMax - calculated over the same eight hour window as the initial metric value was calculated over)

Anthropogenic Metric Value: 66.3

Anthropogenic Target Metric Value: 50.6

Percentage Reduction Required: 23.6%

Reduced Anthropogenic Hourly Observations, Iteration One:

374	0	0	15	0	0	0	15	23	46	46	46

46	46	46	46	46	15	0	0	0	0	0	0

Reduced Hourly Observations, Iteration One:

414	20	25	55	35	35	40	55	63	86	86	86	86

86	86	86	86	55	33	40	30	30	25	20

Reduced Metric Value (EightHourDailyMax): 85.8

Target Metric Value (EightHourDailyMax): 85

Non-Anthropogenic Hourly Observations, Iteration Two:

40	20	25	40	35	35	40	40	40	40	40	40	40

40	40	40	40	40	33	40	30	30	25	20

Anthropogenic Hourly Observations, Iteration Two:

374	0	0	15	0	0	0	15	23	46	46	46	46

46	46	46	46	15	0	0	0	0	0	0

Non-Anthropogenic Metric Value: 40   (EightHourDailyMax - calculated over the same eight hour window the initial metric value was calculated over)

Anthropogenic Metric Value: 45.8

Anthropogenic Target Metric Value: 45

Percentage Reduction Required: 1.9%

Reduced Anthropogenic Hourly Observations, Iteration Two:

368	0	0	15	0	0	0	15	23	45	45	45	45	45

45   45   45   15   0   0   0   0   0   0

Reduced Hourly Observations, Iteration Two:

408   20   25   55   35   35   40   55   63   85   85   85   85   85

85   85   85   55   33   40   30   30   25   20

Reduced Metric Value (EightHourDailyMax): 85

The above example, in addition to illustrating the Intraday Percentage Rollback, also illustrates one reason why the iterative procedure can be necessary. When using the EightHourDailyMax metric in the Attainment Test, it is possible for the window over which the maximum eight hour average occurs to move after a single pass through the rollback procedure. When this happens, it becomes necessary to go through additional iterations to hit the target metric value.

B.3.3   Intraday Rollback - Incremental

To adjust hourly observations using Incremental rollback, BenMAP calculates the increment required to reduce the anthropogenic metric value to exactly the anthropogenic target metric value. This incremental reduction is then applied to all of the anthropogenic observations (but - they are not allowed to fall below zero). Finally, these reduced anthropogenic observations are added to the non-anthropogenic observations to give the final reduced observations.

Example:

Initial Hourly Observations:

20   20   25   60   35   35   40   70   35   30   65   90   76

65   35   35   54   60   33   40   30   30   25   20

Initial Metric Value (EightHourDailyMax): 60

Target Metric Value (EightHourDailyMax): 55

Intraday Background Level: 40

Intraday Rollback Method: Incremental

Non-Anthropogenic Hourly Observations, Iteration One:

20   20   25   40   35   35   40   40   35   30   40   40   40

| 40 | 35 | 35 | 40 | 40 | 33 | 40 | 30 | 30 | 25 | 20 |

Anthropogenic Hourly Observations, Iteration One:

| 0 | 0 | 0 | 20 | 0 | 0 | 0 | 30 | 0 | 0 | 25 | 50 | 36 |

| 25 | 0 | 0 | 14 | 20 | 0 | 0 | 0 | 0 | 0 | 0 |

Non-Anthropogenic Metric Value (EightHourDailyMax): 38.8

Anthropogenic Metric Value (EightHourDailyMax): 21.3

Anthropogenic Target Metric Value (EightHourDailyMax): 16.3

Incremental Reduction Required: 5.0

Reduced Anthropogenic Hourly Observations, Iteration One:

| 0 | 0 | 0 | 15 | 0 | 0 | 0 | 25 | 0 | 0 | 20 | 45 | 31 |

| 20 | 0 | 0 | 9 | 15 | 0 | 0 | 0 | 0 | 0 | 0 |

Reduced Hourly Observations, Iteration One:

| 20 | 20 | 25 | 55 | 35 | 35 | 40 | 65 | 35 | 30 | 60 | 85 | 71 |

| 60 | 35 | 35 | 49 | 55 | 33 | 40 | 30 | 30 | 25 | 20 |

Reduced Metric Value (EightHourDailyMax): 56.25

Target Metric Value (EightHourDailyMax): 55

Non-Anthropogenic Hourly Observations, Iteration Two:

| 20 | 20 | 25 | 40 | 35 | 35 | 40 | 40 | 35 | 30 | 40 | 40 | 40 |

| 40 | 35 | 35 | 40 | 40 | 33 | 40 | 30 | 30 | 25 | 20 |

Anthropogenic Hourly Observations, Iteration Two:

| 0 | 0 | 0 | 15 | 0 | 0 | 0 | 25 | 0 | 0 | 20 | 45 | 31 |

20	0	0	9	15	0	0	0	0	0	0

Non-Anthropogenic Metric Value (EightHourDailyMax): 38.8

Anthropogenic Metric Value (EightHourDailyMax): 17.5

Anthropogenic Target Metric Value (EightHourDailyMax): 16.3

Incremental Reduction Required: 1.25

Reduced Anthropogenic Hourly Observations, Iteration Two:

0	0	0	14	0	0	0	24	0	0	19	44	30

19	0	0	8	14	0	0	0	0	0	0

Reduced Hourly Observations, Iteration Two:

20	20	25	54	35	35	40	64	35	30	59	84	70

59	35	35	48	54	33	40	30	30	25	20

Reduced Metric Value (EightHourDailyMax): 55.3

Target Metric Value (EightHourDailyMax): 55

This example should actually continue for one further iteration, with a new Incremental Reduction of 0.3. This illustrates another reason why the iterative procedure can be necessary - for incremental reductions, the prohibition against values becoming negative can cause target metric values to not be met. Incremental reductions thus very often require multiple iterations.

B.3.4   Interday and Intraday Rollback - Quadratic

Quadratic rollback is based on an algorithm developed by Horst and Duff . The idea behind quadratic rollback is to reduce large values proportionally more than small values while just achieving the standard - that is, the out-of-attainment value should be more or less at the standard after the rollback (some small amount of error is involved).

The original quadratic rollback algorithm is designed to roll back hourly observations given a desired peak value. That is, it assumes that the Attainment Test metric is the one-hour average and the Attainment Test ordinality is one. As such, the algorithm was modified slightly to allow for ordinalities other than one to be used.

The basic formula for quadratic rollback is:

The basic formula for quadratic rollback is:

Reduced Observation = [1 - (A + B * Initial Observation)] * Initial Observation

where:

i ranges over the days being reduced.

A = 1 - V

V = Min(1, Vi)

Vi = (2 * Maximum Observation Value * Standard) / Xi

Xi = (2 * Maximum Observation Value * Metricsi) - Metricsi2

B = Max(0, [(V * Out of Attainment Value - Standard) / Out of Attainment Value2])

B.3.4.1 Quadratic Rollback - Interday

Because Quadratic Rollback was originally designed to adjust hourly observations to meet a daily metric standard, it is slightly complicated to use it to generate target metric values.

First, Quadratic Rollback calculates the anthropogenic out of attainment value by subtracting the Intraday Background Level from the out of attainment value. Note that this differs from the other interday rollback methods, which subtract the Interday Background Level from the out of attainment value. Similarly, the anthropogenic standard is calculated by subtracting the Intraday Background Level from the standard.

The anthropogenic observations and non-anthropogenic observations are then calculated. For pollutants which have daily observations (PM10, PM2.5) the anthropogenic metric values are used (see above for their calculation). For pollutants which have hourly observations (Ozone), Quadratic Rollback loops through each metric value and calculates the twenty four corresponding anthropogenic observations and non-anthropogenic observations as follows:

IF the metric value is at or below the Interday Background Level,

For each observation,

non-anthropogenic observation = observation

anthropogenic observation = 0

ELSE

For each observation,

IF the observation is at or below the Intraday Background Level

non-anthropogenic observation = observation

anthropogenic observation = 0

ELSE

non-anthropogenic observation = Intraday Background Level

anthropogenic observation = observation - Intraday Background Level

A new set of anthropogenic metric values is then calculated by generating the Attainment Test metric from the anthropogenic observations. The Quadratic Rollback algorithm is then called, passing in the anthropogenic metric values as Metrics, anthropogenic observations as Observations, anthropogenic standard as Standard, and anthropogenic out of attainment value as Out of Attainment Value. The result is a set of reduced anthropogenic observations. These are then added together with the non-anthropogenic observations to give a final set of reduced observations.

Then, if Quadratic Rollback was also selected as the Intraday Rollback method, these observations are used as the final reduced observations for the monitor. Otherwise, metric targets are generated from these hourly observations, and the observations themselves are discarded

B.3.4.2 Quadratic Rollback - Intraday

Quadratic Rollback can also be used to adjust hourly observations to meet metric targets generated via a different rollback method. In this case, the algorithm is used to adjust each set of twenty four hourly observations to meet the corresponding metric target. Intraday Quadratic Rollback uses the normal set of anthropogenic observations as Observations, a single normal anthropogenic metric value as Metrics, and the normal anthropogenic metric target as Standard. Intraday Quadratic Rollback tends to always slightly miss its metric target, so it is not run in an iterative fashion as the other Intraday Rollback Methods are (doing so would sometimes result in an infinite loop).

Appendix C: Air Pollution Exposure Estimation Algorithms

BenMAP has grouped individuals into what we refer to as "population grid-cells," where the grid-cells typically conform to some type of grid used in an air quality model, such as the REMSAD air quality model, or just the counties of the United States. For each type of grid, the population is built in each grid-cell by aggregating census block data. In the next step, BenMAP estimates the air pollution exposure for each grid-cell, with the assumption that people living within a particular grid-cell experience the same air pollution levels.

You have a variety of approaches to estimate the exposure to air pollution for the people living within a given population grid-cell. Perhaps the simplest approach is to use model data directly, and to assume that the people living within a particular model grid-cell experience the level estimated by the model. An alternative approach is to use air pollution monitoring data, where you may choose the closest monitor data to the center of a grid-cell or take an average of nearby monitors. In a third general approach, you may combine both modeling and monitoring data to estimate exposure.

When combining modeling and monitoring data, BenMAP scales or adjusts the monitoring data with modeling data. The advantage of modeling data is that they can provide predictions for years in which monitoring data are not available, as well as to provide predictions in areas of the country for which monitoring data are not available. And the advantage of monitor data is that they are based on actual observations. Combining both sources of information, allows BenMAP to make more informed predictions.

The goal of estimating exposure is to provide the necessary input for concentration-response functions, so that BenMAP can estimate the impact of air pollution on adverse health effects. Table C-1 lists the types of metrics commonly used in concentration-response functions. In the case of air pollution metrics calculated on a daily basis, such as the one-hour maximum and the 24-hour average, it is often the case that there are missing days of data. Air quality modeling is often conducted on a subset of the days in the year, and air quality monitors often miss a number of observations through out the year. To account for missing days, BenMAP represents the distribution of daily metrics with a certain number points or "bins," where each bin represents a certain range of the distribution, with the underlying assumption that missing days have the same distribution as the available data. For example, for analyses of the United States the Environmental Protection Agency has typically used 153 bins to represent the ozone season from May through September, and for particulate matter they have used 365 bins to represent the year. In addition to being able to account for incomplete or missing data, and using bins to represent the distribution provides a uniform approach that allows for easy comparison of different monitors.

Table C-1. Metrics Typically Used in Concentration-Response Functions for Criteria Air Pollutants

Measurement Frequency	Metric Name	Metric Description
Daily (e.g., PM2.5, PM10)	Daily Average	Daily average
	Annual Average	Average of four quarterly averages. The four quarters are defined as: Jan-Mar, April-June, Jul-Sep, Oct-Dec.
	Annual Median	Median of values through out the year.
Hourly (e.g., Ozone)	1-hour Daily Max	Highest hourly value from 12:00 A.M. through 11:59 P.M.
	8-hour Daily Average	Average of hourly values from 9:00 A.M. through 4:59 P.M.
	12-hour Daily Average	Average of hourly values from 8:00 A.M. through 7:59 P.M.
	24-hour Daily Average	Average of hours from 12:00 A.M. through 11:59 P.M.

C.1   Direct Modeling

When using direct modeling data to estimate exposure, BenMAP assumes that the people living within a particular air pollution model grid-cell experience the same air pollution levels. BenMAP then estimates the air pollution metrics of interest, as defined for each pollutant. (See the section on defining pollutants in the Loading Data chapter.)

Generally modeling data providing hourly observations are complete for any given day. However, it is common to have missing days of modeling data during the course of a year. Given the estimated metrics from the available data, BenMAP then represents the distribution of daily metrics with the number of days specified for each pollutant. By calculating bins with the available days, BenMAP assumes that the distribution of missing days is similar to the distribution of available data.

C.2   Closest Monitor

When using the closet monitor to represent air pollution levels at a population grid-cell, BenMAP identifies the center of the population grid-cell, and then chooses the monitor that is closest to the center. In the simplest case, BenMAP assigns the closest monitor to a population grid-cell, uses the monitoring data to calculate the annual and daily air pollution metrics, and then calculates the bins that represent the distribution of the daily metrics. The annual metrics and the binned daily metrics are then used in the calculation of health effects.

The figure below presents nine population grid-cells and three monitors, with the focus on identifying the monitor closest to grid-cell "E." In this example, the closest monitor happens to be 10 miles away from the center of grid-cell E, and the data from this monitor would be used to estimate air pollution levels for the population in this grid-cell. An analogous procedure would be used to estimate air pollution levels in the other grid-cells (A, B, C, D, F, G, H, and I).

To capture some of the information generated by air pollution models, BenMAP can also scale the data from the closest monitor with air pollution modeling data. BenMAP includes two types of scaling – "temporal" and "spatial" scaling. We discuss each below.

```
┌─────────────────┬─────────────┬──────────────────┐
│ A               │           B │ 25 miles      C  │
│                 │             │     *            │
│                 │   15 miles  │                  │
│                 │       *     │                  │
├─────────────────┼─────────────┼──────────────────┤
│ D               │           E │                F │
│                 │             │                  │
│ 10 miles   *    │ ──────▶ #   │  * 15 miles      │
├─────────────────┼─────────────┼──────────────────┤
│ G               │           H │                I │
│                 │      *      │      *           │
│                 │   20 miles  │   25 miles       │
│        *        │             │                  │
│    30 miles     │             │                  │
└─────────────────┴─────────────┴──────────────────┘
```
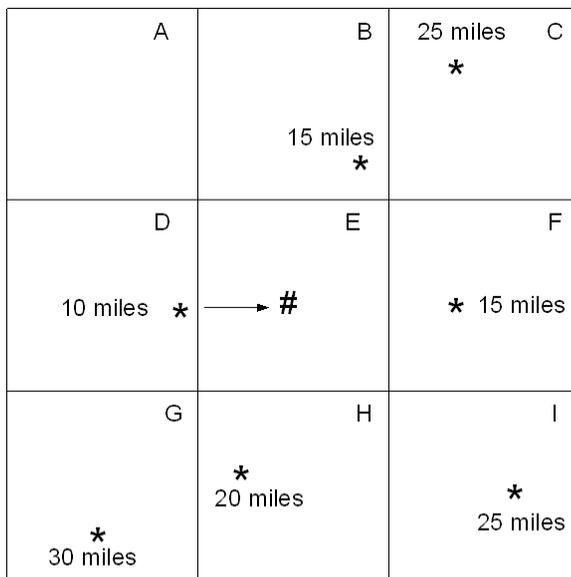
\# = Center Grid-Cell "E"

* = Air Pollution Monitor

C.2.1   Closest Monitor – Temporal Scaling

With temporal scaling, BenMAP scales monitoring data with the ratio of the future-year to base-year modeling data, where the modeling data is from the modeling grid-cell containing the monitor. In the case of pollutants typically measured hourly, such as ozone, BenMAP scales the hourly monitor values, calculates the annual and daily metrics of interest, and then bins the daily metrics. In the case of pollutants typically measured daily, BenMAP scales the daily values, calculates the annual metrics of interest, and then bins the daily metric.

Consider the case in the figure below. To forecast air pollution levels for 2030, BenMAP would multiply the 1995 monitor value (80 ppb) by the ratio of the 2030 model value (70 ppb) to the 1995 model value (95 ppb):

Forecast2030 = Monitor Value 1995 * (Model Value D, 2030 / Model Value D, 1995)

Forecast2030 = 80 ppb * (70 ppb / 95 ppb) = 58.9 ppb.

= Center Grid-Cell "E"

* = Air Pollution Monitor

In this example, we have examined the adjustment of a single monitor value with the ratio of single model values. The approach is essentially the same when there are multiple monitor values and multiple model values.

C.2.2   Closest Monitor – Spatial Scaling

With spatial scaling, we are estimating a monitor value for the center of each population grid-cell. We start by choosing the closest monitor to the center of each population grid-cell, and then we scale this closest monitor with modeling data. In particular, BenMAP multiplies the monitoring data with the ratio of the base-year modeling data for the destination grid-cell to the base-year modeling data for grid-cell containing the monitor. The spatial scaling occurs in the same fashion as with temporal scaling. In the case of pollutants typically measured hourly, such as ozone, BenMAP scales the hourly monitor values, calculates the annual and daily metrics of interest, and then bins the daily metrics. In the case of pollutants typically measured daily, BenMAP scales the daily values, calculates the annual metrics of interest, and then bins the daily metric.

To estimate air pollution levels for 1995 in grid-cell "E" below, BenMAP would multiply the 1995 closest monitor value (80 ppb) by the ratio of the 1995 model value for grid-cell "E" (70 ppb) to the 1995 model value for grid-cell "D" (95 ppb):

Forecast1995 = Monitor Value1995 * (Model Value E, 1995 / Model Value D, 1995)

Forecast1995 = 80 ppb * (85 ppb / 95 ppb) = 71.6 ppb.

= Center Grid-Cell "E"

* = Air Pollution Monitor

C.2.3 Closest Monitor – Temporal and Spatial Scaling

Combining both temporal and spatial scaling, BenMAP first multiplies monitoring data with both the ratio of the future-year to base-year modeling data, where the modeling data is from the modeling grid-cell containing the monitor. This gives a temporary forecast for 2030. BenMAP then multiplies this temporary forecast with the ratio of the future-year modeling data for the destination grid-cell to the future-year modeling data for grid-cell containing the monitor. As seen below, this simplifies to multiplying monitoring data with both the ratio of future-year modeling data from the destination grid-cell to the base-year modeling data from the grid-cell containing the monitor. Again, as described for temporal and spatial scaling, BenMAP first scales the hourly and daily values, generates the metrics of interest and then bins the daily metrics.

To forecast air pollution levels for 2030 in the figure below, BenMAP would multiply the 1995 monitor value (80 ppb) by the ratio of the 2030 model value (70 ppb) to the 1995 model value (95 ppb):

Temporary Forecast 2030 = Monitor Value 1995 * (Model Value D, 2030 / Model Value D, 1995)

Temporary Forecast 2030 = 80 ppb * (70 ppb / 95 ppb) = 58.9 ppb.

Forecast 2030 = Temporary Forecast 2030 * (Model Value E, 2030 / Model Value D, 2030)

Forecast 2030 = 58.9 ppb * (60 ppb / 70 ppb) = 50.5 ppb.

Note that through cancellation, this equation simplifies to:

Forecast 2030 = Monitor Value 1995 * (Model Value E, 2030 / Model Value D, 1995)

$\#$ = Center Grid-Cell "E"

$*$ = Air Pollution Monitor

C.3   Voronoi Neighbor Averaging (VNA)

Like the closest monitor approach, the Voronoi Neighbor Averaging (VNA) algorithm uses monitor data directly or in combination with modeling data. However, instead of using the single closest monitor to estimate exposure at a population grid-cell, the VNA algorithm interpolates air quality at every population grid cell by first identifying the set of monitors that best "surround" the center of the population grid-cell.

= Center Grid-Cell "E"

*= Air Pollution Monitor

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

\# = Center Grid-Cell "E"

* = Air Pollution Monitor

BenMAP then chooses those monitors that share a boundary with the center of grid-cell "E." These are the nearest neighbors, BenMAP uses these monitors to estimate the air pollution level for this grid-cell.

= Center Grid-Cell "E"

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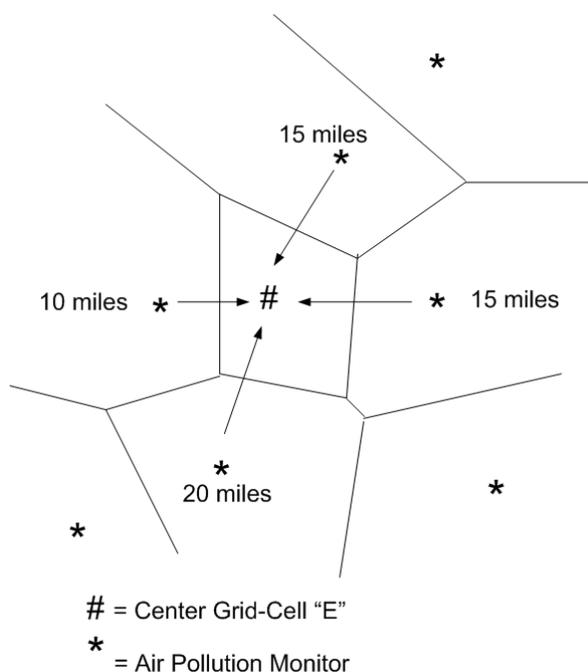
To estimate the air pollution level in each grid-cell, BenMAP calculates the metrics for each of the neighboring monitors, and then calculates an inverse-distance weighted average of the metrics. The further the monitor is from the BenMAP grid-cell, the smaller the weight.

In the figure below, the weight for the monitor 10 miles from the center of grid-cell E is calculated as follows:

$$weight_{10\ miles} = \frac{\frac{1}{10}}{\left(\frac{1}{10} + \frac{1}{15} + \frac{1}{15} + \frac{1}{20}\right)} = 0.35 \ .$$

The weights for the other monitors would be calculated in a similar fashion. BenMAP would then calculate an inverse-distance weighted average for 1995 air pollution levels in grid-cell E as follows:

Forecast 1995 = 0.35*80 ppb + 0.24*90 ppb+ 0.24*60 ppb + 0.18*100 ppb = 81.2 ppb .

A	B	C
	Monitor: 1995 90 ppb 15 miles *	*

= Center Grid-Cell "E"

* = Air Pollution Monitor

C.3.1 VNA | Temporal Scaling

Like the closest monitor approach, the Voronoi Neighbor Averaging (VNA) algorithm uses monitor data directly or in combination with modeling data. However, instead of using the single closest monitor to estimate exposure at a population grid-cell, the VNA algorithm interpolates air quality at every population grid cell by first identifying the set of monitors that best "surround" the center of the population grid-cell.

= Center Grid-Cell "E"

* = Air Pollution Monitor

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

\# = Center Grid-Cell "E"

* = Air Pollution Monitor

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

= Center Grid-Cell "E"

* = Air Pollution Monitor
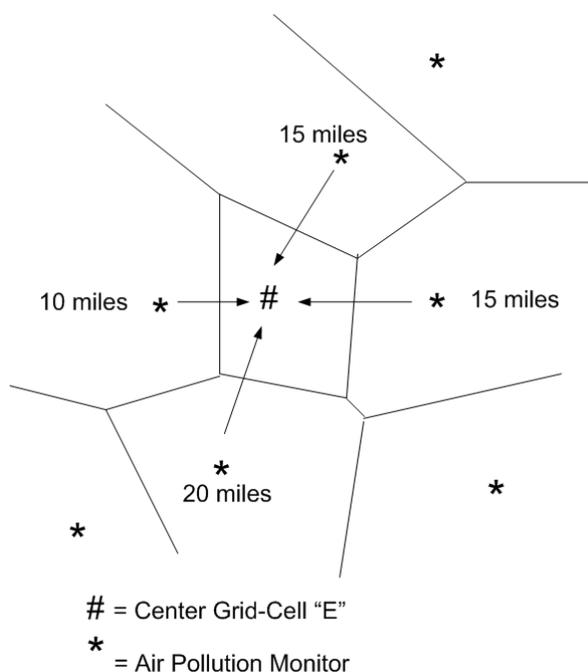
To estimate the air pollution level in each grid-cell, BenMAP calculates the annual and the binned daily metrics for each of the neighboring monitors, and then calculates an inverse-distance weighted average of the metrics. The further the monitor is from the BenMAP grid-cell, the smaller the weight.

In the figure below, the weight for the monitor 10 miles from the center of grid-cell E is calculated as follows:

$$weight_{10\ miles} = \frac{\frac{1}{10}}{\left(\frac{1}{10} + \frac{1}{15} + \frac{1}{15} + \frac{1}{20}\right)} = 0.35 \ .$$

The weights for the other monitors would be calculated in a similar fashion. BenMAP would then calculate an inverse-distance weighted average for 1995 air pollution levels in grid-cell E as follows:

Forecast 1995 = 0.35*80 ppb + 0.24*90 ppb+ 0.24*60 ppb + 0.18*100 ppb = 81.2 ppb .

= Center Grid-Cell "E"

* = Air Pollution Monitor

Note that BenMAP is calculating an inverse-distance weighted average of the annual metrics and the binned daily metrics. Alternatively, BenMAP could calculate an inverse-distance weighted average of the hourly and daily observations, calculated the annual and daily metrics, and then binned the daily metrics.

C.3.2 Voronoi Neighbor Averaging (VNA) – Spatial Scaling

BenMAP can also combine VNA with spatial scaling. For each of the neighbor monitors, BenMAP multiplies the monitoring data with the ratio of the base-year modeling data for the destination grid-cell to the base-year modeling data for grid-cell containing the monitor. The spatial scaling occurs in the same fashion as with temporal scaling. In the case of pollutants typically measured hourly, such as ozone, BenMAP scales the hourly monitor values, calculates the annual and daily metrics of interest, and then bins the daily metrics. In the case of pollutants typically measured daily, BenMAP scales the daily values, calculates the annual metrics of interest, and then bins the daily metric.

Consider the example in the figure below. To forecast air pollution levels for 1995, BenMAP would multiply the 1995 monitor value by the ratio of the 1995 model value to the 1995 model value:

$$Forecast_{1995} = \sum_{i=1}^{4} Weight_i * Monitor_i * \frac{Model_{E,1995}}{Model_{i,1995}}$$

$$Forecast_{1995} = \left(0.35*80*\frac{85}{95}\right) + \left(0.24*90*\frac{85}{100}\right) + \left(0.24*60*\frac{85}{80}\right) + \left(0.18*100*\frac{85}{120}\right) = 70.8\,ppb$$

C.3.3   Voronoi Neighbor Averaging (VNA) – Temporal & Spatial Scaling

Combining both temporal and spatial scaling, BenMAP multiplies monitoring data with the ratio of the future-year to base-year modeling data, where the future-year modeling data are from the destination grid-cell and the base-year modeling data are from the grid-cell containing the monitor. One the hourly and daily monitoring data are scaled, BenMAP generates the metrics of interest, bins the daily metrics, and then uses the metrics to estimate adverse health effects in the population grid-cell.

The figure below gives an example of combining temporal and spatial scaling.

$$Forecast_{2030} = \sum_{i=1}^{4} Weight_i * Monitor_i * \frac{Model_{E,2030}}{Model_{i,1995}}$$

$$Forecast_{2030} = \left(0.35*80*\frac{60}{95}\right) + \left(0.24*90*\frac{60}{100}\right) + \left(0.24*60*\frac{60}{80}\right) + \left(0.18*100*\frac{60}{120}\right) = 50.0$$

C.4   Fixed Radius

When using the fixed radius option to represent air pollution levels at a population grid-cell, BenMAP identifies all monitors within a specified distance of the center of the population grid-cell, calculates the metrics at each monitor, and then calculates a weighted average of the metrics using the algorithms described for VNA. When no monitors are within the specified distance, BenMAP assigns the closest monitor to a population grid-cell, and calculates the metrics using the algorithms described for the closest monitor approach.

C.5   Temporal and Spatial Scaling Adjustment Factors

As presented in the preceding examples of temporal and spatial scaling, the closest monitor , VNA, and fixed radius approaches can use model data to scale monitor observations. In the examples above, we scaled single monitor values with the ratio of single model values. In fact, however, the scaling involves multiple monitor values and multiple model values.

To proceed with the scaling, BenMAP takes the modeling values and splits them into groups, depending on how the pollutant is defined. (See the section on defining pollutants in the Loading Data chapter.) The United States setup has defined ozone to have a default of 10 adjustment factors for the ozone season, where the first group represents the first 10 percent of the model observations; the second group represents the observations between the 10th and 20th percentile; and so on through the tenth group, which represents the observations between the 90th and 100th percentiles. BenMAP then averages the values in each group. The United States setup has defined particulate matter to have five adjustment factors for each of the four seasons in the year, where the first group in each season represents the first 20 percent of the model observations; the second group represents the observations between the 20th and 40th percentiles; and so on. Then, as for ozone model values, BenMAP averages the particulate matter model values in each group.

BenMAP treats the monitor values in a similar way. It sorts the monitor values from low to high, and divides them into the same number groups as there are scaling factors.

C.5.1   Calculation of Scaling Factors

In developing scaling factors for the standard United States setup, BenMAP sorts the modeling data into either 10 groups or 20 groups, depending on the pollutant (e.g., 10 for

ozone and 20 for particulate matter, 5 for each of the 4 seasons). Given the number of groups, BenMAP then determines how to assign the model values. In determining to which group a value belongs, BenMAP assigns a two-digit "percentile" to each value. With values in a given grid-cell sorted from low to high, the percentile for each value will equal: (the observation rank number minus 0.5) divided by (the total number of values) multiplied by (100). If there are 250 hourly values, the first hourly value will have a percentile = $(1-0.5)/(250)*(100) = 0.20\%$; the 27th value will have a percentile = $(27-0.5)/(250)*(100) = 10.60\%$; and so on.

Each data group is represented by "group-lo" and "group-hi" values. These are the minimum and the maximum percentiles in each group, where group-lo equals: (group rank minus 1) multiplied by (100) divided by (the number of groups); and group-hi equals: (group rank) multiplied by (100) divided by (the number of groups) minus 0.001. If there are ten groups: the first group will have: group-lo = $(1-1)/100*10 = 0.000\%$, and group-hi = $(1/100*10)-0.001 = 9.999\%$; the second group will have: group-lo = $(2-1)/100*10 = 10.000\%$, and group-hi = $(2/100*10)-0.001 = 19.999\%$; and so on to the tenth group, which will have: group-lo = $(10-1)/100*10 = 90.000\%$, and group-hi = $(10/100*10)-0.001 = 99.999\%$. BenMAP assigns each observation to a particular group with the following algorithm: if "group-lo" <"percentile" < "group-hi", then assign the observation to that data group.

Below we give some examples of the calculations that BenMAP performs when scaling.

C.5.1.1 Example: PM2.5 Scaling Factors in U.S. Setup

After preparing the PM2.5 model and monitor data, BenMAP calculates the following:

$$adjusted\ monitor_{i,j,future} = monitor_{i,j,base} \cdot \frac{REMSAD_{j,k,future}}{REMSAD_{j,l,base}}$$

where:

adjusted monitor	= predicted daily PM2.5 level, after adjustment by model data (µg/m3)
monitor	= observed daily PM2.5 monitor level (µg/m3)
i	= day identifier
j	= model season/quintile group (1 to 20)
k	= grid cell identifier for population grid cell
l	= grid cell identifier for grid cell containing monitor

base	= base-year (e.g., 2000)
future	= future-year (e.g., 2020)
REMSAD	= representative model season/quintile value (μg/m3)

After adjusting the monitor values to reflect air quality modeling, BenMAP calculates for each monitor the PM2.5 metrics needed to estimate adverse health effects. In the case of VNA, BenMAP then calculates a weighted average (e.g., inverse-distance weighted average) of the neighbors identified for each population grid cell:

$$population\ grid\ cell_{future} = \sum_{m=1}^{n} adjusted\ monitor_{m,future} \cdot weight_m \cdot$$

where:

population grid cell	= inverse distance-weighted PM2.5 metric at population grid cell (μg/m3)
adjusted monitor	= predicted PM2.5 metric, after adjustment by model data (μg/m3)
m	= monitor identifier
base	= base-year (e.g., 2000)
future	= future-year (e.g., 2020)
weight	= inverse-distance weight for monitor

After generating the bins for both the baseline and control scenarios, BenMAP uses these to calculate the change in air quality needed in most health impact functions to calculate the change in adverse health effects. To calculate the change in air quality, BenMAP subtracts the baseline value in the first bin from the control value in the first bin, and so on for each of the bins created for the daily $PM_{2.5}$ average.

C.5.1.2 Example: Ozone Scaling in U.S. Setup

After preparing the ozone model and monitor data, BenMAP calculates the following:

$$adjusted\ monitor_{i,j,future} = monitor_{i,j,base} \cdot \frac{CAMX_{j,k,future}}{CAMX_{j,l,base}}$$

where:

adjusted monitor	= predicted hourly ozone level, after adjustment by model data (ppb)
monitor	= observed hourly ozone monitor level (ppb)
i	= hour identifier
j	= model decile group (1 to 10)
k	= grid cell identifier for population grid cell
l	= grid cell identifier for grid cell containing monitor
base	= base-year (e.g., 1996)
future	= future-year (e.g., 2030)
CAMX	= representative model decile value (ppb)

After adjusting the monitor values to reflect air quality modeling, BenMAP calculates for each monitor the ozone metrics needed to estimate adverse health effects. In the case of VNA, BenMAP then calculates a weighted average (e.g., inverse-distance weighted average) of the neighbors identified for each population grid cell:

$$population\ grid\ cell_{future} = \sum_{m=1}^{n} adjusted\ monitor_{m, future} \cdot weight_m \cdot$$

where:

population grid cell	= inverse distance-weighted ozone metric at population grid cell (ppb)
adjusted monitor	= predicted ozone metric, after adjustment by model data (ppb)
m	= monitor identifier
future	= future-year (2020, 2030)
weight	= inverse-distance weight for monitor

After generating the bins for both the baseline and control scenarios, BenMAP can use these to calculate the change in air quality needed in most C-R functions to calculate the

change in adverse health effects. To calculate the change in air quality, BenMAP subtracts the baseline value in the first bin from the control value in the first bin, and so on for each of the bins created for the daily ozone metrics.

C.6   Binned Metrics

When estimating air pollution exposure, it will often happen that metrics are often not available for each day in the year. To remedy this, BenMAP calculates representative values or bins with the available daily metrics, under the assumption that the missing days have a similar distribution. Each bin represents a day. In the case where there are 365 bins, the set of bins represents the entire year.

When combining air pollution metrics from multiple monitors, BenMAP first calculates the bins for the daily metrics, and then combines the bins, such as with some form of VNA . Once BenMAP has calculated binned exposure measures for both a baseline and a control scenario, BenMAP then takes the difference between the two scenarios for each bin – taking the difference between the baseline value in the first bin and the control value in the first bin, and so on for each of the bins.

Appendix D: Deriving Health Impact Functions

This Appendix provides of an overview regarding the health impact functions that BenMAP uses to estimate the impact of a change in air pollution on adverse health effects. It provides a description of the particular types of health impact functions that BenMAP uses.

The functional form of the relationship between the change in pollutant concentration, Δx, and the change in population health response (usually an incidence rate), Δy depends on the functional form of the C-R function from which it is derived, and this depends on the underlying relationship assumed in the epidemiological study chosen to estimate a given effect. For expository simplicity, the following subsections refer simply to a generic adverse health effect, "y" and uses particulate matter (PM) as the pollutant – that is, $\Delta x = \Delta PM$ – to illustrate how the relationship between Δx and Δy is derived from each of several different C-R functions.

Estimating the relationship between ΔPM and Δy can be thought of as consisting of three steps:

(1) choosing a functional form of the relationship between PM and y (the C-R function),

(2) estimating the values of the parameters in the C-R function assumed, and

(3) deriving the relationship between ΔPM and Δy (the health impact function) from the relationship between PM and y (the C-R function).

Epidemiological studies have used a variety of functional forms for C-R functions. Some studies have assumed that the relationship between adverse health and pollution is best described by a linear form, where the relationship between y and PM is estimated by a linear regression in which y is the dependent variable and PM is one of several independent variables. Log-linear regression and logistic regression are other common forms.

Note that the the log-linear form used in the epidemiological literature is often referred to as "Poisson regression" because the underlying dependent variable is a count (e.g., number of deaths), believed to be Poisson distributed. The model may be estimated by regression techniques but is often estimated by maximum likelihood techniques. The form of the model, however, is still log-linear.

D.1   Overview

The relationship between the concentration of a pollutant, x, and the population response, y, is called the concentration-response (C-R) function. For example, the concentration of the pollutant may be fine particulate matter ($PM_{2.5}$) in µg/m3 per day, and the population response may be the number of premature deaths per 100,000 population per day. C-R functions are estimated in epidemiological studies. A functional form is chosen by the

researcher, and the parameters of the function are estimated using data on the pollutant (e.g., daily levels of $PM_{2.5}$) and the health response (e.g., daily mortality counts). There are several different functional forms, discussed below, that have been used for C-R functions. The one most commonly used is the log-linear form, in which the natural logarithm of the health response is a linear function of the pollutant concentration.

For the purposes of estimating benefits, we are not interested in the C-R function itself, however, but the relationship between the change in concentration of the pollutant, Δx, and the corresponding change in the population health response, Δy. We want to know, for example, if the concentration of $PM_{2.5}$ is reduced by 10 μg/m3, how many premature deaths will be avoided? The relationship between Δx and Δy can be derived from the C-R function, as described below, and we refer to this relationship as a health impact function.

Many epidemiological studies, however, do not report the C-R function, but instead report some measure of the change in the population health response associated with a specific change in the pollutant concentration. The most common measure reported is the relative risk associated with a given change in the pollutant concentration. A general relationship between Δx and Δy can, however, be derived from the relative risk. The relative risk and similar measures reported in epidemiological studies are discussed in the sections below. The derivation of the relationship of interest for BenMAP – the relationship between Δx and Δy – is discussed in the subsequent sections.

D.2   Review Relative Risk and Odds Ratio

The terms relative risk and odds ratio are related but distinct. Table D-1 provides the basis for demonstrating their relationship.

Table D-1. Relative Risk and Odds Ratio Notation

Exposure	Fraction of Population		Adverse Effect Measure	
	Affected	*Not Affected*	*Relative Risk*	*Odds*
Baseline Pollutant Exposure	y0	1-y0		y0/(1-yc)
			y0/yc	
Control Pollutant Exposure	yc	1-yc		yc/(1-yc)

The "risk" that people with baseline pollutant exposure will be adversely affected (e.g., develop chronic bronchitis) is equal to y_0, while people with control pollutant exposure face a risk, y_c, of being adversely affected. The relative risk (RR) is simply:

$$RR = \frac{y_0}{y_c}.$$

The odds that an individual facing high exposure will be adversely affected is:

$$Odds = \frac{y_0}{1 - y_0}.$$

The odds ratio is then:

$$Odds\ Ratio = \frac{\left(\dfrac{y_0}{1 - y_0}\right)}{\left(\dfrac{y_c}{1 - y_c}\right)}.$$

This can be rearranged as follows:

$$Odds\ Ratio = \frac{y_0}{y_c}\left(\frac{1 - y_c}{1 - y_0}\right) = RR\left(\frac{1 - y_c}{1 - y_0}\right).$$

As the risk associated with the specified change in pollutant exposure gets small (i.e., both y_0 and y_c approach zero), the ratio of $(1-y_c)$ to $(1-y_0)$ approaches one, and the odds ratio approaches the relative risk. This relationship can be used to calculate the pollutant coefficient in the C-R function from which the reported odds ratio or relative risk is derived, as described below.

D.3   Linear Model

A linear relationship between the rate of adverse health effects (incidence rate) and various explanatory variables is of the form:

$$y = \alpha + \beta \cdot PM$$

where α incorporates all the other independent variables in the regression (evaluated at their mean values, for example) times their respective coefficients. The relationship between the change in the rate of the adverse health effect from the baseline rate (y_0) to the rate after control (y_c) associated with a change from PM_0 to PM_c is then:

$$\Delta y = y_e - y_0 = \beta(PM_e - PM_0) = \beta \Delta PM.$$

For example, Ostro et al. (1991, Table 5) reported a $PM_{2.5}$ coefficient of 0.0006 (with a standard error of 0.0003) for a linear relationship between asthma and $PM_{2.5}$ exposure.

The lower and upper bound estimates for the $PM_{2.5}$ coefficient are calculated as follows:

$$\beta_{lower\ bound} = \beta - (1.96 \cdot \sigma_\beta) = 0.0006 - (1.96 \cdot 0.0003) = 1.2 \cdot 10^{-5}$$
$$\beta_{upper\ bound} = \beta + (1.96 \cdot \sigma_\beta) = 0.0006 + (1.96 \cdot 0.0003) = 0.00119$$

It is then straightforward to calculate lower and upper bound estimates of the change in asthma.

D.4   Log-linear Model

The log-linear relationship defines the incidence rate (y) as:

$$y = B \cdot e^{\beta \cdot PM}$$

or, equivalently,

$$\ln(y) = \alpha + \beta \cdot PM,$$

where the parameter B is the incidence rate of y when the concentration of PM is zero, the parameter β is the coefficient of PM, ln(y) is the natural logarithm of y, and $\alpha = \ln(B)$. Other covariates besides pollution clearly affect mortality. The parameter B might be thought of as containing these other covariates, for example, evaluated at their means. That is, B = Boexp{ß1x1 + ... + ßnxn}, where Bo is the incidence of y when all covariates in the model are zero, and x1, ... , xn are the other covariates evaluated at their mean values. The parameter B drops out of the model, however, when changes in y are calculated, and is therefore not important.

The relationship between ΔPM and Δy is:

$$\Delta y = y_e - y_0 = Be^{\beta \cdot PM_e} - Be^{\beta \cdot PM_0}.$$

This may be rewritten as:

$$\Delta y = e^{\beta \cdot PM_c}\left(e^{\beta \cdot (PM_c - PM_t)} - 1\right) = y_0 \cdot \left(e^{\beta \cdot \Delta PM} - 1\right),$$

where y_0 is the baseline incidence rate of the health effect (i.e., the incidence rate before the change in PM).

The change in the incidence of adverse health effects can then be calculated by multiplying the change in the incidence rate, Δy, by the relevant population (e.g., if the rate is number per 100,000 population, then the relevant population is the number of 100,000s in the population).

When the PM coefficient (β) and its standard error (σ_β) are published (e.g., Ostro et al., 1989), then the coefficient estimates associated with the lower and upper bound may be calculated easily as follows:

$$\beta_{lower\ bound} = \beta - (1.96 \cdot \sigma_\beta)$$

$$\beta_{upper\ bound} = \beta + (1.96 \cdot \sigma_\beta).$$

Epidemiological studies often report a relative risk for a given ΔPM, rather than the coefficient, β (e.g., Schwartz et al., 1995, Table 4). Recall that the relative risk (RR) is simply the ratio of two risks:

$$RR = \frac{y_0}{y_c} = e^{\beta \Delta PM}.$$

Taking the natural log of both sides, the coefficient in the C-R function underlying the relative risk can be derived as:

$$\beta = \frac{\ln(RR)}{\Delta PM}.$$

The coefficients associated with the lower and upper bounds (e.g., the 2.5 and 97.5 percentiles) can be calculated by using a published confidence interval for relative risk, and then calculating the associated coefficients.

Because of rounding of the published RR and its confidence interval, the standard error for the coefficient implied by the lower bound of the RR will not exactly equal that implied by the upper bound, so an average of the two estimates is used. The underlying standard error for the coefficient (σ_β) can be approximated by:

$$\sigma_{\beta, 2.5\ percentile} = \frac{\beta - \beta_{2.5\ percentile}}{1.96}$$

$$\sigma_{\beta, 97.5\ percentile} = \frac{\beta_{97.5\ percentile} - \beta}{1.96}$$

$$\sigma_{\beta} \cong \frac{\sigma_{\beta, 2.5\ percentile} + \sigma_{\beta, 97.5\ percentile}}{2}.$$

D.5   Logistic Model

In some epidemiological studies, a logistic model is used to estimate the probability of an occurrence of an adverse health effect. Given a vector of explanatory variables, X, the logistic model assumes the probability of an occurrence is:

$$y = prob(occurrence \mid X \cdot \beta) = \left(\frac{e^{X\beta}}{1 + e^{X\beta}} \right),$$

where β is a vector of coefficients. Greene (1997, p. 874) presents models with discrete dependent variables, such as the logit model. See also Judge et al. (1985, p. 763). This may be rewritten as:

$$y = \frac{e^{X\beta}}{1 + e^{X\beta}} \cdot \frac{e^{-X\beta}}{e^{-X\beta}} = \frac{1}{1 + e^{-X\beta}}$$

The odds of an occurrence is:

$$odds = \frac{y}{1-y} = \frac{\left(\frac{1}{1+e^{-X\beta}}\right)}{1 - \frac{1}{1+e^{-X\beta}}}$$

$$\Rightarrow odds = \frac{\left(\frac{1}{1+e^{-X\beta}}\right)}{1 - \frac{1}{1+e^{-X\beta}}} = \frac{\left(\frac{1}{1+e^{-X\beta}}\right)}{\left(\frac{e^{-X\beta}}{1+e^{-X\beta}}\right)} = \frac{1}{e^{-X\beta}} = e^{X\beta}$$

$$\Rightarrow \ln(odds) = X \cdot \beta .$$

The odds ratio for the control scenario (odds$_c$) versus the baseline (odds$_0$) is then:

$$odds\ ratio = \frac{odds_c}{odds_0} = \frac{\left(\frac{y_c}{1-y_c}\right)}{\left(\frac{y_0}{1-y_0}\right)} = \frac{\left(\frac{1}{e^{-X_c\beta}}\right)}{\left(\frac{1}{e^{-X_0\beta}}\right)} = \frac{e^{X_c\beta}}{e^{X_0\beta}} .$$

The *change* in the probability of an occurrence from the baseline to the control (Δy), assuming that all the other covariates remain constant, may be derived from this odds ratio:

$$odds\ ratio = \frac{\left(\dfrac{y_c}{1-y_c}\right)}{\left(\dfrac{y_0}{1-y_0}\right)} = \frac{e^{X_c\beta}}{e^{X_0\beta}} = \frac{e^{\gamma} \cdot e^{PM_c\beta}}{e^{\gamma} \cdot e^{PM_0\beta}} = e^{\Delta PM\cdot\beta}$$

$$\frac{y_c}{1-y_c} = \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}$$

$$y_c = (1-y_c) \cdot \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}$$

$$y_c + y_c \cdot \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta} = \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}$$

$$y_c \cdot \left[1 + \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}\right] = \left(\frac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}$$

$$y_c = \frac{\left(\dfrac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}}{1 + \left(\dfrac{y_0}{1-y_0}\right) \cdot e^{\Delta PM\cdot\beta}} = \frac{y_0 \cdot e^{\Delta PM\beta}}{1 - y_0 + y_0 \cdot e^{\Delta PM\beta}} \cdot$$

Multiplying by:

$$\frac{e^{-\Delta PM\cdot\beta}}{e^{-\Delta PM\cdot\beta}},$$

gives:

$$y_c = \frac{y_0}{(1-y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0}$$

$$\Delta y = y_c - y_0 = \frac{y_0}{(1-y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0} - y_0.$$

The change in the number of cases of the adverse health effect is then obtained by multiplying by the relevant population:

$$\Delta Incidence = \Delta y \cdot pop = \left[\frac{y_0}{(1-y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0} - y_0 \right] \cdot pop.$$

When the coefficient (β) and its standard error (σ_β) are published (e.g., Pope et al., 1991, Table 5), then the coefficient estimates associated with the lower and upper bound may be calculated easily as follows:

$$\beta_{lower\ bound} = \beta - (1.96 \cdot \sigma_\beta)$$
$$\beta_{upper\ bound} = \beta + (1.96 \cdot \sigma_\beta).$$

Often the logistic regression coefficients are not published, and only the odds ratio corresponding to a specified change in PM is presented (e.g., Schwartz et al., 1994). It is easy to calculate the underlying coefficient as follows:

$$\ln(odds\ ratio) = \Delta PM \cdot \beta$$

$$\Rightarrow \beta = \frac{\ln(odds\ ratio)}{\Delta PM}.$$

The coefficients associated with the lower and upper bound estimates of the odds ratios are calculated analogously.

The underlying standard error for the coefficient (σ_β) can be approximated by:

$$\sigma_{\beta,2.5\ percentile} = \frac{\beta - \beta_{2.5\ percentile}}{1.96}$$

$$\sigma_{\beta,97.5\ percentile} = \frac{\beta_{97.5\ percentile} - \beta}{1.96}$$

$$\sigma_{\beta} \cong \frac{\sigma_{\beta,2.5\ percentile} + \sigma_{\beta,97.5\ percentile}}{2}.$$

Sometimes, however, the relative risk is presented. The relative risk does not equal the odds ratio, and a different procedure should be used to estimate the underlying coefficient. Note that ESEERCO (1994, p. V-21) calculated (incorrectly) the underlying regression coefficient for Abbey et al. (1993, Table 5) by taking the logarithm of the relative risk and dividing by the change in TSP.

The relative risk (RR) is simply:

$$RR = \frac{y_0}{y_c},$$

where y_0 is the risk (i.e., probability of an occurrence) at the baseline PM exposure and yc is the risk at the control PM exposure.

When the baseline incidence rate (y_0) is given, then it is easy to solve for the control incidence rate (y_c):

$$y_c = \frac{y_0}{RR}.$$

The odds ratio, may then be calculated:

$$odds\ ratio = \frac{\dfrac{y_0}{1-y_0}}{\dfrac{y_c}{1-y_c}}.$$

Given the odds ratio, the underlying coefficient (β) may be calculated as before:

$$\beta = \frac{\ln(odds\ ratio)}{\Delta PM}.$$

The odds ratio and the coefficient calculated from it are dependent on the baseline and control incidence rates. Unfortunately, it is not always clear what the baseline and control incidence rates should be. Abbey et al. (1995b, Table 2) reported that there are 117 new cases of chronic bronchitis out of a sample of 1,631, or a 7.17 percent rate. In addition, they reported the relative risk (RR = 1.81) for a new case of chronic bronchitis associated with an annual mean concentration "increment" of 45 µg/m3 of $PM_{2.5}$ exposure.

Assuming that the baseline rate for chronic bronchitis (y_0) should be 7.17 percent, the question becomes whether the "increment" of 45 µg/m3 should be added to or subtracted from the existing $PM_{2.5}$ concentration. If added the control incidence rate (y_c) would be greater than the baseline rate (y_0), while subtraction would give a control rate less than the incidence rate. In effect, one might reasonably derive two estimates of the odds ratio:

$$odds\ ratio_1 = \frac{\left(\frac{y_0}{1-y_0}\right)}{\left(\frac{y_c}{1-y_c}\right)} = \frac{\left(\frac{1.81 \cdot 0.0717}{1-(1.81 \cdot 0.0717)}\right)}{\left(\frac{0.0717}{1-0.0717}\right)} = 1.931$$

$$odds\ ratio_2 = \frac{\left(\frac{y_0}{1-y_0}\right)}{\left(\frac{y_c}{1-y_c}\right)} = \frac{\left(\frac{0.0717}{1-0.0717}\right)}{\left(\frac{\frac{0.0717}{1.81}}{1-\frac{0.0717}{1.81}}\right)} = 1.873$$

$$\Rightarrow \beta_1 = \frac{\ln(1.931)}{45} = 0.01462$$

$$\Rightarrow \beta_2 = \frac{\ln(1.873)}{45} = 0.01394.$$

An alternative is to simply assume that the relative risk (1.81) is reasonably close to the odds ratio and calculate the underlying coefficient. It is easy to show that the relative risk equals:

$$RR = \frac{y_0}{y_c} = (1 - y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0 \ .$$

Assuming that:

$$e^{-\Delta PM \cdot \beta} \cong (1 - y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0$$

$$\Rightarrow RR \cong e^{-\Delta PM \cdot \beta} \ .$$

It is then possible to calculate the underlying coefficient:

$$\frac{\ln(RR)}{-\Delta PM} \cong \beta$$

$$\Rightarrow \beta_3 = \frac{\ln(1.81)}{45} = 0.01319 \ .$$

Since this coefficient estimate is based on the assumption that

$$e^{-\Delta PM \cdot \beta} \cong (1 - y_0) \cdot e^{-\Delta PM \cdot \beta} + y_0 \ ,$$

it should be used in a C-R function that maintains this assumption. In effect, it should be applied to a log-linear C-R function:

$$\Delta y = \left[y_0 \cdot (e^{\beta \Delta PM} - 1) \right] \ .$$

Using the formula for the change in the incidence rate and assuming a 10 µg/m3 decline in $PM_{2.5}$, it is shown that this results in changes within the bounds suggested by the two estimates based on using the estimated odds ratios:

$$\Delta y_1 = \frac{.0717}{(1-0.0717) \cdot e^{100.01482} + 0.0717} - 0.0717 = -0.00914$$

$$\Delta y_2 = \frac{.0717}{(1-0.0717) \cdot e^{100.01394} + 0.0717} - 0.0717 = -0.00874$$

$$\Delta y_3 = 0.0717 \cdot \left(e^{-100.01319} - 1 \right) = -0.00886.$$

In this instance, it seems that simply using the relative risk to estimate the underlying coefficient results in a good approximation of the change in incidence. Since it is unclear which of the two other coefficients (β_1 or β_2) should be used -- as the published work was not explicit – the coefficient based on the relative risk and the log-linear functional form seems like a reasonable approach.

D.6 Cox proportional Hazards Model

Use of a Cox proportional hazards model in an epidemiological study results in a C-R function that is log-linear in form. It is often used to model survival times, and as a result, this discussion focuses on mortality impacts.

The Cox proportional hazards model is based on a hazard function, defined as the probability that an individual dies at time t, conditional on having survived up to time t (Collet, 1994, p. 10). More formally, the hazard function equals the probability density function for the risk of dying divided by one minus the cumulative probability density function:

$$h(X,t) = \frac{f(X,t)}{1-F(X,t)}.$$

The proportional hazards model takes the form:

$$h(X,t) = h_0(t)e^{X \cdot \beta}$$

where X is a vector of explanatory variables, β is a vector of coefficients, and $h_0(t)$ is the so-called "baseline hazard" rate. This terminology differs from that used in most of this discussion: this "baseline hazard" is the risk when all of the covariates (X) are set to zero; this is *not* the risk in the baseline scenario.

The Cox proportional hazards model is sometimes termed a "semi-parametric" model, because the baseline hazard rate is calculated using a non-parametric method, while the

impact of explanatory variables is parameterized. Collet (1994) details the estimation of Cox proportional hazards models; in particular, see Collet's discussion (pp. 95-97) of nonparametric estimation of the baseline hazard.

Taking the ratio of the hazard functions for the baseline and control scenarios gives the relative risk:

$$RR = \frac{h(X_0, t)}{h(X_c, t)} = \frac{h_0(t) e^{X_0 \cdot \beta}}{h_0(t) e^{X_c \cdot \beta}} = e^{\Delta PM \cdot \beta}.$$

where it is assumed that the only difference between the baseline and control is the level of PM pollution.

The relative risk is often presented rather than the coefficient β, so it is necessary to estimate β in order to develop functional relationship between ΔPM and Δy, as described previously for log-linear C-R functions.

Appendix E: Health Incidence & Prevalence Data in U.S. Setup

Health impact functions developed from log-linear or logistic models estimate the percent change in an adverse health effect associated with a given pollutant change. In order to estimate the absolute change in incidence using these functions, we need the baseline incidence rate of the adverse health effect. And for certain health effects, such as asthma exacerbation, we need a prevalence rate, which estimates the percentage of the general population with a given ailment like asthma. This appendix describes the data used to estimate baseline incidence and prevalence rates for the health effects considered in this analysis.

E.1   Mortality

This section describes the development of the year 2000 through 2050 county mortality rates for use in BenMAP. First, we describe the source of 1996-1998 county-level mortality rates, and then we describe how we use national-level Census mortality rate projections to develop 2000-2050 county-level mortality rate projections.

E.1.1   Mortality Rates 1996-1998

Age, cause, and county-specific mortality rates were obtained from the U.S. Centers for Disease Control (CDC) for the years 1996 through 1998. CDC maintains an online data repository of health statistics, CDC Wonder, accessible at http://wonder.cdc.gov/. The mortality rates provided are derived from U.S. death records and U.S. Census Bureau postcensal population estimates. Mortality rates were averaged across three years (1996 through 1998) to provide more stable estimates. Population-weighted national mortality rates are presented in Table E-1.

Table E-1. Population-Weighted Mortality Rates (per 100 people per year) for Selected Conditions, by Age Group

Mortality Category (ICD codes)	0-17	18-24	25-29	30-34	35-44	45-54	55-64	65-74	75-84	85+
All-Cause	0.045	0.093	0.119	0.119	0.211	0.437	1.056	2.518	5.765	15.160
Non-Accidental (ICD <800)	0.025	0.022	0.057	0.057	0.150	0.383	1.006	2.453	5.637	14.859
Chronic Lung Disease (ICD 490-496)	0.000	0.001	0.001	0.001	0.002	0.009	0.046	0.166	0.367	0.561
Cardio-Pulmonary	0.004	0.005	0.013	0.013	0.044	0.143	0.420	1.163	3.179	9.846

E.1.2   Mortality Rate Projections 2000-2050

To estimate age- and county-specific mortality rates in years 2000 through 2050, we calculated adjustment factors, based on a series of Census Bureau projected national mortality rates, to adjust the CDC Wonder age- and county-specific mortality rates in 1996-1998 to corresponding rates for each future year. The procedure we used was as follows:

For each age group, we derived an estimate of the national mortality rate in 1997 (the midpoint year in the period 1996 - 1998) consistent with the series of Census Bureau projected national mortality rates, which starts in 1999. We did this by regressing projected mortality rate on year, separately for each age group, using the ten years of Census Bureau projected rates from 1999 - 2008. The resulting estimated national age-group-specific mortality rates for 1997 are shown in Table E-2. Note that the Census Bureau projected mortality rates were derived from crude death rates using the following formula, given by Chiang (1967, p.2 equation 7): $M = Q/(1-(1-A)*Q)$, where M denotes projected mortality rate, Q denotes crude death rate, and A denotes the fraction of the interval (one year) lived by individuals who die in the interval. A=0.1 if age < 1, and A=0.5 otherwise.

We then calculated, separately for each age-group, the ratio of Census Bureau national mortality rate in year Y (Y = 2000, 2001, ..., 2050) to the national mortality rate in 1997, estimated in the previous step to be consistent with the Census Bureau series of rates starting in 1999. These ratios are shown for selected years in Table E-3.

Finally, to estimate mortality rates in year Y (Y = 2000, 2001, ..., 2050) that are both age-group-specific and county-specific, we multiplied the CDC Wonder county-specific age-group-specific mortality rates for 1996-1998 by the appropriate ratio calculated in the previous step. For example, to estimate the projected mortality rate in 2010 among ages 18-24 in Wayne County, MI, we multiplied the CDC Wonder mortality rate for ages 18-24 in Wayne County in 1996-1998 by the ratio of Census Bureau projected national mortality rate in 2010 for ages 18-24 to (estimated) Census Bureau national mortality rate in 1997 for ages 18-24.

Table E-2. All-Cause Mortality Rate (per 100 people per year), by Source, Year, and Age Group

Source & Year	Infant *	1-17	18-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
Census Bureau 2000	0.687	0.030	0.093	0.106	0.192	0.408	0.998	2.454	5.636	13.541
Est. Census Bureau 1997	0.706	0.031	0.095	0.108	0.199	0.421	1.032	2.555	5.787	13.846
CDC Wonder 1996-1998	0.246	0.034	0.093	0.119	0.211	0.437	1.056	2.518	5.765	15.160
Estimated 2000 **	0.239	0.033	0.091	0.116	0.204	0.424	1.022	2.419	5.615	14.826

* Note that the Census Bureau estimate is for all deaths in the first year of life. The CDC Wonder estimate if for post-neonatal mortality (deaths after the first month), because the health impact function (see Appendix F)

estimates post-neonatal mortality.

** The estimate for 2000 is a population-weighted average of the county-level forecasts for 2000 that are calculated from the CDC Wonder county-level estimates and the ratio of the Census Bureau estimates for 2000 and 1997.

Table E-3. Ratio of Future Year All-Cause Mortality Rate to 1997 Estimated All-Cause Mortality Rate, by Age Group

Year	Infant	1-17	18-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
2000	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.96	0.97	0.98
2005	0.93	0.94	0.93	0.95	0.92	0.92	0.90	0.90	0.93	0.95
2010	0.88	0.88	0.88	0.91	0.87	0.88	0.86	0.84	0.89	0.91
2015	0.83	0.81	0.84	0.88	0.82	0.83	0.82	0.79	0.83	0.89
2020	0.78	0.76	0.79	0.86	0.77	0.78	0.78	0.76	0.77	0.86
2025	0.72	0.71	0.75	0.80	0.73	0.73	0.74	0.72	0.73	0.82
2030	0.66	0.66	0.70	0.75	0.68	0.68	0.69	0.70	0.71	0.77
2035	0.61	0.61	0.66	0.70	0.64	0.64	0.65	0.67	0.68	0.72
2040	0.56	0.56	0.62	0.66	0.60	0.60	0.60	0.63	0.65	0.70
2045	0.51	0.52	0.58	0.62	0.56	0.57	0.57	0.60	0.63	0.69
2050	0.47	0.48	0.55	0.58	0.53	0.53	0.54	0.56	0.59	0.68

E.2   Hospitalizations

Regional hospitalization counts were obtained from the National Center for Health Statistics' (NCHS) National Hospital Discharge Survey (NHDS). NHDS is a sample-based survey of non-Federal, short-stay hospitals (<30 days), and is the principal source of nationwide hospitalization data. The survey collects data on patient characteristics, diagnoses, and medical procedures. However, note that the following hospital types are excluded from the survey: hospitals with an average patient length of stay of greater than 30 days, federal, military, Department of Veterans Affairs hospitals, institutional hospitals (e.g. prisons), and hospitals with fewer than six beds.

Public use data files for the year 1999 survey were downloaded (from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHDS/) and processed to estimate hospitalization counts by region. NCHS groups states into four regions using the following groupings defined by the U.S. Bureau of the Census:

Northeast - Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, Pennsylvania

Midwest - Ohio, Indiana, Illinois, Michigan, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, Kansas

South - Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, Texas

West - Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, Washington, Oregon, California, Alaska, Hawaii

We calculated per capita hospitalization rates, by dividing these counts by the estimated regional population estimates for 1999 that we derived from the U.S. Bureau of the Census and the population projections used by NHDS to generate the counts. Note that NHDS started with hospital admission counts, based on a sample of admissions, and then they used population estimates to generate population-weighted hospital admission counts that are representative of each region. This weighting used forecasts of 1999 population data. Ideally, we would use these same forecasts to generate our admission rates. However, while NHDS presented counts of hospital admissions with a high degree of age specificity, it presented regional population data for only four age groups: 0-14, 15-44, 45-64, and 65+. Using only the NHDS data, we would be limited to calculating regional admission rates for four groups. Because we are interested in a broader range of age groups, we turned to 2000 Census.

We used the 2000 Census to obtain more age specificity, and then corrected the 2000 Census figures so that the total population equaled the total for 1999 forecasted by NHDS. That is, we sued the following procedure: (1) we calculated the count of hospital admissions by region in 1999 for the age groups of interest, (2) we calculated the 2000 regional populations corresponding to these age groups, (3) calculated regional correction factors, that equal the regional total population in 1999 divided by the regional total population in 2000 by region, (4) multiplied the 2000 population estimates by these correction factors, and (5) divided the 1999 regional count of hospital admissions by the estimated 1999 population.

The endpoints in hospitalization studies are defined using different combinations of ICD codes. Rather than generating a unique baseline incidence rate for each ICD code combination, for the purposes of this analysis, we identified a core group of hospitalization rates from the studies and applied the appropriate combinations of these rates in the health impact functions:

all respiratory (ICD-9 460-519)

chronic lung disease (ICD-9 490-496)

asthma (ICD-9 493)

pneumonia (ICD-9 480-487)

acute bronchitis (ICD-9 466)

acute laryngitis (ICD-9 464)

all cardiovascular (ICD-9 390-459)

ischemic heart disease (ICD-9 410-414)

dysrhythmia (ICD-9 427)

congestive heart failure (ICD-9 428)

For each C-R function, we selected the baseline rate or combination of rates that most closely matches to the study endpoint definition. For studies that define chronic lung disease as ICD 490-492, 494-496, we subtracted the incidence rate for asthma (ICD 493) from the chronic lung disease rate (ICD 490-496). In some cases, the baseline rate will not match exactly to the endpoint definition in the study. For example, Burnett et al. (2001) studied the following respiratory conditions in infants <2 years of age: ICD 464.4, 466, 480-486, 493. For this C-R function we apply an aggregate of the following rates: ICD 464, 466, 480-487, 493. Although they do not match exactly, we assume that relationship observed between the pollutant and study-defined endpoint is applicable for the additional codes. Table E-4 presents a summary of the national hospitalization rates for 1999 from NHDS.

Table E-4. Hospitalization Rates (per 100 people per year), by Region and Age Group

Hospital Admission Category		ICD-9 Code	0-18	18-24	25-34	35-44	45-54	55-64	65+
Respiratory	all respiratory	460-519	1.066	0.271	0.318	0.446	0.763	1.632	5.200
	acute laryngitis	464	0.055	0.002	0.001	0.002	0.008	0.000	0.005
	acute bronchitis	466	0.283	0.017	0.014	0.017	0.027	0.040	0.156
	pneumonia	480-487	0.308	0.069	0.103	0.155	0.256	0.561	2.355
	asthma	493	0.281	0.081	0.110	0.099	0.144	0.161	0.205
	chronic lung disease	490-496	0.291	0.089	0.124	0.148	0.301	0.711	1.573
Cardiovascular	all cardiovascular	390-429	0.030	0.052	0.146	0.534	1.551	3.385	8.541
	ischemic heart disease	410-414	0.004	0.008	0.031	0.231	0.902	2.021	3.708
	dysrhythmia	427	0.011	0.017	0.027	0.076	0.158	0.392	1.387
	congestive heart failure	428	0.003	0.005	0.011	0.011	0.160	0.469	2.167

E.3 Emergency Room Visits for Asthma

Regional asthma emergency room visit counts were obtained from the National Hospital Ambulatory Medical Care Survey (NHAMCS). NHAMCS is a sample-based survey, conducted by NCHS, designed to collect national data on ambulatory care utilization in hospital emergency and outpatient departments of non-Federal, short-stay hospitals (<30 days). The target universe of the NHAMCS is in-person visits made in the United States to emergency and outpatient departments of non-Federal, short-stay hospitals (hospitals with an average stay of less than 30 days) or those whose specialty is general (medical or surgical) or children's general.

Public use data files for the year 2000 survey were downloaded (from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHAMCS/) and processed to estimate hospitalization counts by region. We obtained population estimates from the 2000 U.S. Census. The NCHS regional groupings described above were used to estimate regional emergency room visit rates. Table E-5 presents the estimated asthma emergency room rates by region.

Table E-5. Emergency Room Visit Rates (per 100 people per year) for Asthma, by Region and Age Group

ER Category	ICD-9 Code	Region	0-18	18-64	65+
asthma	493	Northeast	0.761	0.802	0.300
		Midwest	1.476	0.877	0.334
		South	1.243	0.420	0.192
		West	0.381	0.381	0.137

E.4   Nonfatal Heart Attacks

The relationship between short-term particulate matter exposure and heart attacks was quantified in a case-crossover analysis by Peters et al. (2001). The study population was selected from heart attack survivors in a medical clinic. Therefore, the applicable population to apply to the C-R function is all individuals surviving a heart attack in a given year. Several data sources are available to estimate the number of heart attacks per year. For example, several cohort studies have reported estimates of heart attack incidence rates in the specific populations under study. However, these rates depend on the specific characteristics of the populations under study and may not be the best data to extrapolate nationally. The American Heart Association reports approximately 540,000 new heart attacks per year using data from a multi-center study (Haase, 2002, to be published in the American Heart Association's 2003 Statistical Handbook). Exclusion of heart attack deaths reported by CDC Wonder yields approximately 330,000 nonfatal cases per year.

An alternative approach to the estimation of heart attack rates is to use data from the National Hospital Discharge Survey, assuming that all heart attacks that are not instantly fatal will result in a hospitalization. According to the National Hospital Discharge Survey, in 1999 there were approximately 829,000 hospitalizations due to heart attacks (acute myocardial infarction: ICD-9 410) (Popovic, 2001, Table 8). We used regional hospitalization rates over estimates extrapolated from cohort studies because the former is part of a nationally representative survey with a larger sample size, which is intended to provide reliable national estimates. As additional information is provided regarding the American Heart Association methodology, we will evaluate the usefulness of this estimate of heart attack incidence.

Rosamond et al. (1999) reported that approximately six percent of male and eight percent of female hospitalized heart attack patients die within 28 days (either in or outside of the hospital). We, therefore, applied a factor of 0.93 to the count of hospitalizations to estimate the number of nonfatal heart attacks per year. To estimate the rate of nonfatal heart attack, we divided the count by the population estimate for 2000 from the U.S. Census. Table E-6 presents the regional nonfatal heart attack incidence rates.

Table E-6. Nonfatal Heart Attack Rates (per 100 people per year), by Region and Age Group

Effect	Region	0-18	18-64	65+
nonfatal heart attacks (ICD-9 410)	Northeast	0.0000	0.2167	1.6359
	Midwest	0.0003	0.1772	1.4898
	South	0.0006	0.1620	1.1797
	West	0.0000	0.1391	1.1971

* Rates are based on data from the 1999 National Hospital Discharge Survey (NHDS) and an estimate from Rosamond et al. that approximately 7% of individuals hospitalized for a heart attack die within 28 days.

E.5 School Loss Days

Epidemiological studies have examined the relationship between air pollution and a variety of measures of school absence. These measures include: school loss days for all causes, illness-related, and respiratory illness-related. We have two sources of information. The first is the National Center for Education Statistics, which provided an estimate of all-cause school loss days, and the other is the National Health Interview Survey (Adams et al., 1999, Table 47), which has data on different categories of acute school loss days. Table E-7 presents the estimated school loss day rates. Further detail is provided below on these rates.

Table E-7. School Loss Day Rates (per student per year)

Type	Northeast	Midwest	South	West
Respiratory illness-related absences	1.3	1.7	1.1	2.2
Illness-related absences	2.4	2.6	2.6	3.7
All-cause	9.9	9.9	9.9	9.9

* We based illness-related school loss day rates on data from the

1996 NHIS and an estimate of 180 school days per year. This
excludes school loss days due to injuries. We based the all-cause
school loss day rate on data from the National Center for
Education Statistics.

All-Cause School Loss Day Rate

Based on data from the U.S. Department of Education (1996, Table 42-1), the National
Center for Education Statistics estimates that for the 1993-1994 school year, 5.5 percent of
students are absent from school on a given day. This estimate is comparable to
study-specific estimates from Chen et al. (2000) and Ransom and Pope (1992), which
ranged from 4.5 to 5.1 percent.

Illness-Related School Loss Day Rate

The National Health Interview Survey (NHIS) has regional estimates of school loss days
due to a variety of acute conditions (Adams et al., 1999). NHIS is a nationwide
sample-based survey of the health of the noninstitutionalized, civilian population,
conducted by NCHS. The survey collects data on acute conditions, prevalence of chronic
conditions, episodes of injury, activity limitations, and self-reported health status.
However, it does not provide an estimate of all-cause school loss days.

In estimating illness-related school loss days, we started with school loss days due to acute
problems (Adams et al., 1999, Table 47) and subtracted lost days due to injuries, in order
to match the definition of the study used in the C-R function to estimate illness-related
school absences (Gilliland et al., 2001). We then divided by 180 school days per to
estimate *illness*-related school absence rates per school day. Similarly, when estimating
respiratory illness-related school loss days, we use data from Adams et al. (1999, Table
47). Note that we estimated 180 school days in a year to calculate respiratory
illness-related school absence rates per year.

E.6   Other Acute and Chronic Effects

For many of the minor effect studies, baseline rates from a single study are often the only
source of information, and we assume that these rates hold for locations in the U.S. The
use of study-specific estimates are likely to increase the uncertainty around the estimate
because they are often estimated from a single location using a relatively small sample.
These endpoints include: acute bronchitis, chronic bronchitis, upper respiratory symptoms,
lower respiratory symptoms. Table E-8 presents a summary of these baseline rates.

**Table E-8. Selected Acute and Chronic Incidence (Cases / Person-Year)
& Prevalence (Percentage Population)**

Endpoint	Age	Parameter	Rate	Source
Acute Bronchitis	8-12	Incidence	0.043	American Lung Association (2002b, Table 11)
Chronic Bronchitis	27+	Incidence	0.00378	Abbey et al (1993, Table 3)
Chronic Bronchitis	18+	Prevalence	4.43%	
	18-44		3.67%	American Lung Association (2002a, Table 4)
	45-64		5.05%	
	65+		5.87%	
Lower Respiratory Symptoms (LRS)	7-14	Incidence	0.438	Schwartz et al (1994, Table 2)
Minor Restricted Activity Days (MRAD)	18-64	Incidence	7.8	Ostro and Rothschild (1989, p. 243)
Work Loss Day (WLD)	18-64	Incidence	2.172	
	18-24		1.971	Adams et al (1999, Table)
	25-44		2.475	U.S. Bureau of the Census (1997, No. 22)
	45-64		1.796	

NOTE: The incidence rate is the number of cases <u>per person</u> per year. Prevalence refers to the fraction of people that have a particular illness during a particular time period.

E.6.1 Acute Bronchitis

The annual rate of acute bronchitis for children ages 5 to 17 was obtained from the American Lung Association (2002b, Table 11). The authors reported an annual incidence rate per person of 0.043, derived from the 1996 National Health Interview Survey.

E.6.2 Chronic Bronchitis Incidence Rate

The annual incidence rate for chronic bronchitis is estimated from data reported by Abbey et al. (1993, Table 3). The rate is calculated by taking the number of new cases (234), dividing by the number of individuals in the sample (3,310), dividing by the ten years covered in the sample, and then multiplying by one minus the reversal rate (estimated to be 46.6% based on Abbey et al. (1995a, Table 1).

Age-specific incidence rates are not available. Abbey et al. (1995a, Table 1) did report the incidences by three age groups (25-54, 55-74, and 75+) for "cough type" and "sputum type" bronchitis. However, they did not report an overall incidence rate for bronchitis by age-group. Since, the cough and sputum types of bronchitis overlap to an unknown extent, we did not attempt to generate age-specific incidence rates for the over-all rate of

bronchitis.

E.6.3   Chronic Bronchitis Prevalence Rate

We obtained the annual prevalence rate for chronic bronchitis from the American Lung Association (2002a, Table 4). Based on an analysis of 1999 National Health Interview Survey data, they estimated a rate of 0.0443 for persons 18 and older, they also reported the following prevalence rates for people in the age groups 18-44, 45-64, and 65+: 0.0367, 0.0505, and 0.0587, respectively.

E.6.4   Lower Respiratory Symptoms

Lower respiratory symptoms (LRS) are defined as two or more of the following: cough, chest pain, phlegm, wheeze. The proposed yearly incidence rate for 100 people, 43.8, is based on the percentiles in Schwartz et al. (Schwartz et al., 1994, Table 2). The authors did not report the mean incidence rate, but rather reported various percentiles from the incidence rate distribution. The percentiles and associated per person per day values are 10 th = 0 percent, 25th = 0 percent, 50th = 0 percent, 75th = 0.29 percent, and 90th = 0.34 percent. The most conservative estimate consistent with the data are to assume the incidence per person per day is zero up to the 75th percentile, a constant 0.29 percent between the 75th and 90th percentiles, and a constant 0.34 percent between the 90th and 100 th percentiles. Alternatively, assuming a linear slope between the 50th and 75th, 75th and 90 th, and 90th to 100th percentiles, the estimated mean incidence rate per person per day is 0.12 percent. (For example, the 62.5th percentile would have an estimated incidence rate per person per day of 0.145 percent.) We used the latter approach in this analysis.

E.6.5   Minor Restricted Activity Days (MRAD)

Ostro and Rothschild (1989, p. 243) provide an estimate of the annual incidence rate of MRADs per person of 7.8.

E.6.6   Work Loss Days

The yearly work-loss-day incidence rate per 100 people is based on estimates from the 1996 National Health Interview Survey (Adams et al., 1999, Table 41). They reported a total annual work loss days of 352 million for individuals ages 18 to 65. The total population of individuals of this age group in 1996 (162 million) was obtained from (U.S. Bureau of the Census, 1997, No. 22). The average annual rate of work loss days per individual is 2.17. Using a similar approach, we calculated work-loss-day rates for ages 18-24, 25-44, and 45-64, respectively.

E.7   Asthma-Related Health Effects

Several studies have examined the impact of air pollution on asthma development or exacerbation. Many of the baseline incidence rates used in the C-R functions are based on study-specific estimates. The baseline rates for the various endpoints are described below

and summarized in TableE-9.

Table E-9. Asthma-Related Health Effects Rates

Endpoint	Age	Parameter a	Rate	Source
Asthma Exacerbation, Shortness of Breath, African American	8-13	Incidence	13.51	Ostro et al (2001, p. 202)
	8-13	Prevalence	7.40%	
Asthma Exacerbation, Wheeze, African American	8-13	Incidence	27.74	Ostro et al (2001, p. 202)
	8-13	Prevalence	17.30%	
Asthma Exacerbation, Cough, African American	8-13	Incidence	24.46	Ostro et al (2001, p. 202)
	8-13	Prevalence	14.50%	
Asthma Exacerbation, Cough	6-13	Incidence	31.39	Vedal et al (1998, Table 1)
Upper Respiratory Symptoms (URS)	9-11	Incidence	124.79	

NOTE: The incidence rate is the number of cases <u>per person</u> per year. Prevalence refers to the fraction of people that have a particular illness during a particular time period.

E.7.1 Shortness of Breath

To estimate the annual rate of new shortness of breath episodes among African-American asthmatics, ages 8-13, we used the rate reported by Ostro et al. (2001, p.202). We estimated the daily prevalence of shortness of breath episodes among African-American asthmatics, ages 8-13, by taking a weighted average of the reported rates in Ostro et al. (2001, p.202).

E.7.2 Wheeze

The daily rate of new wheeze episodes among African-American asthmatics, ages 8-13, is reported by Ostro et al. (2001, p.202) as 0.076. We multiplied this value by 100 and by 365 to get the annual incidence rate per 100 people. The daily rate of prevalent wheeze episodes (0.173) among African-American asthmatics, ages 8-13, is estimated by taking a weighted average of the reported rates in Ostro et al. (2001, p.202)

E.7.3 Cough

The daily rate of new cough episodes among African-American asthmatics, ages 8-13, is reported by Ostro et al. (2001, p.202) as 0.067. We multiplied this value by 100 and by 365 to get the annual incidence rate per 100 people. The daily rate of prevalent cough episodes (0.145) among African-American asthmatics, ages 8-13, is estimated by taking a weighted average of the reported rates in Ostro et al. (2001, p.202).

E.7.4 Upper Respiratory Symptoms

Upper Respiratory Symptoms are defined as one or more of the following: runny or stuffy nose; wet cough; burning, aching, or red eyes.   Using the incidence rates for upper respiratory symptoms among asthmatics, published in Pope et al. (1991, Table 2), we calculated a sample size-weighted average incidence rate.

E.7.5 Asthma Population Estimates

In studies examining the association between air pollution and the development or exacerbation of asthma, often times an estimate of the percent of the population with asthma is required. Asthma percentages were obtained either directly from the National Health Interview Survey (NHIS) or an American Lung Association (2002c) report summarizing data from NHIS. Table E-10 presents asthma prevalence rates used to define asthmatic populations in the health impact functions.

Table E-10. Asthma Prevalence Rates Used to Estimate Asthmatic Populations

Population Group	Prevalence	Source
All Ages	3.86%	
<18	5.27%	
5-17	5.67%	American Lung Association (2002c) *
18-44	3.71%	
45-64	3.33%	
65+	2.21%	
African-American, 5 to 17	7.26%	American Lung Association (2002c) *
African-American, <18	7.35%	
Male, 27+	2.10%	2000 NHIS public use data files **

* American Lung Association (2002c) is based on the 1999 National Health Interview Survey (Adams et al, 1999).

** See ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/

Appendix F: Particulate Matter Health Impact Functions in U.S. Setup

In this Appendix, we present the PM-related health impact functions in BenMAP. Each sub-section has a table with a brief description of the health impact function and the underlying parameters. Following each table, we present a brief summary of each of the studies and any items that are unique to the study.

Note that Appendix D mathematically derives the standard types of health impact functions encountered in the epidemiological literature, such as, log-linear, logistic and linear, so we simply note here the type of functional form. And Appendix E presents a description of the sources for the incidence and prevalence data used in the health impact functions.

F.1   Long-term Mortality

There are two types of exposure to PM that may result in premature mortality. Short-term exposure may result in excess mortality on the same day or within a few days of exposure. Long-term exposure over, say, a year or more, may result in mortality in excess of what it would be if PM levels were generally lower, although the excess mortality that occurs will not necessarily be associated with any particular episode of elevated air pollution levels. In other words, long-term exposure may capture a facet of the association between PM and mortality that is not captured by short-term exposure. Table F-1 lists the long-term mortality health impact functions.

Table F-1. Health Impact Functions for Particulate Matter and Long-Term Mortality

Effect	Author	Year	Location	Age	Metric	Beta	Std Err	Form	Notes
Mortality, All Cause	Expert A	2006		30-99	Annual	0.015180		Log-linear	
Mortality, All Cause	Expert B	2006		30-99	Annual	0.012620		Log-linear	Range >10 to 30 ug. Unconditional dist. 2% no causality included.
Mortality, All Cause	Expert B	2006		30-99	Annual	0.011950		Log-linear	Range 4 to 10 ug. Unconditional dist. 2% no causality included.
Mortality, All Cause	Expert C	2006		30-99	Annual	0.011930		Log-linear	
Mortality, All Cause	Expert D	2006		30-99	Annual	0.008380		Log-linear	Unconditional dist. 5% no causality included.
Mortality, All Cause	Expert E	2006		30-99	Annual	0.019670		Log-linear	Unconditional dist. 1% no causality included.
Mortality, All Cause	Expert F	2006		30-99	Annual	0.011440		Log-linear	Range >7 to 30 ug
Mortality, All Cause	Expert F	2006		30-99	Annual	0.009370		Log-linear	Range 4 to 7 ug
Mortality, All Cause	Expert G	2006		30-99	Annual	0.006970		Log-linear	Unconditional dist. 30% no causality included.

Effect	Author	Year	Location	Age	Metric	Beta	Std Err	Form	Notes
Mortality, All Cause	Expert H	2006		30-99	Annual	0.008700		Log-linear	
Mortality, All Cause	Expert I	2006		30-99	Annual	0.011810		Log-linear	Unconditional dist. 5% no causality included.
Mortality, All Cause	Expert J	2006		30-99	Annual	0.009620		Log-linear	
Mortality, All Cause	Expert K	2006		30-99	Annual	0.006890		Log-linear	Range >16 to 30. No threshold. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0.003940		Log-linear	Range 4 to 16 ug. No threshold. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0.003940		Log-linear	Range 4 to 16 ug. Threshold 0 to 5 ug. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0.003940		Log-linear	Range 4 to 16 ug. Threshold 5 to 10 ug. Conditional dist.
Mortality, All Cause	Expert L	2006		30-99	Annual	0.009340		Log-linear	Range >10 to 30 ug. Unconditional dist. 1% no causality included.
Mortality, All Cause	Expert L	2006		30-99	Annual	0.007390		Log-linear	Range 4 to 10 ug. Unconditional dist. 25% no causality included.
Mortality, All Cause	Laden et al.	2006	6 cities	25-99	Annual	0.014842	0.004170	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Mortality, All Cause	Laden et al.	2006	6 cities	25-99	Annual	0.014842	0.004170	Log-linear	
Mortality, All Cause	Pope et al.	2002	51 cities	30-99	Annual	0.006555	0.002427	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Mortality, All Cause	Pope et al.	2002	51 cities	30-99	Annual	0.007284	0.002696	Log-linear	Adjusted coefficient with 12 ug/m3 threshold.
Mortality, All Cause	Pope et al.	2002	51 cities	30-99	Annual	0.008740	0.003236	Log-linear	Adjusted coefficient with 15 ug/m3 threshold.
Mortality, All Cause	Pope et al.	2002	51 cities	30-99	Annual	0.005827	0.002157	Log-linear	Adjusted coefficient with 7.5 ug/m3 threshold.
Mortality, All Cause	Pope et al.	2002	51 cities	30-99	Annual	0.005827	0.002157	Log-linear	
Mortality, All Cause	Woodruff et al.	1997	86 cities	Infant	Annual	0.003922	0.001221	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Mortality, All Cause	Woodruff et al.	1997	86 cities	Infant	Annual	0.003922	0.001221	Logistic	
Mortality, All Cause	Woodruff et al.	2006	204 counties	Infant	Annual	0.006766	0.007339	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Mortality, All Cause	Woodruff et al.	2006	204 counties	Infant	Annual	0.006766	0.007339	Logistic	

F.1.1 Expert Functions

In this section, we describe the approach taken to incorporate into BenMAP concentration-response (C-R) functions that were obtained through expert elicitation for EPA (IEc, 2006).

We have specified expert distributions for the PM2.5 effect either as truncated parametric distributions or as non-parametric distributions. Therefore they can only be included in

BenMAP in the form of custom distribution tables containing 15,000 random draws (with replacement) from an underlying distribution. We first describe the way these custom distribution tables were created. Then we explain how these custom distribution tables should be handled in a configuration file to represent the expert-specified distribution as closely as possible.

Note that the table on page 3-30 of the expert elicitation report (IEc, 2006) refers to the non-parametric distributions as "custom" distributions. However, BenMAP refers to distribution tables that are supplied in the form of a simulated draw as "custom distribution tables". In order to avoid confusion in terminology, we will call the expert-specified distributions, which did not have a parametric shape, "non-parametric" expert distributions.

We divided the experts into two groups – those who specified a parametric distribution and those who specified a non-parametric distribution. This division was necessary because the two groups required different methods for generating the custom distribution tables. We describe the respective algorithms below and then provide an assessment of the results for each expert.

F.1.1.1 Parametric Distributions

Experts A, C, D, E, G, I, J, and K chose parametric distribution functions to represent their subjective beliefs about the percent change in risk associated with an increase in PM2.5. In particular, they specified the following characteristics of the distribution:

The shape (e.g., Normal, Triangular, Weibull)

The truncation points (i.e., minimum and/or maximum)

Two or three percentile points

The likelihood that the association is causal and whether the function includes that (i.e., whether the function is conditional on the association being causal or unconditional).

There were two types of inconsistencies encountered in these specifications:

(1) The experts who chose Normal or Weibull shapes for their distributions also specified minimum and/or maximum values at which there could be an effect. The Normal distribution has an unlimited support from $-\infty$ to $+\infty$. The Weibull distribution has support $(l, +\infty)$, where l is a location parameter that can be any value on the real line. The specification of a minimum or a maximum value for the effect is therefore inconsistent with specifying these distributions. Therefore, we interpreted these experts' distributions as truncated Normal or truncated Weibull distributions. In other words, we assumed that the shape of the distribution is Normal or Weibull between the truncation points.

(2) Experts A, C, and J indicated that they included the likelihood of causality in their subjective distributions. However, the continuous parametric distributions specified were inconsistent with the causality likelihoods provided by these experts. Because there was no way to reconcile this, we chose to interpret the distributions of these experts as

unconditional and ignore the additional information on the likelihood of causality. For example, Expert A specified a truncated Normal distribution with a minimum 0 and a maximum 4. The expert also indicated that the likelihood of causality is 95 percent and it is included in the distribution. This implies that the 5th percentile of the truncated Normal distribution should be zero. The minimum and 5th percentile of the distribution both being zero imply a density with a large (discrete) mass at zero. This, however, is not consistent with specifying a continuous Normal density. (In the case of Expert A, in addition, he specified a 5th percentile value of 0.29, whereas a 5 percent chance of non-causality would imply a 5th percentile value of 0.)

In order to create a random draw from a parametric distribution it is not sufficient to know its shape and truncation points. In addition, one needs to know the values of parameters that distinguish this particular distribution from a class of similarly shaped distributions with identical truncation points. Experts D and I reported parameter values of their subjective distributions (see details in Table 1). Therefore, we simply drew 15,000 times from each of their distributions.

However, the only information, in addition to the shape and truncation points, which the other experts provided was the percentile points. To derive the parameter values of interest, we used this information as follows:

Let $F(x;\boldsymbol{\theta},\mathbf{min},\mathbf{max})$ be a truncated continuous parametric (cumulative) distribution function with (vector of) parameters $\boldsymbol{\theta}$ and truncation points $\mathbf{min}$ and $\mathbf{max}$. The nth percentile point is defined as the value xn such that $F(xn;\boldsymbol{\theta},\mathbf{min},\mathbf{max})=n/100$. Thus, if we know that the expert distribution's nth percentile point is xn and mth percentile point is xm then the following has to hold:

$$F(xn;\boldsymbol{\theta},\mathbf{min},\mathbf{max})=n/100$$
$$F(xm;\boldsymbol{\theta},\mathbf{min},\mathbf{max})=m/100$$

This is a system of non-linear equations that can be solved for the unknown distribution parameters $\boldsymbol{\theta}$. We used the Nelder and Mead (1965) numeric optimization algorithm, available in R, to find the best-fitting estimates of parameters $\boldsymbol{\theta}$ for the truncated distributions specified by the experts. Once estimates of $\boldsymbol{\theta}$ were obtained, the distributions were specified fully and we had enough information to make 15,000 draws from each.

Table F-2 below summarizes the results for each expert who specified a parametric distribution. In each case, we provide an "input" line that has all the information that was provided by the expert. We also show the "output" line that contains the inferred parameters and five percentile points of the distribution from which draws were made.

Highlighted in yellow are the percentiles specified by the expert and used to create the equation system for the optimization. After finding the best-fitting parameters, we calculated the associated percentiles and confirmed that they are close to the input values.

Table F-2. Description of the Parametric Expert Functions

Expert	Information	Distribution	Min	P5	P25	P50	P75	P95	Max	Parameters
A	input	Normal	0	0.290				2.900	4	mean=? sd=?
	output			0.290	0.929	1.481	2.059	2.900		mean=1.42 sd=0.895
C	input	Normal	0			1.200		2.000	+∞	mean=? sd=?
	output			0.423	0.875	1.200	1.528	2.000		mean=1.196 sd=0.488
D	input	Triangular	0.100						1.600	mode=0.95
	output			0.350	0.662	0.897	1.107	1.382		
E	input	Normal	0			2.000		3.000	+∞	mean=? sd=?
	output			1.002	1.590	2.000	2.410	3.000		mean=2 sd=0.608
G	input	Normal	-∞			1.000		1.300	1.500	mean=? sd=?
	output			0.695	0.875	1.000	1.124	1.300		mean=1.001 sd=0.185
I	input	Normal	0.200						2.300	mean=1.25 sd=0.53
	output			0.473	0.912	1.250	1.588	2.027		
J	input	Weibull	0	0.150		0.900		2.000	3.000	shape=? scale=? location=?
	output			0.150	0.525	0.900	1.331	2.000		shape=2.21 scale=1.413 location=-0.326
K1 4-16 ug/m3	input	Normal	-∞	0.100		0.400			0.800	mean=? sd=?
	output			0.100	0.277	0.400	0.521	0.682		mean=0.404 sd=0.184
K2 >16-30 ug/m3	input	Normal	-∞	0.100		0.700			1.500	mean=? sd=?
	output			0.100	0.455	0.700	0.942	1.264		mean=0.707 sd=0.367

For example, Expert A indicated that the distribution of the effect is Normal, with minimum 0 and maximum 4. Under the assumption that this is actually a truncated Normal distribution, we looked for the corresponding mean and standard deviation for it. The 5th and the 95th percentile values (0.29 and 2.90, respectively) were used to specify the following equations:

$$N(0.29; \textbf{mean=?,sd=?}, \text{min=}\textbf{0}, \text{max=}\textbf{4}) = 0.05$$

$N(2.90;$ **mean=?,sd=?**,min=**0**,max=**4**$)=0.95$

The solution to this system was a mean of 1.42 and a standard deviation of 0.89. We also verified that these parameters produced percentile values consistent with the ones supplied by the expert. We similarly solved for the parameters of the other experts who specified parametric distributions, with the exception of experts D and I, who specified their distributions fully.

The experts were asked to describe uncertainty distributions for the percent change in mortality risk associated with a 1 μg/m3 change in PM2.5. All of the experts assumed log-linear (or piecewise log-linear) C-R functions. If Z denotes the percent change elicited from an expert, the relative risk associated with a 1 μg/m3 change in PM2.5 is (1+Z/100), and the PM2.5 coefficient in the log-linear C-R function is ln(1+(Z/100)). We applied this transformation to the values drawn from each distribution.

Finally, some experts stated that their distribution does not incorporate the likelihood of causality – *i.e.*, they specified conditional distributions. We made 15,000 draws from an expert's conditional distribution. BenMAP contains a function that is zero. If an expert specified, for example, a five percent chance that there is not a causal association, BenMAP will draw from this zero function with five percent probability and draw from the 15,000-draw custom distribution (of positive values) with 95 percent probability. Table F-3 below shows summary statistics for the draws from the parametric distributions that became BenMAP "custom" distribution tables. Additional details on the form of the distributions are below and in Belova et al (2007).

Table F-3. Descriptive Statistics of the Random Draws from the Parametric Expert Distributions

Expert	Mean	Standard Deviation	Min	P25	P50	P75	Max
A	0.01518	0.00773	0.00000	0.00944	0.01483	0.02051	0.03917
C	0.01193	0.00466	0.00001	0.00870	0.01189	0.01509	0.02848
D (cond)	0.00884	0.00305	0.00105	0.00671	0.00899	0.01108	0.01577
D	0.00838	0.00354	0.00000	0.00623	0.00875	0.01092	0.01577
E (cond)	0.01975	0.00591	0.00026	0.01577	0.01986	0.02376	0.04534
E	0.01967	0.00619	0.00000	0.01575	0.01989	0.02381	0.04534
G (cond)	0.00996	0.00181	0.00256	0.00873	0.00996	0.01123	0.01489
G	0.00697	0.00480	0.00000	0.00000	0.00892	0.01062	0.01489
I (cond)	0.01240	0.00458	0.00200	0.00905	0.01244	0.01575	0.02273
I	0.01181	0.00523	0.00000	0.00845	0.01214	0.01559	0.02273
J	0.00962	0.00567	0.00000	0.00525	0.00902	0.01329	0.02936
K1 (cond)	0.00394	0.00175	-0.00262	0.00278	0.00398	0.00520	0.00797
K1	0.00139	0.00215	-0.00262	0.00000	0.00000	0.00298	0.00796
K2 (cond)	0.00689	0.00350	-0.00766	0.00452	0.00698	0.00937	0.01489
K2	0.00237	0.00382	-0.00402	0.00000	0.00000	0.00489	0.01488

F.1.1.2 Non-Parametric Distributions

Experts B, F, H, and L chose a non-parametric distribution function to represent their subjective beliefs about the percent change in risk associated with 1 μg/m3 increase in PM 2.5. They specified the following characteristics of the distribution:

The truncation points (*i.e.*, minimum and/or maximum)
Five percentile points
The likelihood that the association is causal and whether the function includes that (*i.e.*, whether the function is conditional on the association being causal or unconditional)

The only information that we had about these distributions was the minimum, the maximum, and the five percentiles. The shape of the distribution was unknown. Therefore, we made an assumption that the cumulative distribution function (cdf) is piece-wise linear. In other words, we assumed that all values between the percentiles are equally likely. Following this assumption, we used linear interpolation between the percentile points to derive the cdf for each expert. We then made 15,000 draws from each cdf.

Table F-4 shows the inputs and the outputs of this process for each expert. The inputs are the minimum, the maximum, and the percentiles. The outputs are the percentiles that we calculated from the draws from the respective linearly interpolated cdfs.

Table F-4. Description of the Non-Parametric Expert Functions

Expert	Information	Min	P5	P10	P25	P50	P75	P95	Max
B1 4-10 ug/m3	input	0.010	0.100		0.200	1.200	2.100	2.600	2.800
	output		0.099		0.203	1.213	2.092	2.599	
B2 >10-30 ug/m3	input	0.100	0.200		0.500	1.200	2.100	2.600	2.800
	output		0.198		0.501	1.191	2.096	2.597	
F1 4-7 ug/m3	input	0.370	0.580		0.730	0.930	1.100	1.400	1.700
	output		0.581		0.732	0.928	1.097	1.407	
F2 >7-30 ug/m3	input	0.290	0.770		0.960	1.100	1.400	1.600	1.800
	output		0.771		0.958	1.100	1.398	1.606	
H	input	0	0	0	0.400	0.700	1.300	2.000	3.000
	output		0	0	0.407	0.710	1.320	2.010	
L1 4-10 ug/m3	input	0	0.200		0.570	1.000	1.400	1.600	2.700
	output		0.201		0.570	0.996	1.400	1.619	
L2 >10-30 ug/m3	input	0.020	0.200		0.570	1.000	1.400	1.600	2.700
			0.018		0.568	1.003	1.396	1.634	

Table F-5 below shows summary statistics for the draws from the non-parametric distributions that became BenMAP "custom" distribution tables. The section below on

distributional details contains histograms for all the experts' distributions.

Table F-5. Descriptive Statistics of the Random Draws from the Non-Parametric Expert Distributions

Expert	Mean	Standard Deviation	Min	P25	P50	P75	Max
B1 (cond)	0.01217	0.00897	0.00010	0.00200	0.01195	0.02090	0.02761
B1	0.01195	0.00901	0.00000	0.00195	0.01167	0.02075	0.02761
B2 (cond)	0.01290	0.00813	0.00100	0.00489	0.01187	0.02068	0.02761
B2	0.01262	0.00827	0.00000	0.00464	0.01159	0.02042	0.02761
F1	0.00937	0.00268	0.00370	0.00727	0.00924	0.01092	0.01686
F2	0.01144	0.00292	0.00290	0.00951	0.01091	0.01387	0.01784
H	0.00870	0.00662	0.00000	0.00406	0.00702	0.01302	0.02954
L1 (cond)	0.00985	0.00511	0.00001	0.00582	0.00999	0.01391	0.02662
L1	0.00739	0.00613	0.00000	0.00001	0.00727	0.01250	0.02659
L2 (cond)c	0.00953	0.00544	0.00000	0.00567	0.00991	0.01389	0.02661
L2	0.00934	0.00549	0.00000	0.00531	0.00964	0.01371	0.02661

F.1.1.3 Using Expert Functions in BenMAP

When an expert has specified certain functional specifics with certain probabilities, the resulting "C-R function" becomes a *set* of possible functions, each with an associated probability. For example, expert K specified a piecewise log-linear function (i.e., two different log-linear functions on two different parts of the range of PM2.5); this expert also specified a threshold within different ranges with different probabilities (and no threshold with a specified probability). BenMAP incorporates such a set of possible functions specified by an expert function by assigning appropriate weights to each specification. We illustrate this using expert K's specification.

Expert K specified one log-linear function if the baseline PM2.5 value falls within the range from 4 μg/m3 to 16 μg/m3 and another log-linear function if the baseline value falls within the range from >16 μg/m3 to 30 μg/m3. BenMAP thus incorporates two sets of functions – one set for each of these two PM2.5 ranges – and selects from the set appropriate for a given PM2.5 baseline value. Expert K also specified a 64% probability that there is no causal relationship; an 18% probability that there is a causal relationship with no threshold, a 4% probability that there is a causal relationship with a threshold somewhere between 5 μg/m3 to 10 μg/m3, and a 14% probability that there is a causal relationship with a threshold somewhere between 0 μg/m3 to 5 μg/m3. Thus, the set of log-linear functions in BenMAP for expert K on the range from 4 μg/m3 to 16 μg/m3 contains:

a function with PM2.5 coefficient = 0 (no causality), which BenMAP selects with 65% probability;
a function with the PM2.5 coefficient expert K specified for the log-linear function on that range and no threshold, which BenMAP selects with 18% probability;
a function with the PM2.5 coefficient expert K specified for the log-linear function on that range and a threshold (with uniform probability) between 0 μg/m3 to 5 μg/m3, which

BenMAP selects with 14% probability; and

a function with the PM2.5 coefficient expert K specified for the log-linear function on that range and a threshold (with uniform probability) between 5 μg/m3 to 10 μg/m3, which BenMAP selects with 4% probability.

If the PM2.5 baseline value is greater than 16 μg/m3, BenMAP goes through an analogous procedure to select a function from among the two functions in that set.

F.1.1.4 Distributional Details by Expert

Distributional details on each expert distribution are presented below. The derivation of the distributions is described above with additional details provided by Belova et al (2007).

F.1.1.4.1   Expert A

Figure F-1. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert A
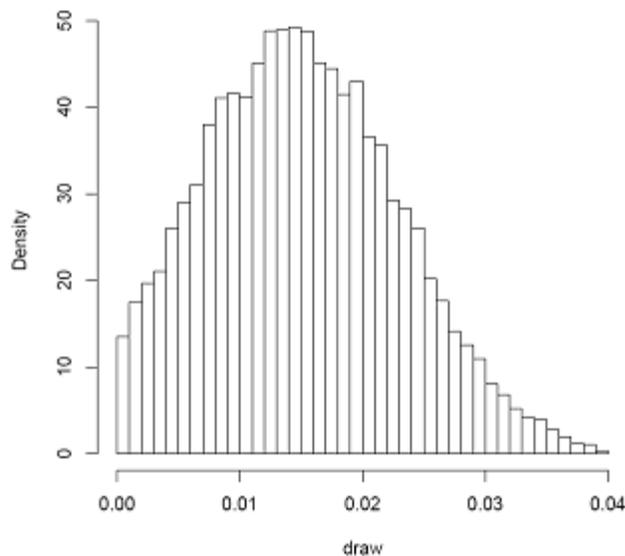
Notes:

- Expert A specified a truncated Normal Distribution. We inferred the following values for the parameters of this distribution: mean=1.42 and standard deviation=0.89.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.2   Expert B

Figure F-2. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert B

(1) Results for the range 4-10 ug/m^3

Notes:

- Expert B specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 2 percent zeros to the draw. Panels (c) and (d) show the respective distributions.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding $PM_{2.5}$ effects was the following transformation of the percent change in relative risk $Z – \log(1+(Z/100))$.

Figure F-2. Characteristics of the Random Draw from the Approximated Distribution of the $PM_{2.5}$ Effect Specified by Expert B (continued)

(2)   Results for the range >10-30 ug/m³

Notes:

- Expert B specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 2 percent zeros to the draw. Panels (c) and (d) show the

respective distributions.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding $PM_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.3   Expert C

Figure F-3. Histogram of the Random Draw from the Distribution of the $PM_{2.5}$ Effect Specified by Expert C
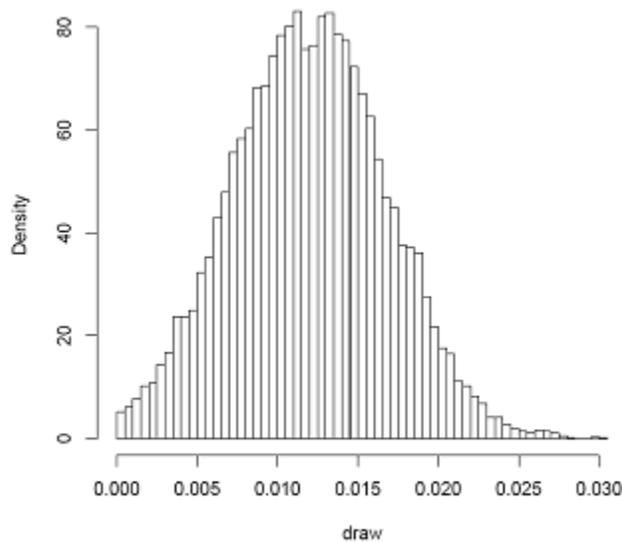
Notes:

- Expert C specified a truncated Normal Distribution. We inferred the following values for the parameters of this distribution: mean=1.20 and standard deviation=0.49.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding $PM_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.4   Expert D

Figure F-4. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert D
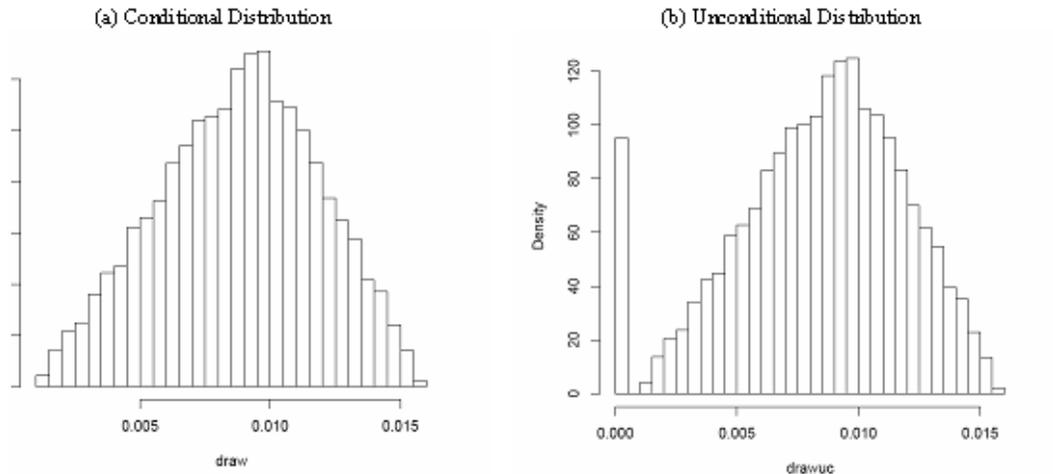
Notes:

- Expert D specified a Triangular Distribution with minimum=0.1, maximum=1.6, and mode=0.95. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 5 percent zeros to the draw.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.5   Expert E

Figure F-5. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert E
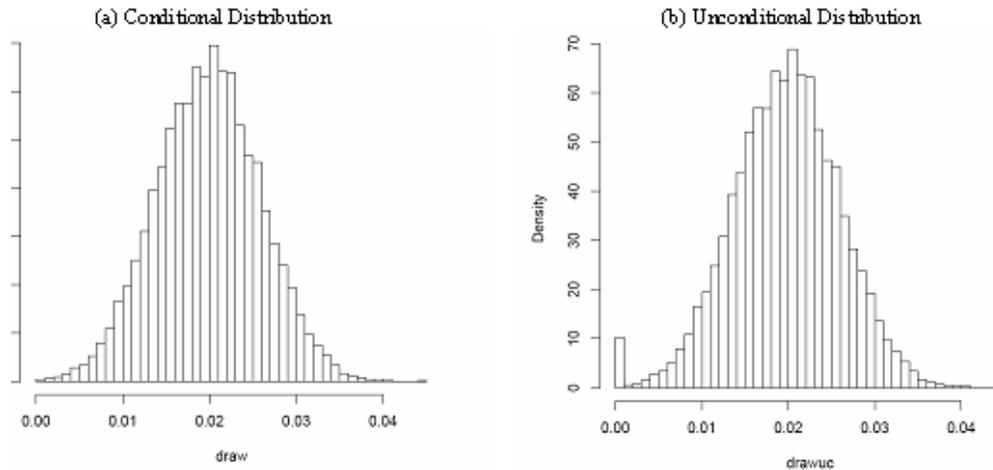
Notes:

- Expert E specified a truncated Normal Distribution. We inferred the following parameters for this distribution: mean=2.00 and standard deviation=0.61. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 1 percent zeros to the draw.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.6   Expert F

Figure F-6. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert F

(1)   Results for the range 4-7 ug/m3
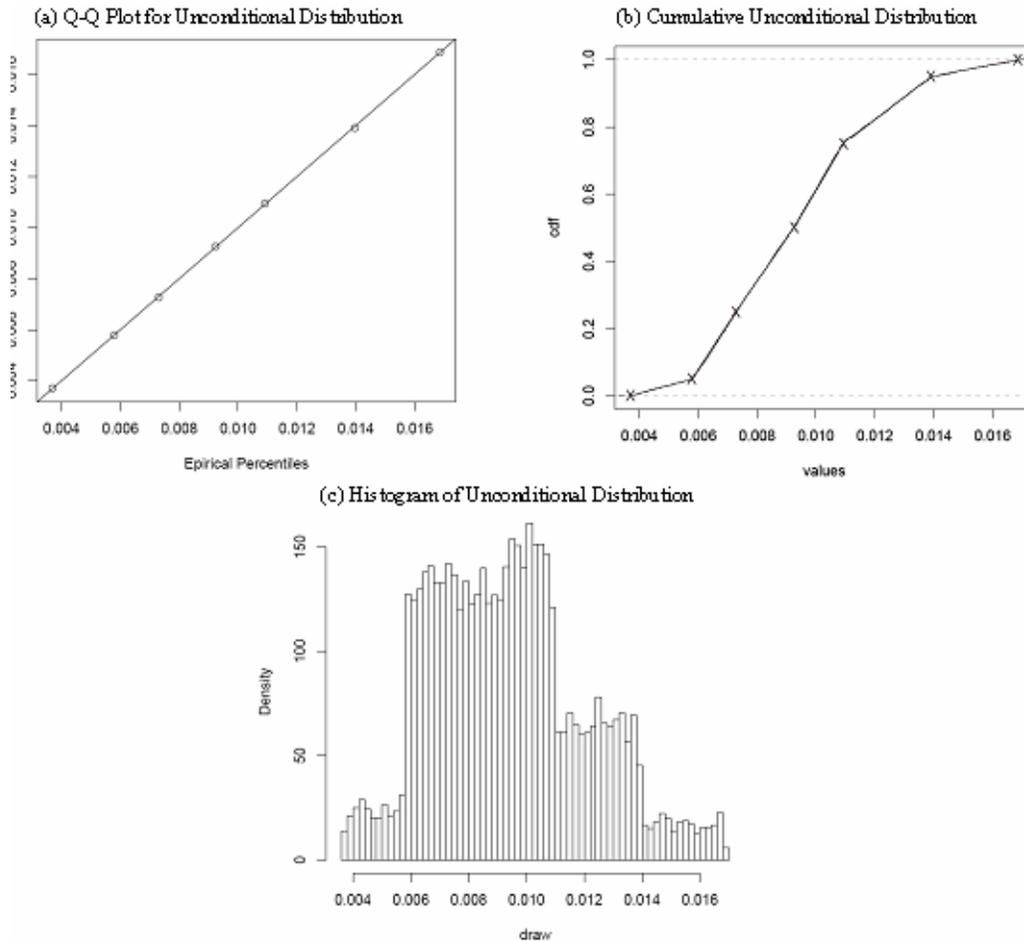
(a) Q-Q Plot for Unconditional Distribution

(b) Cumulative Unconditional Distribution

(c) Histogram of Unconditional Distribution

Notes:

- Expert F specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. Panel (c) shows the histogram of the distribution.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

Figure F-6. Characteristics of the Random Draw from the Approximated Distribution of the PM2.5 Effect Specified by Expert F (continued)

(2) Results for the range >7-30 ug/m3
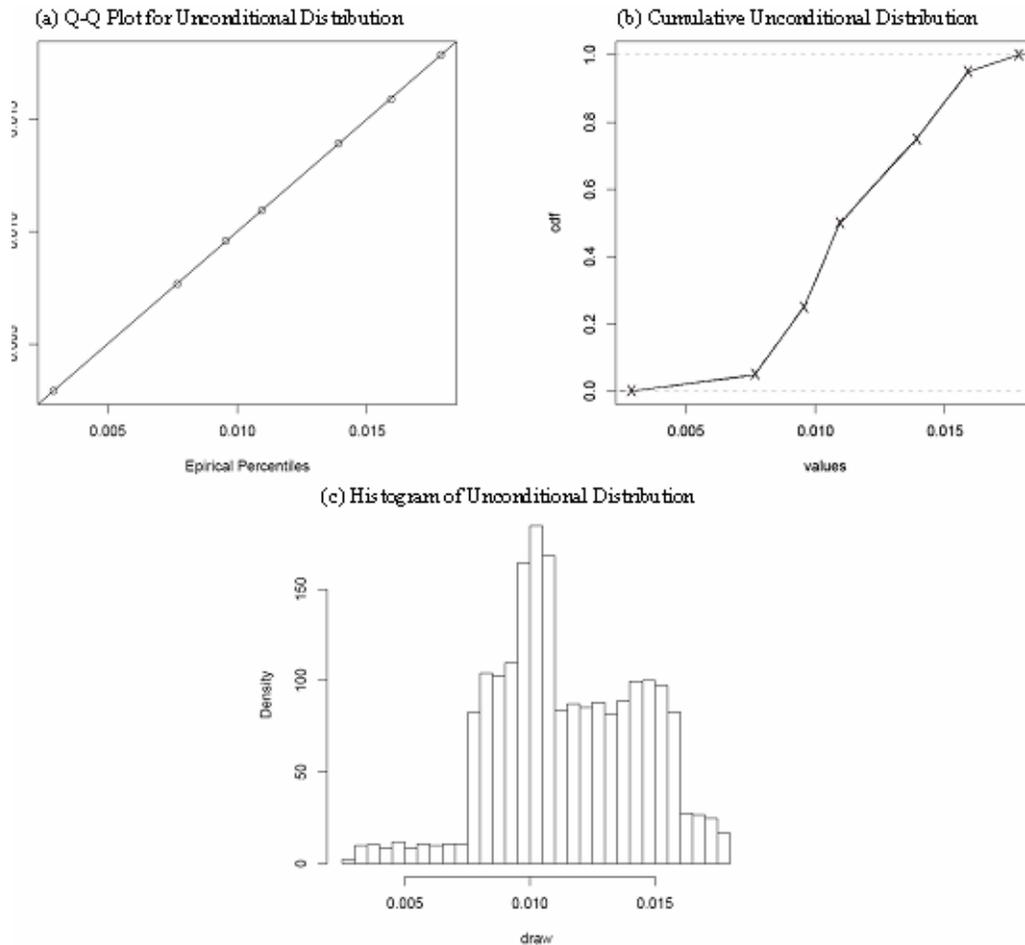
Notes:

- Expert F specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. Panel (c) shows the histogram of the distribution.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding $PM_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.7   Expert G

Figure F-7. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert G
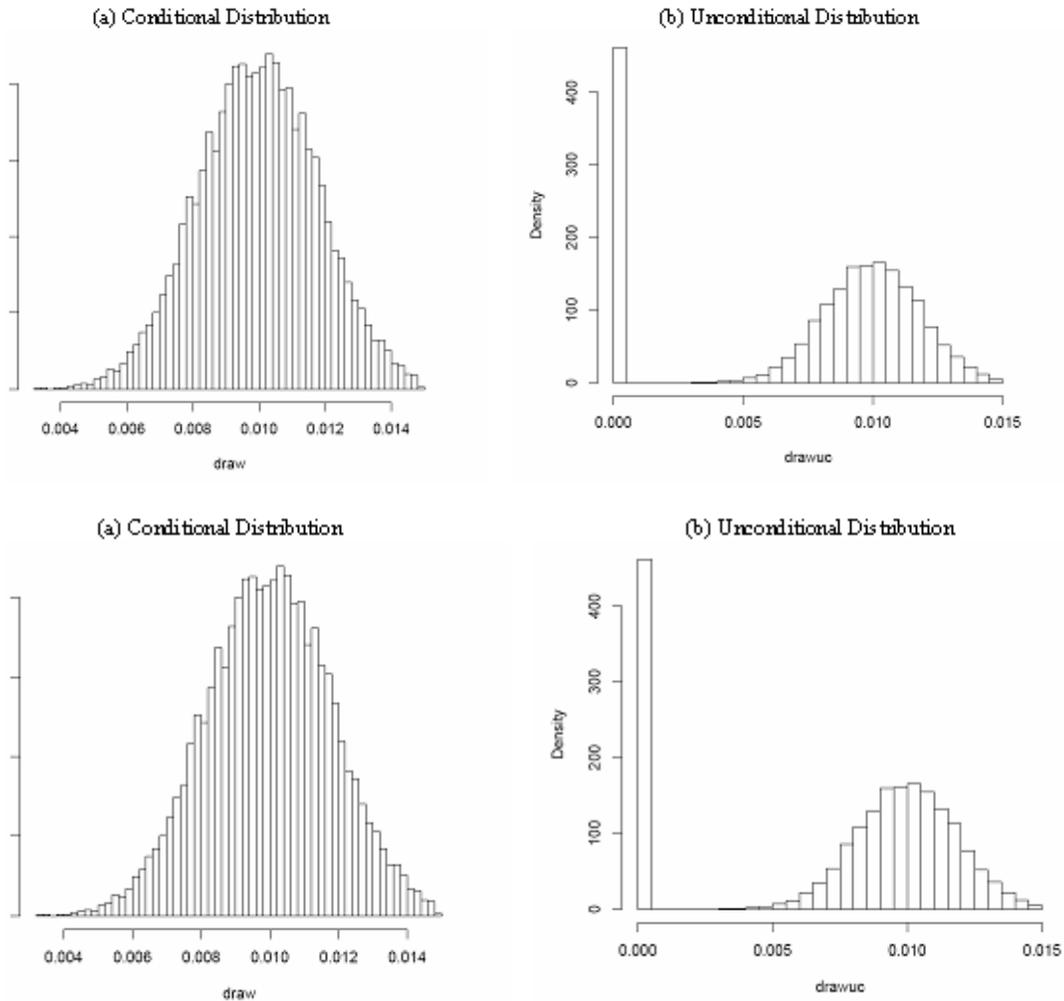
Notes:

- Expert G specified a truncated Normal Distribution. We inferred the following parameters for this distribution: mean=1.00 and standard deviation=0.19. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 30 percent zeros to the draw.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.8   Expert H

Figure F-8. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert H
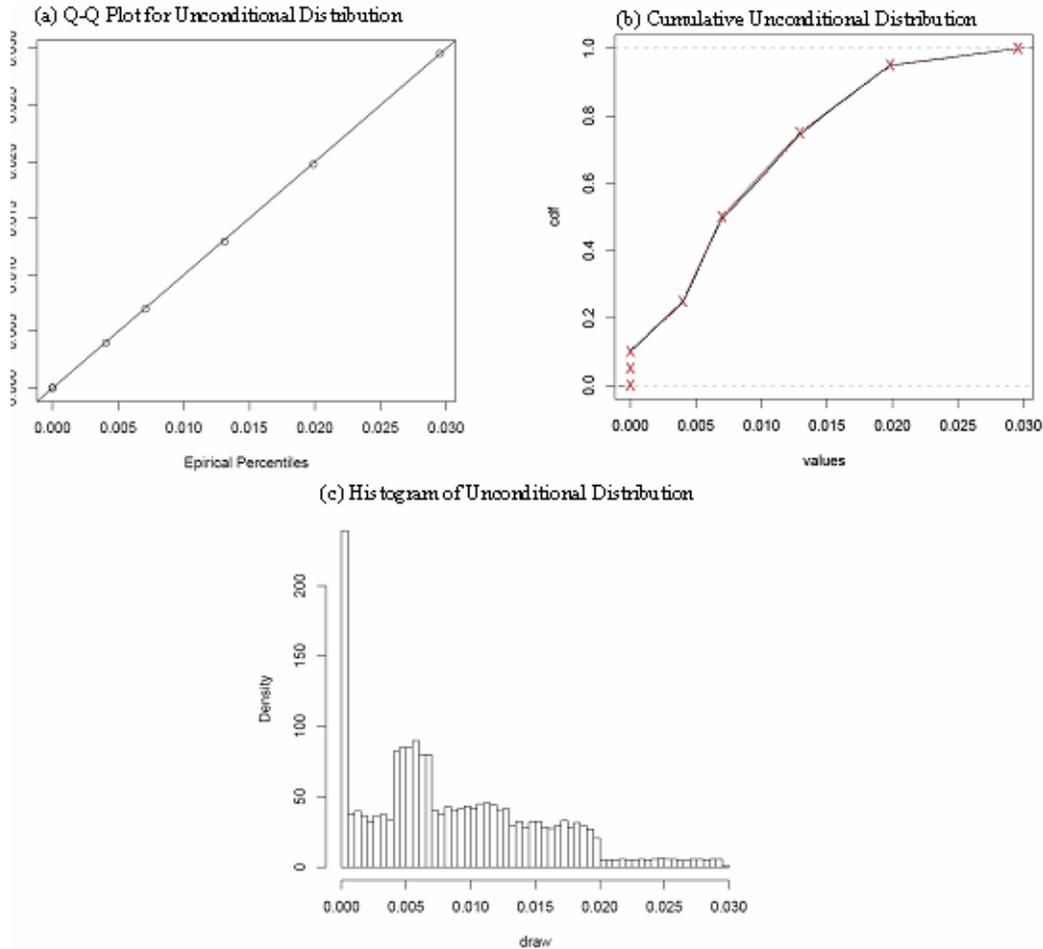
Notes:

- Expert H specified a non-parametric distribution using six percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. Panel (c) shows the histogram of the distribution.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.9   Expert I

Figure F-9. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert I
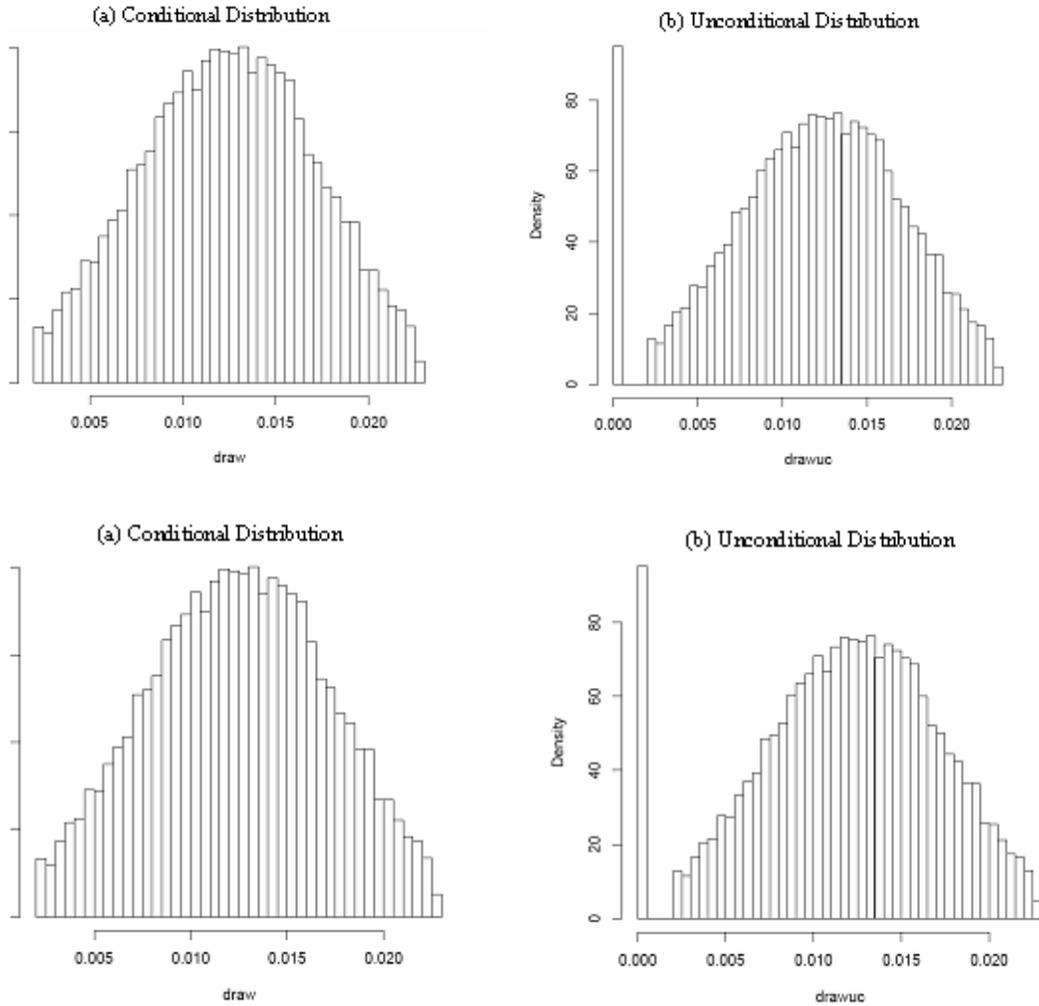
Notes:

- Expert I specified a truncated Normal Distribution with mean=1.25 and standard deviation=0.53. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 5 percent zeros to the draw.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.10   Expert J

Figure F-10. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert J
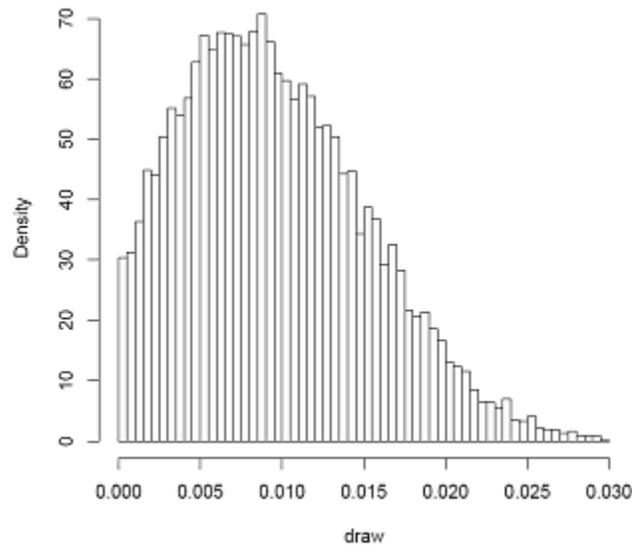
Notes:

- Expert J specified a truncated Weibull Distribution. We inferred the following values for the parameters of this distribution: shape=2.21, scale=1.41, and location=-0.33.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.11   Expert K

Figure F-11. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert K
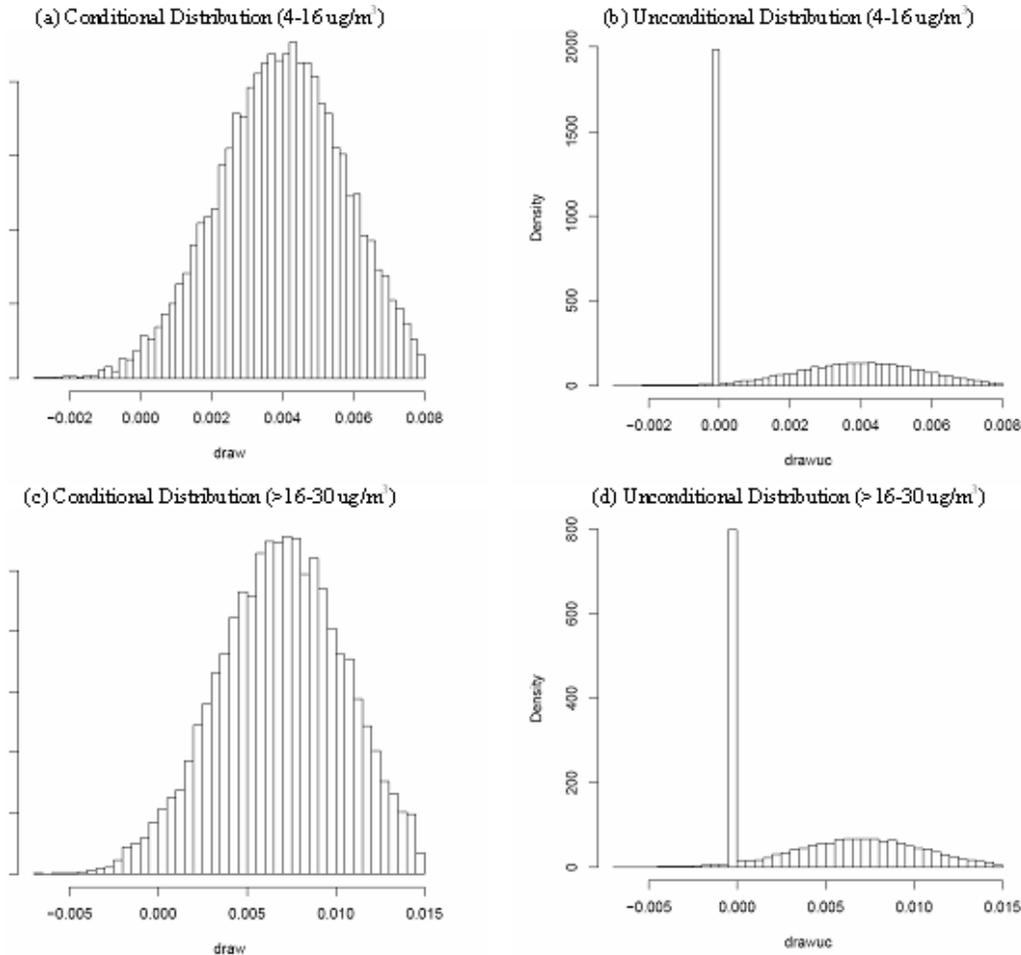
Notes:

- Expert K specified a truncated Normal Distribution two ranges (4-16 ug/m3 and >16-30 ug/m3). We inferred the following parameters for this distribution: mean=0.40 and standard deviation=0.18 in the lower range and mean=0.71 and standard deviation=0.37 in the upper range. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 65 percent zeros to the draws in each range.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk $Z - \log(1+(Z/100))$.

F.1.1.4.12   Expert L

Figure F-12. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert L

(1) Results for the range 4-10 ug/m^3
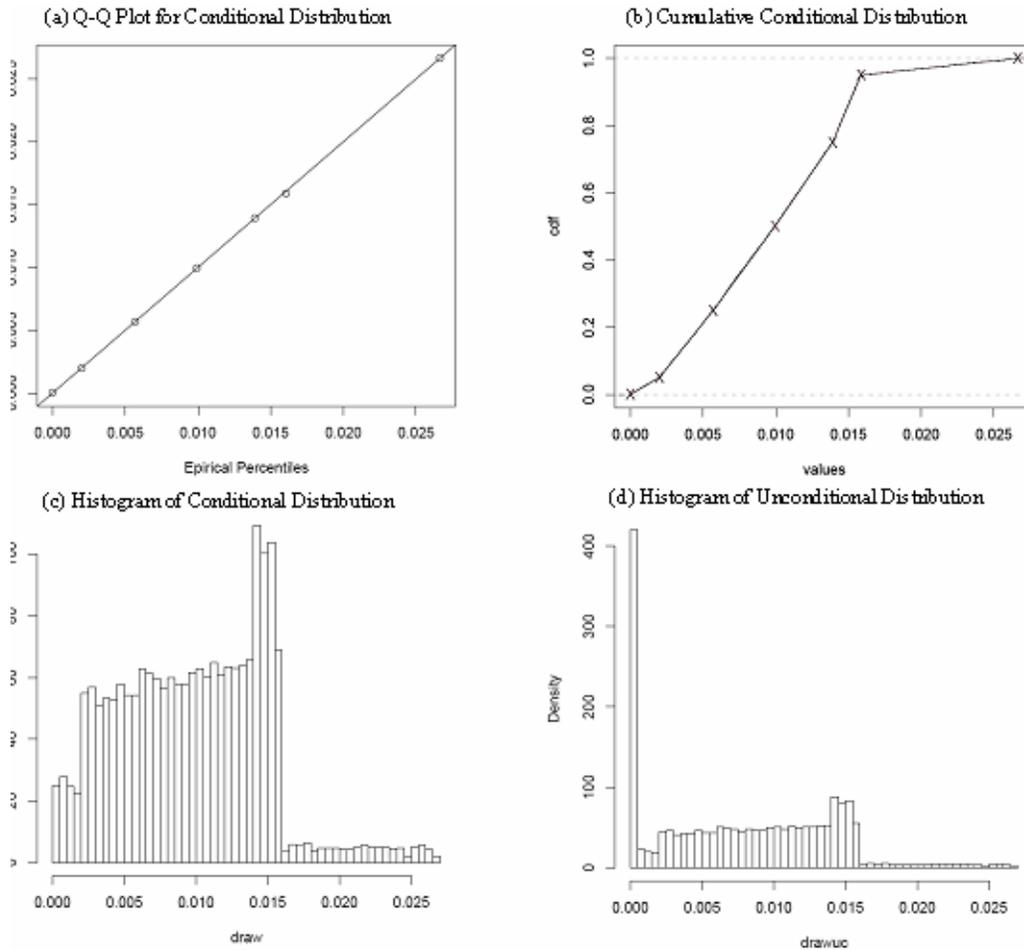
Notes:

- Expert L specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 25 percent zeros to the draw. Panels (c) and (d) show the respective distributions.

- The experts specified distributions for the percent changes in the relative risk. The distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the

percent change in relative risk $Z - \log(1+(Z/100))$.

Figure F-12. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert L (continued)

(2)   Results for the range >10-30 ug/m^3

Notes:

- Expert L specified a non-parametric distribution using five percentile points. We linearly interpolated the cdf between them. Panel (a) shows q-q plot of the expert percentiles and empirical percentiles for the draw. Panel (b) shows empirical cdf associated with the draw, the red "X" marks indicate corresponding expert percentiles. The distribution was conditional on causality. We created a corresponding unconditional distribution by adding extra 1 percent zeros to the draw. Panels (c) and (d) show the respective distributions.

- The experts specified distributions for the percent changes in the relative risk. The

distribution of the corresponding PM$_{2.5}$ effects was the following transformation of the percent change in relative risk Z – log(1+(Z/100)).

F.1.2   Laden et al (2006)

A large body of epidemiologic literature has found an association of increased fine particulate air pollution (PM2.5) with acute and chronic mortality. The effect of improvements in particle exposure is less clear. Earlier analysis of the Harvard Six Cities adult cohort study showed an association between long-term ambient PM2.5 and mortality between enrollment in the mid-1970's and follow-up until 1990. The authors extended mortality follow-up for eight years in a period of reduced air pollution concentrations. Annual city-specific PM2.5 concentrations were measured between 1979-1988, and estimated for later years from publicly available data. E xposure was defined as (1) city-specific mean PM2.5 during the two follow-up periods, (2) mean PM2.5 in the first period and change between these periods, (3) overall mean PM2.5 across the entire follow-up, and (4) year-specific mean PM2.5. Mortality rate ratios were estimated with Cox proportional hazards regression controlling for individual risk factors. The authors found an increase in overall mortality associated with each 10 ug/m3 increase in PM2.5 modeled either as the overall mean (RR=1.16, 95%CI=1.07-1.26) or as exposure in the year of death (RR=1.14, 95%CI=1.06-1.22). PM2.5 exposure was associated with lung cancer (RR=1.27, 95%CI=0.96-1.69) and cardiovascular deaths (RR=1.28, 95%CI=1.13-1.44). Improved overall mortality was associated with decreased mean PM2.5 (10 microg/m(3)) between periods (RR=0.73, 95% CI=0.57-0.95). Total, cardiovascular, and lung cancer mortality were each positively associated with ambient PM2.5 concentrations. Reduced PM2.5 concentrations were associated with reduced mortality risk.

All-Cause Mortality

The coefficient and standard error for PM$_{2.5}$ are estimated from the relative risk (1.16) and 95% confidence interval (1.07-1.26) associated with a change in *annual mean* exposure of 10.0 µg/m^3 (Laden et al, 2006, p. 667).

F.1.3   Pope et al (2002)

The Pope et al. (2002) analysis is a longitudinal cohort tracking study that uses the same American Cancer Society (ACS) cohort as the original Pope et al. (1995) study, and the Krewski et al. (2000) reanalysis. Pope et al. (2002) analyzed survival data for the cohort from 1982 through 1998, 9 years longer than the original Pope study. Pope et al. (2002) also obtained PM$_{2.5}$ data in 116 metropolitan areas collected in 1999, and the first three quarters of 2000. This is more metropolitan areas with PM$_{2.5}$ data than was available in the Krewski reanalysis (61 areas), or the original Pope study (50 areas), providing a larger size cohort.

They used a Cox proportional hazard model to estimate the impact of long-term PM

exposure using three alternative measures of $PM_{2.5}$ exposure; metropolitan area-wide annual mean PM levels from the beginning of tracking period ('79-'83 PM data, conducted for 61 metropolitan areas with 359,000 individuals), annual mean PM from the end of the tracking period ('99-'00, for 116 areas with 500,000 individuals), and the average annual mean PM levels of the two periods (for 51 metropolitan areas, with 319,000 individuals). PM levels were lower in '99-00 than in '79 - '83 in most cities, with the largest improvements occurring in cities with the highest original levels.

Pope et al. (2002) followed Krewski et al. (2000) and Pope et al. (1995, Table 2) and reported results for all-cause deaths, lung cancer (ICD-9 code: 162), cardiopulmonary deaths (ICD-9 codes: 401-440 and 460-519), and "all other" deaths. All-cause mortality includes accidents, suicides, homicides and legal interventions. The category "all other" deaths is all-cause mortality less lung cancer and cardiopulmonary deaths. Like the earlier studies, Pope et al. (2002) found that mean $PM_{2.5}$ is significantly related to all-cause and cardiopulmonary mortality. In addition, Pope et al. (2002) found a significant relationship with lung cancer mortality, which was not found in the earlier studies. None of the three studies found a significant relationship with "all other" deaths.

Pope et al. (2002) obtained ambient data on gaseous pollutants routinely monitored by EPA during the 1982-1998 observation period, including SO_2, NO_2, CO, and ozone. They did not find significant relationships between NO_2, CO, and ozone and premature mortality, but there were significant relationships between SO_4 (as well as SO_2), and all-cause, cardiopulmonary, lung cancer and "all other" mortality.

All-Cause Mortality, '79-'83 Exposure

The coefficient and standard error for $PM_{2.5}$ using the '79-'83 PM data are estimated from the relative risk (1.04) and 95% confidence interval (1.01-1.08) associated with a change in *annual mean* exposure of 10.0 µg/m³ (Pope et al, 2002, Table 2).

All-Cause Mortality, Average of '79-'83 and '99-'00 Exposure

The coefficient and standard error for $PM_{2.5}$ using the average of '79-'83 and '99-'00 PM data are estimated from the relative risk (1.06) and 95% confidence interval (1.02-1.11) associated with a change in *annual mean* exposure of 10.0 µg/m³ (Pope et al, 2002, Table 2).

F.1.4   Woodruff et al (1997)

In a study of four million infants in 86 U.S. metropolitan areas conducted from 1989 to 1991, Woodruff et al. (1997) found a significant link between PM_{10} exposure in the first two months of an infant's life with the probability of dying between the ages of 28 days and 364 days. PM_{10} exposure was significant for all-cause mortality. PM_{10} was also

significant for respiratory mortality in average birth-weight infants, but not low birth-weight infants.

Post-Neonatal Mortality

The coefficient and standard error are based on the odds ratio (1.04) and 95% confidence interval (1.02-1.07) associated with a 10 µg/m³ change in PM_{10} (Woodruff et al., 1997, Table 3).

F.1.5   Woodruff et al (2006)

Studies suggest that airborne particulate matter (PM) may be associated with postneonatal infant mortality, particularly with respiratory causes and sudden infant death syndrome (SIDS). To further explore this issue, the authors examined the relationship between long-term exposure to fine PM air pollution and postneonatal infant mortality in California. They linked monitoring data for PM2.5 to infants born in California in 1999 and 2000 using maternal addresses for mothers who lived within 5 miles of a PM2.5 monitor. They matched each postneonatal infant death to four infants surviving to 1 year of age, by birth weight category and date of birth (within 2 weeks). For each matched set, they calculated exposure as the average PM2.5 concentration over the period of life for the infant who died. They used conditional logistic regression to estimate the odds of postneonatal all-cause, respiratory-related, SIDS, and external-cause (a control category) mortality by exposure to PM2.5, controlling for the matched sets and maternal demographic factors. They matched 788 postneonatal infant deaths to 3,089 infant survivors, with 51 and 120 postneonatal deaths due to respiratory causes and SIDS, respectively. They found an adjusted odds ratio for a 10-microg/m3 increase in PM2.5 of 1.07 [95% confidence interval (CI), 0.93-1.24] for overall postneonatal mortality, 2.13 (95% CI, 1.12-4.05) for respiratory-related postneonatal mortality, 0.82 (95% CI, 0.55-1.23) for SIDS, and 0.83 (95% CI, 0.50-1.39) for external causes.

Post-Neonatal Mortality

The coefficient and standard error for $PM_{2.5}$ are estimated from the relative risk (1.07) and 95% confidence interval (0.93-1.24) associated with a change in *annual mean* exposure of 10.0 µg/m³ (Woodruff et al., 2006, p. 786).

F.2   Chronic / Severe Illness

Table F-6 summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and chronic / severe health effects. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table F-6. Health Impact Functions for Particulate Matter and Chronic Illness

Effect	Author	Year	Lcoation	Age	Co-P oll	Metric	Beta	Std Err	Form	Notes
Chronic Bronchitis	Abbey et al.	1995	SF, SD, South Coast Air Basin	27-99		Annual	0.013185	0.006796	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Chronic Bronchitis	Abbey et al.	1995	SF, SD, South Coast Air Basin	27-99		Annual	0.013185	0.006796	Logistic	
Acute Myocardial Infarction, Nonfatal	Peters et al.	2001	Boston, MA	18-99		D24HourMe an	0.033230	0.012791	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Acute Myocardial Infarction, Nonfatal	Peters et al.	2001	Boston, MA	18-99		D24HourMe an	0.024121	0.009285	Logistic	

F.2.1   Abbey et al (1995b)

Abbey et al. (1995b) examined the relationship between estimated $PM_{2.5}$ (annual mean from 1966 to 1977), PM_{10} (annual mean from 1973 to 1977) and TSP (annual mean from 1973 to 1977) and the same chronic respiratory symptoms in a sample population of 1,868 Californian Seventh Day Adventists. The initial survey was conducted in 1977 and the final survey in 1987. To ensure a better estimate of exposure, the study participants had to have been living in the same area for an extended period of time. In single-pollutant models, there was a statistically significant $PM_{2.5}$ relationship with development of chronic bronchitis, but not for AOD or asthma; PM_{10} was significantly associated with chronic bronchitis and AOD; and TSP was significantly associated with all cases of all three chronic symptoms. Other pollutants were not examined.

Chronic Bronchitis

The estimated coefficient (0.0137) is presented for a one µg/m³ change in $PM_{2.5}$ (Abbey et al., 1995b, Table 2). The standard error is calculated from the reported relative risk (1.81) and 95% confidence interval (0.98-3.25) for a 45 µg/m³ change in $PM_{2.5}$.

Incidence Rate: annual bronchitis incidence rate per person (Abbey et al., 1993, Table 3) = 0.00378

Population: population of ages 27 and older without chronic bronchitis = 95.57% of population 27+. Using the same data set, Abbey et al. (1995a, p. 140) reported that the respondents in 1977 ranged in age from 27 to 95. The American Lung Association (2002b, Table 4) reports a chronic bronchitis prevalence rate for ages 18 and over of 4.43%.

F.2.2   Peters et al (2001)

Peters et al. (2001) studied the relationship between increased particulate air pollution and

onset of heart attacks in the Boston area from 1995 to 1996. The authors used air quality data for PM_{10}, $PM_{10-2.5}$, $PM_{2.5}$, "black carbon", O_3, CO, NO_2, and SO_2 in a case-crossover analysis. For each subject, the case period was matched to three control periods, each 24 hours apart. In univariate analyses, the authors observed a positive association between heart attack occurrence and $PM_{2.5}$ levels hours before and days before onset. The authors estimated multivariate conditional logistic models including two-hour and twenty-four hour pollutant concentrations for each pollutant. They found significant and independent associations between heart attack occurrence and both two-hour and twenty-four hour $PM_{2.5}$ concentrations before onset. Significant associations were observed for PM_{10} as well. None of the other particle measures or gaseous pollutants were significantly associated with acute myocardial infarction for the two hour or twenty-four hour period before onset.

The patient population for this study was selected from health centers across the United States. The mean age of participants was 62 years old, with 21% of the study population under the age of 50. In order to capture the full magnitude of heart attack occurrence potentially associated with air pollution and because age was not listed as an inclusion criteria for sample selection, we apply an age range of 18 and over in the C-R function. According to the National Hospital Discharge Survey, there were no hospitalizations for heart attacks among children <15 years of age in 1999 and only 5.5% of all hospitalizations occurred in 15-44 year olds (Popovic, 2001, Table 10).

Acute Myocardial Infarction, Nonfatal

The coefficient and standard error are calculated from an odds ratio of 1.62 (95% CI 1.13-2.34) for a 20 $\mu g/m^3$ increase in twenty-four hour average $PM_{2.5}$ (Peters et al., 2001, Table 4, p. 2813).

Incidence Rate: region-specific daily nonfatal heart attack rate per person 18+ = 93% of region-specific daily heart attack hospitalization rate (ICD code 410). This estimate assumes that all heart attacks that are not instantly fatal will result in a hospitalization. In addition, Rosamond et al. (1999) report that approximately six percent of male and eight percent of female hospitalized heart attack patients die within 28 days (either in or outside of the hospital). We applied a factor of 0.93 to the number of hospitalizations to estimate the number of nonfatal heart attacks per year.

Population: population of ages 18 and older

F.3   Hospitalizations

Table F-7 summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and hospital admissions. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table F-7. Health Impact Functions for Particulate Matter and Hospital Admissions

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Congestive Heart Failure	Ito	2003	Detroit, MI	65-99		D24HourMean	0.003458	0.0014553	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Congestive Heart Failure	Ito	2003	Detroit, MI	65-99		D24HourMean	0.003074	0.0012922	Log-linear	
HA, Dysrhythmia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0014055	0.0022887	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Dysrhythmia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0012499	0.0020233	Log-linear	
Ischemic Heart (less Myocardial Infarctions)	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0016140	0.0013000	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Ischemic Heart (less Myocardial Infarctions)	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0014355	0.0011556	Log-linear	
Chronic Lung	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0013116	0.0023222	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Chronic Lung	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0011669	0.0020644	Log-linear	
Pneumonia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0044766	0.0018667	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Pneumonia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.0039799	0.0016569	Log-linear	
All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.0015111	0.0003688	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.0014000	0.0003411	Log-linear	
Chronic Lung (less Asthma)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.0023744	0.0007911	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Chronic Lung (less Asthma)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.0022200	0.0007333	Log-linear	
All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.0017055	0.0003711	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.0015800	0.0003444	Log-linear	
Chronic Lung	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.0019996	0.0005665	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Chronic Lung	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.0018500	0.0005244	Log-linear	
Asthma	Sheppard	2003	Seattle, WA	0-64		D24HourMean	0.0039288	0.0012355	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
HA, Asthma	Sheppard	2003	Seattle, WA	0-64		D24HourMean	0.0033224	0.0010455	Log-linear	

F.3.1   Ito (2003)

Lippmann et al. (2000) studied the association between particulate matter and daily mortality and hospitalizations among the elderly in Detroit, MI. Data were analyzed for two separate study periods, 1985-1990 and 1992-1994. The 1992-1994 study period had a greater variety of data on PM size and was the main focus of the report. The authors collected hospitalization data for a variety of cardiovascular and respiratory endpoints. They used daily air quality data for PM_{10}, $PM_{2.5}$, and $PM_{10-2.5}$ in a Poisson regression model with generalized additive models (GAM) to adjust for nonlinear relationships and temporal trends. In single pollutant models, all PM metrics were statistically significant for pneumonia (ICD codes 480-486), $PM_{10-2.5}$ and PM_{10} were significant for ischemic heart disease (ICD code 410-414), and $PM_{2.5}$ and PM_{10} were significant for heart failure (ICD code 428). There were positive, but not statistically significant associations, between the PM metrics and COPD (ICD codes 490-496) and dysrhythmia (ICD code 427). In separate co-pollutant models with PM and either ozone, SO_2, NO_2, or CO, the results were generally comparable.

In response to concerns with the Splus issue, Ito (2003) reanalyzed the study by Lippmann et al. (2000). The reanalysis by Ito reported that more generalized additive models with stringent convergence criteria and generalized linear models resulted in smaller relative risk estimates.

Chronic Lung Disease (ICD-9 codes 490-496)

The coefficient and standard error are based on the relative risk (1.043) and 95% confidence interval (0.902-1.207) for a 36 $\mu g/m^3$ increase in $PM_{2.5}$ in the 3-day lag GAM stringent model (Ito, 2003, Table 8).

Pneumonia (ICD-9 codes 480-487)

The estimated $PM_{2.5}$ coefficient and standard error are based on a relative risk of 1.154 (95% CI -1.027, 1.298) due to a $PM_{2.5}$ change of 36 $\mu g/m^3$ in the 1-day lag GAM stringent model (Ito, 2003, Table 7).

Disrhythmia (ICD-9 code 429)

The co-pollutant coefficient and standard error are calculated from a relative risk of 1.046 (95% CI 0.906-1.207) for a 36 $\mu g/m^3$ increase in $PM_{2.5}$ in the 1-day lag GAM stringent model (Ito, 2003, Table 10).

Congestive Heart Failure (ICD-9 code 428)

The co-pollutant coefficient and standard error are calculated from a relative risk of 1.117 (95% CI 1.020-1.224) for a 36 $\mu g/m^3$ increase in $PM_{2.5}$ in the 1-day lag GAM stringent model (Ito, 2003, Table 11).

Ischemic Heart Disease (ICD-9 codes 411-414)

The co-pollutant coefficient and standard error are calculated from a relative risk of 1.053 (95% CI 0.971-1.143) for a 36 $\mu g/m^3$ increase in $PM_{2.5}$ in the 1-day lag GAM stringent model (Ito, 2003, Table 9)

Note that Lippmann et al. (2000) report results for ICD codes 410-414. In the benefit analysis, avoided nonfatal heart attacks are estimated using the results reported by Peters et al. (2001). The baseline rate in the Peters et al. function is a modified heart attack hospitalization rate (ICD code 410), since most, if not all, nonfatal heart attacks will require hospitalization. In order to avoid double counting heart attack hospitalizations, we have excluded ICD code 410 from the baseline incidence rate used in this function.

F.3.2   Moolgavkar (2000a), Chronic Lung

Moolgavkar (2000a) examined the association between air pollution and COPD hospital admissions (ICD 490-496) in the Chicago, Los Angeles, and Phoenix metropolitan areas. He collected daily air pollution data for ozone, SO_2, NO_2, CO, and PM_{10} in all three areas. $PM_{2.5}$ data was available only in Los Angeles. The data were analyzed using a Poisson regression model with generalized additive models to adjust for temporal trends. Separate models were run for 0 to 5 day lags in each location. Among the 65+ age group in Chicago and Phoenix, weak associations were observed between the gaseous pollutants and admissions. No consistent associations were observed for PM_{10}. In Los Angeles, marginally significant associations were observed for $PM_{2.5}$, which were generally lower than for the gases. In co-pollutant models with CO, the $PM_{2.5}$ effect was reduced. Similar results were observed in the 0-19 and 20-64 year old age groups.

The $PM_{2.5}$ C-R functions are based on the single and co-pollutant models ($PM_{2.5}$ and CO) reported for the 20-64 and 65+ age groups. Since the true PM effect is most likely best represented by a distributed lag model, then any single lag model should underestimate the total PM effect. As a result, we selected the lag models with the greatest effect estimates for use in the C-R functions.

Hospital Admissions, Chronic Lung Disease Less Asthma (ICD-9 codes 490-492, 494-496)

In a model with CO, the coefficient and standard error are calculated from an estimated

percent change of 2.0 and t-statistic of 2.2 for a 10 µg/m³ increase in $PM_{2.5}$ in the two-day lag model (Moolgavkar, 2000a, Table 4, p. 81). In a log-linear model, the percent change is equal to (RR - 1) * 100.

In this study, Moolgavkar defines and reports the "estimated" percent change as (log RR * 100). Because the relative risk is close to 1, RR-1 and log RR are essentially the same. For example, a true percent change of 2.0 would result in a relative risk of 1.020 and coefficient of 0.001980. The "estimated" percent change, as reported by Moolgavkar, of 2.0 results in a relative risk of 1.020201 and coefficient of 0.002.

Note that although Moolgavkar (2000a) reports results for the 20-64 year old age range, for comparability to other studies, we apply the results to the population of ages 18 to 64. Note also that in order to avoid double counting non-elderly asthma hospitalizations (ICD code 493), which are typically estimated separately in EPA benefit analyses, we have excluded ICD code 493 from the baseline incidence rate used in this function.

F.3.3   Moolgavkar (2000b), Cardiovascular

Moolgavkar (2000b) examined the association between air pollution and cardiovascular hospital admissions (ICD 390-448) in the Chicago, Los Angeles, and Phoenix metropolitan areas. He collected daily air pollution data for ozone, SO_2, NO_2, CO, and PM_{10} in all three areas. $PM_{2.5}$ data was available only in Los Angeles. The data were analyzed using a Poisson regression model with generalized additive models to adjust for temporal trends. Separate models were run for 0 to 5 day lags in each location. Among the 65+ age group, the gaseous pollutants generally exhibited stronger effects than PM_{10} or $PM_{2.5}$. The strongest overall effects were observed for SO_2 and CO. In a single pollutant model, $PM_{2.5}$ was statistically significant for lag 0 and lag 1. In co-pollutant models with CO, the $PM_{2.5}$ effect dropped out and CO remained significant. For ages 20-64, SO_2 and CO exhibited the strongest effect and any $PM_{2.5}$ effect dropped out in co-pollutant models with CO.

Hospital Admissions, All Cardiovascular (ICD codes 390-409, 411-459)

The single pollutant coefficient and standard error are calculated from an estimated percent change of 1.4 and t-statistic of 4.1 for a 10 µg/m³ increase in $PM_{2.5}$ in the zero lag model (Moolgavkar, 2000b, Table 4, p. 1203).

Note that (Moolgavkar (2000b) report results that include ICD code 410 (heart attack). In a benefit analysis, avoided nonfatal heart attacks are typically estimated separately. The baseline rate in the Peters et al. function is a modified heart attack hospitalization rate (ICD code 410), since most, if not all, nonfatal heart attacks will require hospitalization. In order to avoid double counting heart attack hospitalizations, we have excluded ICD code 410 from the baseline incidence rate used in this function.

F.3.4   Moolgavkar (2003)

Moolgavkar (2000a) examined the association between air pollution and COPD hospital admissions (ICD 490-496) in the Chicago, Los Angeles, and Phoenix metropolitan areas. In response to concerns with Splus issue, Moolgavkar (2003) reanalyzed his earlier studies. In the reanalysis, he reported that more generalized additive models with stringent convergence criteria and generalized linear models resulted in smaller relative risk estimates.

Hospital Admissions, Chronic Lung (ICD-9 codes 490-496)

The coefficient and standard error are calculated from an estimated percentage change of 1.85 and t-statistic of 3.53 for a 10 µg/m^3 increase in $PM_{2.5}$ in the 2-day lag GAM-30df stringent (10-^{8}) model (Moolgavkar, 2003, Table 17). In a log-linear model, the percent change is equal to (RR - 1) * 100.

The $PM_{2.5}$ C-R functions for the 65+ age group are based on the reanalysis in Moolgavkar (2003) of the single and co-pollutant models ($PM_{2.5}$ and CO). The true PM effect is most likely best represented by a distributed lag model, then any single lag model should underestimate the total PM effect. As a result, we selected the lag models with the greatest effect estimates for use in the C-R functions.

Hospital Admissions, All Cardiovascular (ICD-9 codes 390-429)

The single pollutant coefficient and standard error are calculated from an estimated percent change of 1.58 and t-statistic of 4.59 for a 10 µg/m^3 increase in $PM_{2.5}$ in the 0-day lag GAM-30df stringent (10-8) model (Moolgavkar, 2003, Table 12). In a log-linear model, the percent change is equal to (RR - 1) * 100.

F.3.5   Sheppard (2003)

Sheppard et al. (1999) studied the relation between air pollution in Seattle and nonelderly (<65) hospital admissions for asthma from 1987 to 1994. They used air quality data for PM_{10}, $PM_{2.5}$, coarse $PM_{10-2.5}$, SO_2, ozone, and CO in a Poisson regression model with control for time trends, seasonal variations, and temperature-related weather effects. PM2.5 levels were estimated from light scattering data. They found asthma hospital admissions associated with PM_{10}, $PM_{2.5}$, $PM_{10-2.5}$, CO, and ozone. They did not observe an association for SO_2. They found PM and CO to be jointly associated with asthma admissions. The best fitting co-pollutant models were found using ozone. However, ozone data was only available April through October, so they did not consider ozone further. For the remaining pollutants, the best fitting models included $PM_{2.5}$ and CO. Results for other co-pollutant models were not reported.

In response to concerns that the work by Sheppard et al. (1999) may be biased because of

the Splus issue, Sheppard (2003) reanalyzed some of this work, in particular Sheppard reanalyzed the original study's $PM_{2.5}$ single pollutant model.

Hospital Admissions, Asthma (ICD-9 code 493)

The coefficient and standard error are based on the relative risk (1.04) and 95% confidence interval (1.01-1.06) for a 11.8 µg/m³ increase in $PM_{2.5}$ in the 1-day lag GAM stringent model (Sheppard, 2003, pp. 228-229).

F.4   Emergency Room Visits

Table F-8 summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and emergency room visits. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table F-8. Health Impact Functions for Particulate Matter and Emergency Room Visits

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma	Norris et al.	1999	Seattle, WA	0-17	NO2, SO2	D24HourMean	0.018542	0.004644	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Asthma	Norris et al.	1999	Seattle, WA	0-17	NO2, SO2	D24HourMean	0.016527	0.004139	Log-linear	

F.4.1   Norris et al (1999)

Norris et al. (1999) examined the relation between air pollution in Seattle and childhood (<18) hospital admissions for asthma from 1995 to 1996. The authors used air quality data for PM_{10}, light scattering (used to estimate fine PM), CO, SO_2, NO_2, and O_3 in a Poisson regression model with adjustments for day of the week, time trends, temperature, and dew point. They found significant associations between asthma ER visits and light scattering (converted to $PM_{2.5}$), PM_{10}, and CO.   No association was found between O_3, NO_2, or SO_2 and asthma ER visits, although O_3 had a significant amount of missing data. In multipollutant models with either PM metric (light scattering or PM_{10}) and NO_2 and SO_2, the PM coefficients remained significant while the gaseous pollutants were not associated with increased asthma ER visits. The PM C-R functions are based on results of the single and multipollutant models reported.

Emergency Room Visits, Asthma

In a model with NO_2 and SO_2, the $PM_{2.5}$ coefficient and standard error are calculated from a relative risk of 1.17 (95% CI 1.08-1.26) for a 9.5 $\mu g/m^3$ increase in $PM_{2.5}$ (Norris et al., 1999, p. 491).

F.5   Minor Effects

Table F-9 summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and minor effects. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table F-9. Health Impact Functions for Particulate Matter and Minor Effects

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Acute Bronchitis	Dockery et al.	1996	24 communities	8-12		Annual	0.037894	0.023806	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Acute Bronchitis	Dockery et al.	1996	24 communities	8-12		Annual	0.027212	0.017096	Logistic	
Work Loss Days	Ostro	1987	Nationwide	18-64		D24HourMean	0.004600	0.000360	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Work Loss Days	Ostro	1987	Nationwide	18-64		D24HourMean	0.004600	0.000360	Log-linear	
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	Ozone	D24HourMean	0.007410	0.000700	Log-linear	Adjusted coefficient with 10 ug/m3 threshold.
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	Ozone	D24HourMean	0.007410	0.000700	Log-linear	
Lower Respiratory Symptoms	Schwartz and Neas	2000	6 U.S. cities	7-14		D24HourMean	0.019712	0.006226	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Lower Respiratory Symptoms	Schwartz and Neas	2000	6 U.S. cities	7-14		D24HourMean	0.019012	0.006005	Logistic	

F.5.1   Dockery et al (1996)

Dockery et al. (1996) examined the relationship between PM and other pollutants on the reported rates of asthma, persistent wheeze, chronic cough, and bronchitis, in a study of 13,369 children ages 8-12 living in 24 communities in U.S. and Canada. Health data were collected in 1988-1991, and single-pollutant models were used in the analysis to test a number of measures of particulate air pollution. Dockery et al. found that annual level of

sulfates and particle acidity were significantly related to bronchitis, and $PM_{2.1}$ and PM_{10} were marginally significantly related to bronchitis. The original study measured $PM_{2.1}$, however when using the study's results we use $PM_{2.5}$. This makes only a negligible difference, assuming that the adverse effects of $PM_{2.1}$ and $PM_{2.5}$ are comparable. They also found nitrates were linked to asthma, and sulfates linked to chronic phlegm. It is important to note that the study examined annual pollution exposures, and the authors did not rule out that acute (daily) exposures could be related to asthma attacks and other acute episodes. Earlier work, by Dockery et al. (1989), based on six U.S. cities, found acute bronchitis and chronic cough significantly related to PM_{15}. Because it is based on a larger sample, the Dockery et al. (1996) study is the better study to develop a C-R function linking $PM_{2.5}$ with bronchitis.

Bronchitis was counted in the study only if there were "reports of symptoms in the past 12 months" (Dockery et al., 1996, p. 501). It is unclear, however, if the cases of bronchitis are acute and temporary, or if the bronchitis is a chronic condition. Dockery et al. found no relationship between PM and chronic cough and chronic phlegm, which are important indicators of chronic bronchitis. For this analysis, we assumed that the C-R function based on Dockery et al. is measuring acute bronchitis. The C-R function is based on results of the single pollutant model reported in Table 1.

Acute Bronchitis

The estimated logistic coefficient and standard error are based on the odds ratio (1.50) and 95% confidence interval (0.91-2.47) associated with being in the most polluted city ($PM_{2.1}$ = 20.7 µg/m³) versus the least polluted city ($PM_{2.1}$ = 5.8 µg/m³) (Dockery et al., 1996, Tables 1 and 4). The original study used $PM_{2.1}$, however, we use the $PM_{2.1}$ coefficient and apply it to $PM_{2.5}$ data.

Incidence Rate: annual bronchitis incidence rate per person = 0.043 (American Lung Association, 2002a, Table 11)

Population: population of ages 8-12.

F.5.2   Ostro (1987)

Ostro (1987) estimated the impact of $PM_{2.5}$ on the incidence of work-loss days (WLDs), restricted activity days (RADs), and respiratory-related RADs (RRADs) in a national sample of the adult working population, ages 18 to 65, living in metropolitan areas. The study population is based on the Health Interview Survey (HIS), conducted by the National Center for Health Statistics. The annual national survey results used in this analysis were conducted in 1976-1981. Ostro reported that two-week average $PM_{2.5}$ levels were significantly linked to work-loss days, RADs, and RRADs, however there was some year-to-year variability in the results. Separate coefficients were developed for each year

in the analysis (1976-1981); these coefficients were pooled. The coefficient used in the concentration-response function presented here is a weighted average of the coefficients in Ostro (1987, Table III) using the inverse of the variance as the weight.

Work Loss Days

The coefficient used in the C-R function is a weighted average of the coefficients in Ostro (1987, Table III) using the inverse of the variance as the weight:

$$\beta = \left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = 0.0046.$$

The standard error of the coefficient is calculated as follows, assuming that the estimated year-specific coefficients are independent:

$$\sigma_\beta^2 = \mathrm{var}\left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = \mathrm{var}\left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\gamma} \right) = \sum\limits_{i=1976}^{1981} \mathrm{var}\left(\frac{\beta_i}{\sigma_{\beta_i}^2 \cdot \gamma} \right).$$

This eventually reduces down to:

$$\sigma_\beta^2 = \frac{1}{\gamma} \Rightarrow \sigma_\beta = \sqrt{\frac{1}{\gamma}} = 0.00036.$$

Incidence Rate: daily work-loss-day incidence rate per person ages 18 to 64 = 0.00595 (U.S. Bureau of the Census, 1997, No. 22; Adams et al., 1999, Table 41)

Population: adult population ages 18 to 64

F.5.3 Ostro and Rothschild (1989)

Ostro and Rothschild (1989) estimated the impact of $PM_{2.5}$ and ozone on the incidence of minor restricted activity days (MRADs) and respiratory-related restricted activity days (RRADs) in a national sample of the adult working population, ages 18 to 65, living in metropolitan areas. The study population is based on the Health Interview Survey (HIS),

conducted by the National Center for Health Statistics. In publications from this ongoing survey, non-elderly adult populations are generally reported as ages 18-64. From the study, it is not clear if the age range stops at 65 or includes 65 year olds. We apply the C-R function to individuals ages 18-64 for consistency with other studies estimating impacts to non-elderly adult populations. The annual national survey results used in this analysis were conducted in 1976-1981. Controlling for $PM_{2.5}$, two-week average ozone has highly variable association with RRADs and MRADs. Controlling for ozone, two-week average $PM_{2.5}$ was significantly linked to both health endpoints in most years.

Minor Restricted Activity Days

Using the results of the two-pollutant model, we developed separate coefficients for each year in the analysis, which were then combined for use in this analysis. The coefficient is a weighted average of the coefficients in Ostro and Rothschild (1989, Table 4) using the inverse of the variance as the weight:

$$\beta = \left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma^2_{\beta_i}}}{\sum\limits_{i=1976}^{1981} \frac{1}{\sigma^2_{\beta_i}}} \right) = 0.00741 .$$

The standard error of the coefficient is calculated as follows, assuming that the estimated year-specific coefficients are independent:

$$\sigma^2_\beta = \text{var}\left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma^2_{\beta_i}}}{\sum\limits_{i=1976}^{1981} \frac{1}{\sigma^2_{\beta_i}}} \right) = \left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma^2_{\beta_i}}}{\gamma} \right) = \sum\limits_{i=1976}^{1981} \text{var}\left(\frac{\beta_i}{\sigma^2_{\beta_i} \cdot \gamma} \right) .$$

This reduces down to:

$$\sigma^2_\beta = \frac{1}{\gamma} \Rightarrow \sigma_\beta = \sqrt{\frac{1}{\gamma}} = 0.00070 .$$

Incidence Rate: daily incidence rate for minor restricted activity days (MRAD) = 0.02137 (Ostro and Rothschild, 1989, p. 243)

Population: adult population ages 18 to 64

F.5.4 Schwartz and Neas (2000)

Schwartz et al. (2000) replicated a previous analysis (Schwartz et al., 1994) linking PM levels to lower respiratory symptoms in children in six cities in the U.S. The original study enrolled 1,844 children into a year-long study that was conducted in different years (1984 to 1988) in six cities. The students were in grades two through five at the time of enrollment in 1984. By the completion of the final study, the cohort would then be in the eighth grade (ages 13-14); this suggests an age range of 7 to 14. The previous study focused on PM_{10}, acid aerosols, and gaseous pollutants, although single-pollutant $PM_{2.5}$ results were reported. Schwartz et al. (2000) focused more on the associations between $PM_{2.5}$ and $PM_{10-2.5}$ and lower respiratory symptoms. In single and co-pollutant models, $PM_{2.5}$ was significantly associated with lower respiratory symptoms, while $PM_{10-2.5}$ was not. $PM_{10-2.5}$ exhibited a stronger association with cough than did $PM_{2.5}$. The $PM_{2.5}$ C-R functions for lower respiratory symptoms are based on the results of the reported single pollutant and co-pollutant model ($PM_{2.5}$ and $PM_{10-2.5}$).

Lower Respiratory Symptoms

The coefficient and standard error are calculated from the reported odds ratio (1.33) and 95% confidence interval (1.11-1.58) associated with a 15 µg/m³ change in $PM_{2.5}$ (Schwartz and Neas, 2000, Table 2).

Incidence Rate: daily lower respiratory symptom incidence rate per person = 0.0012 (Schwartz et al., 1994, Table 2)

Population: population of ages 7 to 14

F.6 Asthma-Related Effects

Table F-10 summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and asthma exacerbation. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table F-10. Health Impact Functions for Particulate Matter and Asthma-Related Effects

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Cough	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.001013	0.000768	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Cough	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.000985	0.000747	Logistic	

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Shortness of Breath	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.002636	0.001372	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Shortness of Breath	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.002565	0.001335	Logistic	
Wheeze	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.001996	0.000825	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Wheeze	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.001942	0.000803	Logistic	
Upper Respiratory Symptoms	Pope et al.	1991	Utah Valley	9-11		D24HourMean	0.003600	0.001500	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Upper Respiratory Symptoms	Pope et al.	1991	Utah Valley	9-11		D24HourMean	0.003600	0.001500	Logistic	
Cough	Vedal et al.	1998	Vancouver, CAN	6-18		D24HourMean	0.008376	0.004273	Logistic	Adjusted coefficient with 10 ug/m3 threshold.
Cough	Vedal et al.	1998	Vancouver, CAN	6-18		D24HourMean	0.008000	0.004082	Logistic	

F.6.1   Ostro et al (2001)

Ostro et al. (2001) studied the relation between air pollution in Los Angeles and asthma exacerbation in African-American children (8 to 13 years old) from August to November 1993. They used air quality data for PM_{10}, $PM_{2.5}$, NO_2, and O_3 in a logistic regression model with control for age, income, time trends, and temperature-related weather effects. The authors note that there were 26 days in which $PM_{2.5}$ concentrations were reported higher than PM_{10} concentrations. The majority of results the authors reported were based on the full dataset. These results were used for the basis for the C-R functions. Asthma symptom endpoints were defined in two ways: "probability of a day with symptoms" and "onset of symptom episodes". New onset of a symptom episode was defined as a day with symptoms followed by a symptom-free day.

The authors found cough prevalence associated with PM_{10} and $PM_{2.5}$ and cough incidence associated with $PM_{2.5}$, PM_{10}, and NO_2. Ozone was not significantly associated with cough among asthmatics. The authors found that both the prevalent and incident episodes of shortness of breath were associated with $PM_{2.5}$ and PM_{10}. Neither ozone nor NO_2 were significantly associated with shortness of breath among asthmatics. The authors found both the prevalence and incidence of wheeze associated with $PM_{2.5}$, PM_{10}, and NO_2. Ozone was not significantly associated with wheeze among asthmatics.

The derived health impact functions are based on the results of single pollutant models looking at the probability of symptoms.

Asthma Exacerbation, Cough

The coefficient and standard error are based on an odds ratio of 1.03 (95% CI 0.98-1.07) for a 30 $\mu g/m^3$ increase in 12-hour average $PM_{2.5}$ concentration (Ostro et al., 2001, Table 4, p.204).

Incidence Rate: daily cough rate per person (Ostro et al., 2001, p.202) = 0.145

Population: asthmatic African-American population ages 8 to 13 = 7.26% of African-American population ages 8 to 13. The American Lung Association (2002a, Table 9) estimates asthma prevalence for African-American children ages 5 to 17 at 7.26% (based on data from the 1999 National Health Interview Survey).

Asthma Exacerbation, Shortness of Breath

The coefficient and standard error are based on an odds ratio of 1.08 (95% CI 1.00-1.17) for a 30 $\mu g/m^3$ increase in 12-hour average $PM_{2.5}$ concentration (Ostro et al., 2001, Table 4, p.204).

Incidence Rate: daily shortness of breath rate per person (Ostro et al., 2001, p.202) = 0.074

Population: asthmatic African-American population ages 8 to 13 = 7.26% of African-American population ages 8 to 13. (Described above.)

Asthma Exacerbation, Wheeze

The coefficient and standard error are based on an odds ratio of 1.06 (95% CI 1.01-1.11) for a 30 $\mu g/m^3$ increase in 12-hour average $PM_{2.5}$ concentration (Ostro et al., 2001, Table 4, p.204).

Incidence Rate: daily wheeze rate per person (Ostro et al., 2001, p.202) = 0.173

Population: asthmatic African-American population ages 8 to 13 = 7.26% of African-American population ages 8 to 13. (Described above.)

F.6.2   Pope et al (1991)

Using logistic regression, Pope et al. (1991) estimated the impact of PM_{10} on the incidence of a variety of minor symptoms in 55 subjects (34 "school-based" and 21 "patient-based") living in the Utah Valley from December 1989 through March 1990. The children in the Pope et al. study were asked to record respiratory symptoms in a daily diary. With this information, the daily occurrences of upper respiratory symptoms (URS) and lower respiratory symptoms (LRS) were related to daily PM_{10} concentrations. Pope et al. describe URS as consisting of one or more of the following symptoms: runny or stuffy nose; wet cough; and burning, aching, or red eyes. Levels of ozone, NO_2, and SO_2 were reported low during this period, and were not included in the analysis. The sample in this study is relatively small and is most representative of the asthmatic population, rather than the general population. The school-based subjects (ranging in age from 9 to 11) were

chosen based on "a positive response to one or more of three questions: ever wheezed without a cold, wheezed for 3 days or more out of the week for a month or longer, and/or had a doctor say the 'child has asthma' (Pope et al., 1991, p. 669)." The patient-based subjects (ranging in age from 8 to 72) were receiving treatment for asthma and were referred by local physicians. Regression results for the school-based sample (Pope et al., 1991, Table 5) show PM_{10} significantly associated with both upper and lower respiratory symptoms. The patient-based sample did not find a significant PM_{10} effect. The results from the school-based sample are used here.

Upper Respiratory Symptoms

The coefficient and standard error for a one $\mu g/m^3$ change in PM_{10} is reported in Table 5.

Incidence Rate: daily upper respiratory symptom incidence rate per person = 0.3419 (Pope et al., 1991, Table 2)

Population: asthmatic population ages 9 to 11 = 5.67% of population ages 9 to 11. (The American Lung Association (2002a, Table 7) estimates asthma prevalence for children ages 5 to 17 at 5.67%, based on data from the 1999 National Health Interview Survey.)

F.6.3 Vedal et al (1998)

Vedal et al. (1998) studied the relationship between air pollution and respiratory symptoms among asthmatics and non-asthmatic children (ages 6 to 13) in Port Alberni, British Columbia, Canada. Four groups of elementary school children were sampled from a prior cross-sectional study: (1) all children with current asthma, (2) children without doctor diagnosed asthma who experienced a drop in FEV after exercise, (3) children not in groups 1 or 2 who had evidence of airway obstruction, and (4) a control group of children with matched by classroom. The authors used logistic regression and generalized estimating equations to examine the association between daily PM_{10} levels and daily increases in various respiratory symptoms among these groups. In the entire sample of children, PM_{10} was significantly associated with cough, phlegm, nose symptoms, and throat soreness. Among children with diagnosed asthma, the authors report a significant association between PM_{10} and cough symptoms, while no consistent effects were observed in the other groups. Since the study population has an over-representation of asthmatics, due to the sampling strategy, the results from the full sample of children are not generalizeable to the entire population. The C-R function presented below is based on results among asthmatics only.

Asthma Exacerbation, Cough

The PM_{10} coefficient and standard error are based on an increase in odds of 8% (95% CI 0-16%) reported in the abstract for a 10 $\mu g/m^3$ increase in daily average PM_{10}.

Incidence Rate: daily cough rate per person (Vedal et al., 1998, Table 1, p. 1038) = 0.086

Population: asthmatic population ages 6 to 13 = 5.67% of population ages 6 to 13. (The American Lung Association (2002a, Table 7) estimates asthma prevalence for children 5-17 at 5.67% (based on data from the 1999 National Health Interview Survey).)

F.7   Calculating Threshold-Adjusted Functions

Following the approach taken in OAQPS' June 2005 particulate matter (PM) risk assessment, we used a 10 ug/m3 cutpoint for short-term (daily) C-R functions and a 7.5 ug/m3 cutpoint for long-term (annual metric) C-R functions from which PM2.5 health impact functions have been derived. The risk assessment noted that 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 PM2.5 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 log-linear or logistic C-R functions, but also with modified functions that approximate non-linear, sigmoidal-shaped functions that would better reflect possible population thresholds.

However, following the approach currently being taken in OAQPS' ongoing O3 risk assessment, we did not derive threshold models for O3 concentration-response functions. After debating the merits of considering possible thresholds for O3 for hospital admissions and mortality C-R functions, OAQPS staff concluded that, because the studies report a relationship down to very low ambient levels, at or below the estimated policy relevant background (PRB) concentrations (roughly around 0.03 ppm), consideration of threshold models for O3 is not warranted. In addition, some of the O3 studies reported effects for mortality and hospital admissions in Canadian cities where the levels never exceeded the current 0.08 ppm standard.

We approximated hypothetical sigmoidal PM2.5 C-R functions by "hockeystick" functions based on the reported log-linear or logistic 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.

If the researchers in the original study fit a log-linear, linear, or logistic 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 Exhibit 3. 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 upward-sloping 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 F-13. 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 PM2.5 coefficient will be minimal, as illustrated in Figure

F-14.

Figure F-13. Relationship Between Estimated Log-Linear Concentration-Response Function and Hockeystick Model With Threshold C -- General Case
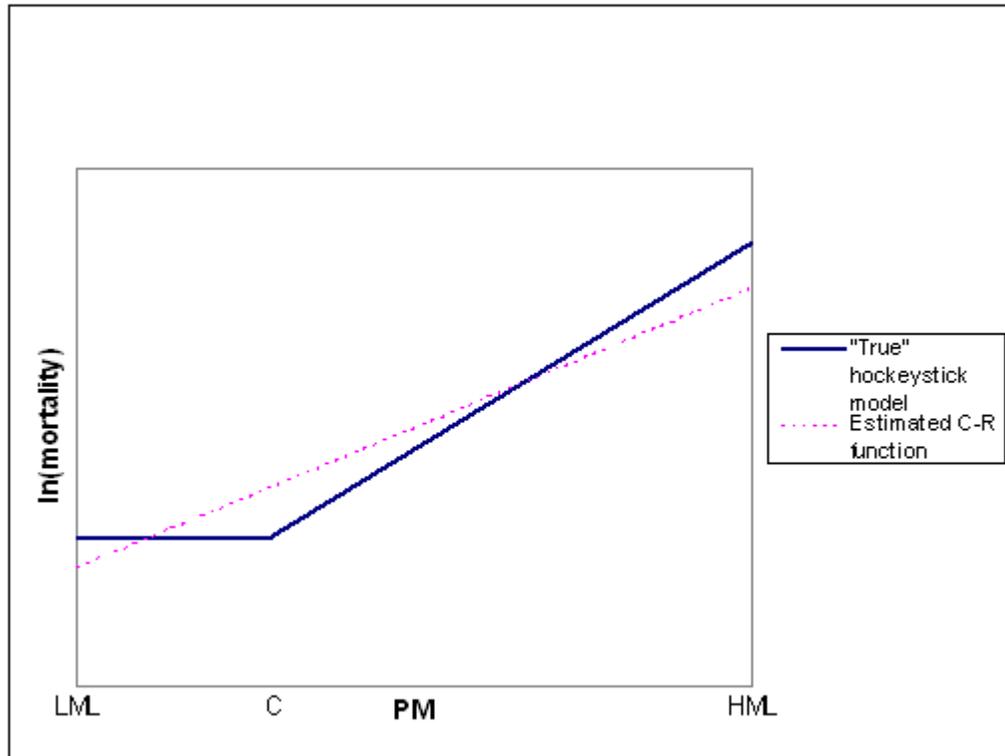
Figure F-14. Relationship Between Estimated Log-Linear Concentration-Response Function and Hockeystick Model With Threshold C -- Lowest Measured Level

(LML) Close to Hypothetical Threshold
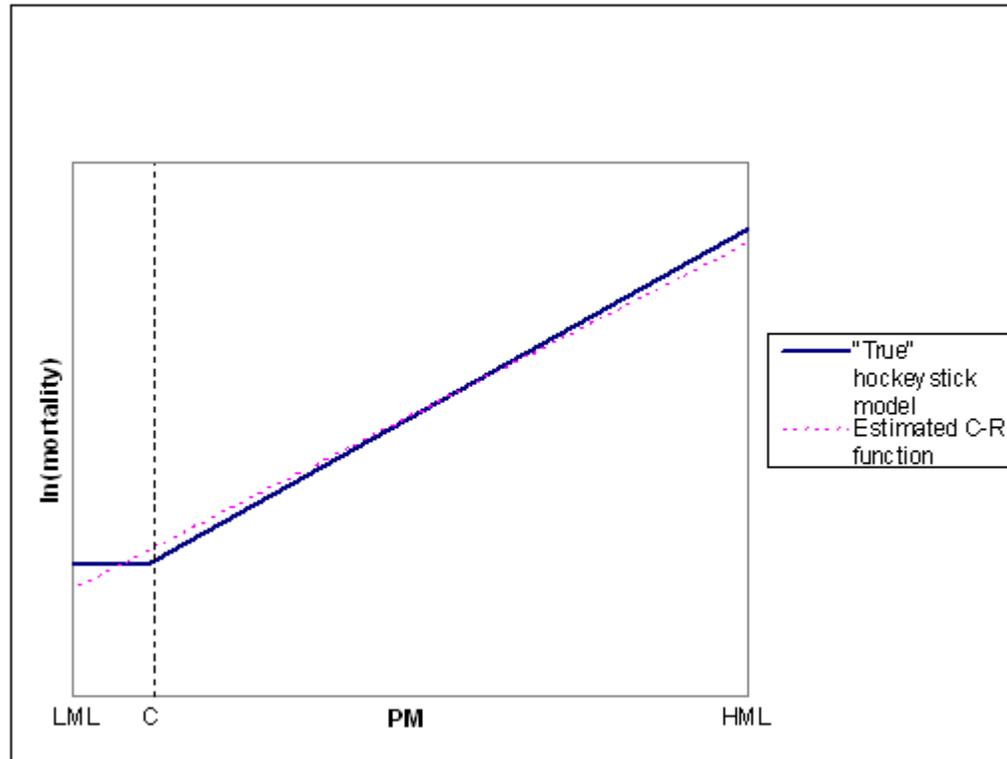
We used a simple slope adjustment method based on the idea discussed above – that, if the data in the study were best described by a hockeystick model with a cutpoint at c, then the slope estimated in the study using a log-linear or logistic model would be approximately a weighted average of the two 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 (for c > LML),[2]
- HML denote the highest measured PM level in the study,
- β^{est} denote the slope (the PM coefficient) estimated in the study (using a log-linear or logistic 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,

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

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 or logistic 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 PRB (or if it was not available for the study), the estimated PRB was substituted for LML in the above equation.

Table F-11 presents the threshold adjustments that were used to multiply with both the mean coefficient estimate and its standard error.

Table F-11. Threshold Adjustment Factors Based on Assumed Threshold of 10 ug/ m3

Author	Year	Study Location	Min	Max	Thresh old	Threshold Adj	Notes
Abbey et al.	1995	SF, SD, South Coast Air Basin			10	1.000	
Dockery et al.	1996	24 communities	5.8	20.7	10	1.393	
Ito	2003	Detroit, MI	6	42	10	1.125	Min and max based on 5% and 95%.
Laden et al.	2006	6 cities	10.8	25.5	10	1.000	
Moolgavkar	2003	Los Angeles, CA	4	86	10	1.079	
Norris et al.	1999	Seattle, WA	9	18.2	10	1.122	

Ostro	1987	Nationwide			10	1.000	Study did not provide a mean, SD, or pollutant range.
Ostro and Rothschild	1989	Nationwide			10	1.000	Study gave mean and std dev. We estimate min is above threshold of 10.
Ostro et al.	2001	Los Angeles, CA	4.5	208.7	10	1.028	
Peters et al.	2001	Boston, MA	4.6	24.3	10	1.378	Min and max based on 5% and 95%.
Pope et al.	1991	Utah Valley	11	195	10	1.000	
Pope et al.	2002	51 cities	7.5	30	7.5	1.000	Mean: 17.7 & SD: 3.7. PM risk assessment assumed min = 7.5 and max = 30
Pope et al.	2002	51 cities	7.5	30	10	1.125	
Pope et al.	2002	51 cities	7.5	30	12	1.250	
Pope et al.	2002	51 cities	7.5	30	15	1.500	
Schwartz and Neas	2000	6 U.S. cities	7.2	86	10	1.037	
Sheppard	2003	Seattle, WA	6	32	10	1.182	Min and max based on 5% and 95%.
Vedal et al.	1998	Vancouver, CAN	3	159	10	1.047	Min = policy-relevant background. Actual min was 0.2 in North and 0.5 in South.
Woodruff et al.	1997	86 cities	11.9	68.8	10	1.000	
Woodruff et al.	2006	204 counties			10	1.000	Only presented interquartile PM2.5 range

Appendix G: Ozone Health Impact Functions in U.S. Setup

In this Appendix, we present the health impact functions used to estimate ozone-related adverse health effects. Each sub-section has a table with a brief description of each health impact function and the underlying parameters. Following each table, we present a brief summary of each of the studies and any items that are unique to the study.

Note that Appendix D mathematically derives the standard types of health impact functions encountered in the epidemiological literature, such as, log-linear, logistic and linear, so we simply note here the type of functional form. And Appendix E presents a description of the sources for the incidence and prevalence data used in the health impact functions.

G.1   Short-term Mortality

Table G-1 summarizes the health impacts functions used to estimate the relationship between ozone and short-term mortality. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table G-1. Health Impact Functions for Ozone and Short-Term Mortality

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Non-Accidental	Bell et al.	2004	95 US cities	0-99		D24HourMean	0.000390	0.000133	Log-linear	Warm season.
Non-Accidental	Bell et al.	2004	95 US cities	0-99		D24HourMean	0.000520	0.000128	Log-linear	All year.
Non-Accidental	Bell et al.	2004	95 US cities	0-99		D8HourMax	0.000261	0.000089	Log-linear	Warm season. 8-hour max from 24-hour mean.
All Cause	Bell et al.	2005	US & non-US	0-99		D24HourMean	0.001500	0.000401	Log-linear	Warm season.
All Cause	Bell et al.	2005	US & non-US	0-99		D8HourMax	0.000795	0.000212	Log-linear	Warm season. 8-hour max from 24-hour mean.
Cardiopulmonary	Huang et al.	2005	19 US cities	0-99		D24HourMean	0.001250	0.000398	Log-linear	Warm season.
Cardiopulmonary	Huang et al.	2005	19 US cities	0-99		D8HourMax	0.000813	0.000259	Log-linear	Warm season. 8-hour max from 24-hour mean.
Non-Accidental	Ito and Thurston	1996	Chicago, IL	18-99	PM10	D1HourMax	0.000634	0.000251	Log-linear	
Non-Accidental	Ito et al.	2005		0-99		D1HourMax	0.000400	0.000066	Log-linear	1-hour max.
Non-Accidental	Ito et al.	2005		0-99		D24HourMean	0.001750	0.000357	Log-linear	Warm season. 24-hour mean.

Non-Accidental	Ito et al.	2005		0-99		D8HourMax	0.001173	0.000239	Log-linear	Warm season. 8-hour max from 24-hour mean.
Non-Accidental	Ito et al.	2005		0-99		D8HourMax	0.000532	0.000088	Log-linear	8-hour max from 1-hour max.
All Cause	Levy et al.	2005	US and non-US	0-99		D1HourMax	0.000841	0.000134	Log-linear	Warm season.
All Cause	Levy et al.	2005	US and non-US	0-99		D8HourMax	0.001119	0.000179	Log-linear	Warm season. 8-hour max from 1-hour max.
Non-Accidental	Moolgavkar et al.	1995	Philadelphia, PA	0-99		D24HourMean	0.001398	0.000266	Log-linear	Warm season.
Non-Accidental	Moolgavkar et al.	1995	Philadelphia, PA	0-99	TSP, SO2	D24HourMean	0.001389	0.000373	Log-linear	Warm season.
Non-Accidental	Moolgavkar et al.	1995	Philadelphia, PA	18-99	TSP, SO2	D24HourMean	0.000611	0.000216	Log-linear	
Non-Accidental	Samet et al.	1997	Philadelphia, PA	18-99	CO, NO2, SO2, TSP	D24HourMean	0.000936	0.000312	Log-linear	
Non-Accidental	Schwartz	2005	14 US cities	0-99		D1HourMax	0.000370	0.000130	Logistic	Warm season.
Non-Accidental	Schwartz	2005	14 US cities	0-99		D8HourMax	0.000426	0.000150	Logistic	Warm season. 8-hour max from 1-hour max.

G.1.1 Bell et al (2004)

Ozone has been associated with various adverse health effects, including increased rates of hospital admissions and exacerbation of respiratory illnesses. Although numerous time-series studies have estimated associations between day-to-day variation in ozone levels and mortality counts, results have been inconclusive. The authors investigated whether short-term (daily and weekly) exposure to ambient ozone is associated with mortality in the United States. Using analytical methods and databases developed for the National Morbidity, Mortality, and Air Pollution Study, they estimated a national average relative rate of mortality associated with short-term exposure to ambient ozone for 95 large US urban communities from 1987-2000. The authors used distributed-lag models for estimating community-specific relative rates of mortality adjusted for time-varying confounders (particulate matter, weather, seasonality, and long-term trends) and hierarchical models for combining relative rates across communities to estimate a national average relative rate, taking into account spatial heterogeneity. A 10-ppb increase in the previous week's ozone was associated with a 0.52% increase in daily mortality (95% posterior interval [PI], 0.27%-0.77%) and a 0.64% increase in cardiovascular and respiratory mortality (95% PI, 0.31%-0.98%). Effect estimates for aggregate ozone during the previous week were larger than for models considering only a single day's exposure. Results were robust to adjustment for particulate matter, weather, seasonality, and long-term trends. These results indicate a statistically significant association between short-term changes in ozone and mortality on average for 95 large US urban communities, which include about 40% of the total US population.

Non-Accidental Mortality

The coefficient and standard error are based on the relative risk (1.003908) and 95% confidence interval (1.0013-1.0065) associated with a 10 ppb increase in daily average ozone (Bell et al., 2004, p. 2376).

G.1.2   Bell et al (2005)

Although many time-series studies of ozone and mortality have identified positive associations, others have yielded null or inconclusive results, making the results of these studies difficult to interpret. The authors performed a meta-analysis of 144 effect estimates from 39 time-series studies, and estimated pooled effects by lags, age groups, cause-specific mortality, and concentration metrics. They compared results with pooled estimates from the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), a time-series study of 95 large U.S. urban centers from 1987 to 2000. Both meta-analysis and NMMAPS results provided strong evidence of a short-term association between ozone and mortality, with larger effects for cardiovascular and respiratory mortality, the elderly, and current-day ozone exposure. In both analyses, results were insensitive to adjustment for particulate matter and model specifications. In the meta-analysis, a 10-ppb increase in daily ozone at single-day or 2-day average of lags 0, 1, or 2 days was associated with an 0.87% increase in total mortality (95% posterior interval = 0.55% to 1.18%), whereas the lag 0 NMMAPS estimate is 0.25% (0.12% to 0.39%). Several findings indicate possible publication bias: meta-analysis results were consistently larger than those from NMMAPS; meta-analysis pooled estimates at lags 0 or 1 were larger when only a single lag was reported than when estimates for multiple lags were reported; and heterogeneity of city-specific estimates in the meta-analysis were larger than with NMMAPS.

All-Cause Mortality

The coefficient and standard error are based on the relative risk (1.008738) and 95% confidence interval (1.0055-1.0119) associated with a 10 ppb increase in daily average ozone (Bell et al., 2005, Table 6).

G.1.3   Huang et al (2005)

The authors developed Bayesian hierarchical distributed lag models for estimating associations between daily variations in summer ozone levels and daily variations in cardiovascular and respiratory (CVDRESP) mortality counts for 19 large U.S. cities included in the National Morbidity, Mortality and Air Pollution Study (NMMAPS) for the summers of 1987–1994. In the first stage, they defined a semi-parametric distributed lag Poisson regression model to estimate city-specific relative rates of CVDRESP mortality associated with short-term exposure to summer ozone. In the second stage, they specified a class of distributions for the true city-specific relative rates to estimate an overall effect by taking into account the variability within and across cities. They performed the calculations with respect to several random effects distributions (normal, t-student, and

mixture of normal), thus relaxing the common assumption of a two-stage normal–normal hierarchical model. They assessed the sensitivity of the results to: (i) lag structure for ozone exposure; (ii) degree of adjustment for long-term trends; (iii) inclusion of other pollutants in the model; (iv) heat waves; (v) random effects distributions; and (vi) prior hyperparameters. On average across cities, the authors found that a 10 ppb increase in summer ozone level over the previous week is associated with a 1.25 per cent increase in CVDRESP mortality (95 per cent posterior regions: 0.47, 2.03). The relative rate estimates are also positive and statistically significant at lags 0, 1 and 2. They found that associations between summer ozone and CVDRESP mortality are sensitive to the confounding adjustment for PM10, but are robust to: (i) the adjustment for long-term trends, other gaseous pollutants (NO2, SO2 and CO); (ii) the distributional assumptions at the second stage of the hierarchical model; and (iii) the prior distributions on all unknown parameters.

Cardiopulmonary Mortality

Assuming a 10 ppb change in ozone, Huang et al (2005, Table 1) reported a 1.25% change in CVDRESP mortality with a 95% confidence interval of 0.47% to 2.03%.

Note that Huang et al (2005, p. 549) define CVDRESP as including ICD-9 codes: 390-448, 480-487, 490-496, and 507. This differs somewhat from the the definition of "cardiopulmonary" mortality in BenMAP -- defined as ICD-9 codes 401-440 and 460-519.

G.1.4 Levy et al, 2005

The authors conducted an empiric Bayes metaregression to estimate the ozone effect on mortality, and to assess whether this effect varies as a function of hypothesized confounders or effect modifiers. They gathered 71 time-series studies relating ozone to all-cause mortality, and tjey selected 48 estimates from 28 studies for the metaregression. Metaregression covariates included the relationship between ozone concentrations and concentrations of other air pollutants, proxies for personal exposure-ambient concentration relationships, and the statistical methods used in the studies. For the metaregression, they applied a hierarchical linear model with known level-1 variances. The authors estimated a grand mean of a 0.21% increase (95% confidence interval = 0.16-0.26%) in mortality per 10-microg/m increase of 1-hour maximum ozone (0.41% increase per 10 ppb) without controlling for other air pollutants. In the metaregression, air-conditioning prevalence and lag time were the strongest predictors of between-study variability. Air pollution covariates yielded inconsistent findings in regression models, although correlation analyses indicated a potential influence of summertime PM2.5.

All-Cause Mortality

Levy et al (2005, Table 1) reported a 0.43% change in all-cause mortality with a 95% confidence interval of 0.29% to 0.56% associated with a 10 ug/m3 change in ozone. We converted ug/m3 to ppb with an assumed relationship of 1.96 ug/m3 per 1.0 ppb.

G.1.5   Ito and Thurston (1996)

In this study, race, gender, and cause-specific counts of daily mortality in Cook County, Illinois (which encompasses the city of Chicago) during 1985-1990 were analyzed to determine if there was any heterogeneity in air pollution/weather/mortality associations across these various population subcategories. Seasonal cross-correlations between mortality and environmental variables first were examined to identify appropriate lag structures. Of the pollution variables considered -- PM10, ozone, CO, SO2, and visual range-derived extinction coefficient -- both PM10 and ozone showed significant associations with same-day and next-day mortality. The Poisson regression models employed included seasonal cycles (sine/cosine series), square and linear terms of lagged temperature, trend line, day-of-week dummy variables, and the average of the same day's and previous day's PM10 or ozone.

The authors reported a significant relationship for ozone and PM_{10} with both pollutants in the model; no significant effects were found for SO_2 and CO. In single pollutant models the effects were slightly larger. The health impact function for ozone is based on results from the co-pollutant models.

Non-Accidental Mortality

For a co-pollutant model with PM_{10}, the ozone coefficient (0.000634) and standard error (0.000251) were obtained directly from the author because the published paper reported incorrect information.

G.1.6   Ito et al (2005)

The authors conducted a review and meta-analysis of short-term ozone mortality studies, identified unresolved issues, and conducted an additional time-series analysis for 7 U.S. cities (Chicago, Detroit, Houston, Minneapolis-St. Paul, New York City, Philadelphia, and St. Louis). They found a combined estimate of 0.39% (95% confidence interval = 0.26-0.51%) per 10-ppb increase in 1-hour daily maximum ozone for the all-age nonaccidental cause/single pollutant model (43 studies). Adjusting for the funnel plot asymmetry resulted in a slightly reduced estimate (0.35%; 0.23-0.47%). In a subset for which particulate matter (PM) data were available (15 studies), the corresponding estimates were 0.40% (0.27-0.53%) for ozone alone and 0.37% (0.20-0.54%) with PM in model. The estimates for warm seasons were generally larger than those for cold seasons. The additional time-series analysis found that including PM in the model did not substantially reduce the ozone risk estimates. However, the difference in the weather adjustment model could result in a 2-fold difference in risk estimates (eg, 0.24% to 0.49% in multicity combined estimates across alternative weather models for the ozone-only all-year case). The authors concluded that the results suggest short-term associations between ozone and daily mortality in the majority of the cities, although the estimates

appear to be heterogeneous across cities.

Non-Accidental Mortality

Ito et al (2005) reported results for functions with both 1-hour daily maximum and 24-hour daily average metrics. We present both below.

One-hour Max Function

Assuming a 10 ppb change in the daily 1-hour maximum, Ito et al (2005, p. 446) reported a 0.40% change in non-accidental mortality with a 95% confidence interval of 0.27% to 0.53%.

Daily Average Function

Assuming a 20 ppb change in the daily 24-hour average, Ito et al (2005, p. 448) reported a 3.5% change in non-accidental mortality with a 95% confidence interval of 2.1% to 4.9%.

G.1.7 Moolgavkar et al (1995)

Moolgavkar et al. (1995) examined the relationship between daily non-accidental mortality and air pollution levels in Philadelphia, Pennsylvania from 1973 to 1988. They examined ozone, TSP, and SO_2 in a three-pollutant model, and found a significant relationship for ozone and SO_2; TSP was not significant. In season-specific models, ozone was significantly associated with mortality only in the summer months.

Mortality, Non-Accidental

The health impact function for ozone is based on the full-year three-pollutant model reported in Table 5 (Moolgavkar et al., 1995, p. 482). The coefficient and standard error are based on the relative risk (1.063) and 95% confidence interval (1.018-1.108) associated with a 100 ppb increase in daily average ozone.

G.1.8 Samet et al (1997)

Samet et al. (1997) examined the relationship between daily non-accidental mortality and air pollution levels in Philadelphia, Pennsylvania from 1974 to 1988. They examined ozone, TSP, SO_2, NO_2, and CO in a Poisson regression model. In single pollutant models, ozone, SO_2, TSP, and CO were significantly associated with mortality. In a five-pollutant model, they found a positive statistically significant relationship for each pollutant except

NO_2.

Mortality, Non-Accidental

The health impact function for ozone is based on the five-pollutant model (ozone, CO, NO_2, SO_2, and TSP) reported in Table 9 (Samet et al., 1997, p. 20). The ozone coefficient and standard error are based on the percent increase (1.91) and t-statistic (3) associated with a 20.219 ppb increase in two-day average ozone.

G.1.9   Schwartz (2005)

The author used the case-crossover approach, where the control for each person is the same person on a day near in time, when he or she did not die. This method controls for season and individual risk factors by matching. One can also choose the control day to have the same temperature as the event day. The author applied this approach to a study of more than 1 million deaths in 14 U.S. cities. He found that, with matching on temperature, a 10-ppb increase in maximum hourly ozone concentrations was associated with a 0.23% (95% confidence interval [CI] 0.01%, 0.44%) increase in the risk of dying. This finding was indistinguishable from the risk when only matching on season and controlling for temperature with regression splines (0.19%; 95% CI 03%, 0.35%). Control for suspended particulate matter with an aerodynamic diameter of 10 mum or less (PM(10)) did not change this risk. However, the association was restricted to the warm months (0.37% increase; 95% CI 0.11%, 0.62%), with no effect in the cold months. The author concluded that the association between ozone and mortality risk is unlikely to be caused by confounding by temperature.

Non-Accidental Mortality

Assuming a 10 ppb change in the daily 1-hour maximum, Schwartz (2005, Table 2) reported a 0.37% change in non-accidental mortality with a 95% confidence interval of 0.11% to 0.62%.

G.2   Hospital Admissions

Table G-2 summarizes the health impacts functions used to estimate the relationship between ozone and hospital admissions. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table G-2. Health Impact Functions for Ozone and Hospital Admissions

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes

All Respiratory	Burnett et al.	2001	Toronto, CAN	0-1	PM2.5	D1HourMax	0.007301	0.002122	Log-linear	Warm season.
All Respiratory	Burnett et al.	2001	Toronto, CAN	0-1	PM2.5	D8HourMax	0.008177	0.002377	Log-linear	Warm season. 8-hour max from 1-hour max.
Chronic Lung	Moolgavkar et al.	1997	Minneapolis, MN	65-99	PM10, CO	D24HourMean	0.002800	0.001769	Log-linear	
Chronic Lung	Moolgavkar et al.	1997	Minneapolis, MN	65-99	PM10, CO	D8HourMax	0.001960	0.001238	Log-linear	All year. 8-hour max from 24-hour mean.
Pneumonia	Moolgavkar et al.	1997	Minneapolis, MN	65-99	PM10, SO2, NO2	D24HourMean	0.003800	0.001088	Log-linear	
Pneumonia	Moolgavkar et al.	1997	Minneapolis, MN	65-99	PM10, SO2, NO2	D8HourMax	0.002660	0.000762	Log-linear	All year. 8-hour max from 24-hour mean.
Chronic Lung (less Asthma)	Schwartz	1994	Detroit, MI	65-99	PM10	D24HourMean	0.005523	0.002085	Log-linear	All year.
Chronic Lung (less Asthma)	Schwartz	1994	Detroit, MI	65-99	PM10	D8HourMax	0.003424	0.001293	Log-linear	All year. 8-hour max from 24-hour mean.
Pneumonia	Schwartz	1994	Detroit, MI	65-99	PM10	D24HourMean	0.005210	0.001300	Log-linear	All year.
Pneumonia	Schwartz	1994	Minneapolis, MN	65-99	PM10	D24HourMean	0.003977	0.001865	Log-linear	All year.
Pneumonia	Schwartz	1994	Detroit, MI	65-99	PM10	D8HourMax	0.003230	0.000806	Log-linear	All year. 8-hour max from 24-hour mean.
Pneumonia	Schwartz	1994	Minneapolis, MN	65-99	PM10	D8HourMax	0.002784	0.001305	Log-linear	All year. 8-hour max from 24-hour mean.
All Respiratory	Schwartz	1995	New Haven, CT	65-99	PM10	D24HourMean	0.002652	0.001398	Log-linear	Warm season.
All Respiratory	Schwartz	1995	Tacoma, WA	65-99	PM10	D24HourMean	0.007147	0.002565	Log-linear	Warm season.
All Respiratory	Schwartz	1995	New Haven, CT	65-99	PM10	D8HourMax	0.001777	0.000936	Log-linear	Warm season. 8-hour max from 24-hour mean.
All Respiratory	Schwartz	1995	Tacoma, WA	65-99	PM10	D8HourMax	0.004931	0.001770	Log-linear	Warm season. 8-hour max from 24-hour mean.

G.2.1   Burnett et al (2001)

Burnett et al. (2001) studied the association between air pollution and acute respiratory hospital admissions (ICD codes 493, 466, 464.4, 480-486) in Toronto from 1980-1994, among children less than 2 years of age. They collected hourly concentrations of the gaseous pollutants, CO, NO_2, SO_2, and ozone. Daily measures of particulate matter were estimated for the May to August period of 1992-1994 using TSP, sulfates, and coefficient of haze data. The authors report a positive association between ozone in the May through August months and respiratory hospital admissions, for several single days after elevated ozone levels.

The strongest association was found using a five-day moving average of ozone. No association was found in the September through April months. In co-pollutant models with a particulate matter or another gaseous pollutant, the ozone effect was only slightly diminished. The effects for PM and gaseous pollutants were generally significant in single

pollutant models but diminished in co-pollutant models with ozone, with the exception of CO. The C-R functions for ozone are based on a single pollutant and two co-pollutant models, using the five-day moving average of one-hour max ozone.

Hospital Admissions, All Respiratory (ICD-9 codes 464, 466, 480-487, 493)

In a model with $PM_{2.5}$, the coefficient and standard error are based on the percent increase (33.0) and t-statistic (3.44) associated with a 45.2 ppb increase in the five-day moving average of one-hour max ozone (Burnett et al., 2001, Table 3).

G.2.2   Moolgavkar et al (1997)

Moolgavkar et al. (1997) examined the relationship between air pollution and hospital admissions (ICD codes 490-496) for individuals 65 and older in Minneapolis-St. Paul, Minnesota, from January 1986 to December 1991. In a Poisson regression, they found no significant effect for any of the pollutants (PM_{10}, ozone, or CO). The effect for ozone was marginally significant. The model with a 100 df smoother was reported to be optimal (p. 368). The health impact function for chronic lung disease is based on the results from a three-pollutant model (ozone, CO, PM_{10}) using the 100 df smoother; the function for Pneumonia uses the 130 df smoother.

Hospital Admissions, Chronic Lung Disease (ICD-9 codes 490-496)

In a model with CO and PM_{10}, the estimated coefficient and standard error are based on the percent increase (4.2) and 95% confidence interval of the percent increase (-1.0-9.4) associated with a change in daily average ozone levels of 15 ppb (Moolgavkar et al., 1997, Table 4).

Hospital Admissions, Pneumonia (ICD-9 codes 480-487)

In a model with NO_2, PM_{10}, and SO_2, the estimated coefficient and standard error are based on the percent increase (5.7) and 95% confidence interval of the percent increase (2.5-8.9) associated with an increase in daily average ozone levels of 15 ppb (Moolgavkar et al., 1997, Table 4).

G.2.3   Schwartz (1994a)

Schwartz (1994a) examined the relationship between air pollution and hospital admissions for individuals 65 and older in Minneapolis-St. Paul, Minnesota, from January 1986 to December 1989. In single-pollutant Poisson regression models, both ozone and PM_{10} were significantly associated with pneumonia admissions. In a two-pollutant model, Schwartz found PM_{10} significantly related to pneumonia; ozone was weakly linked to pneumonia.

The results were not sensitive to the methods used to control for seasonal patterns and weather. The ozone C-R functions are based on the results of the single pollutant model and the two-pollutant model (PM_{10} and ozone) with spline smoothing for temporal patterns and weather.

Hospital Admissions, Pneumonia (ICD-9 codes 480-487)

In a model with PM_{10} and spline functions to adjust for time and weather, the coefficient and standard error are based on the relative risk (1.22) and 95% confidence interval (1.02, 1.47) for a 50 ppb increase in daily average ozone levels (Schwartz, 1994a, Table 4).

G.2.4 Schwartz (1994b)

Schwartz (1994b) examined the relationship between air pollution and hospital admissions (ICD codes 491-492, 494-496) for individuals 65 and older in Detroit, Michigan, from January 1986 to December 1989. In a two-pollutant Poisson regression model, Schwartz found both PM_{10} and ozone significantly linked to pneumonia and COPD. The authors state that effect estimates were relatively unchanged compared to the unreported single pollutant models. No significant associations were found between either pollutant and asthma admissions. The C-R function for chronic lung disease incidence is based on the results of the "basic" co-pollutant model (ozone and PM_{10}) presented in Table 4 (p. 651). The study also reports results using generalized additive models to fit time and temperature variables, however no standard error or confidence intervals were reported.

Hospital Admissions, Chronic Lung Disease less Asthma (ICD-9 codes 490-492, 494-496)

The coefficient and standard error for the "basic" model are reported in Table 4 (Schwartz, 1994b, p.651) for a one ppb change in daily average ozone.

Hospital Admissions, Pneumonia (ICD-9 codes 480-487)

The ozone C-R function for pneumonia incidence is based on the coefficient and standard error for the "basic" co-pollutant model presented in Table 4 (Schwartz, 1994b, p. 651).

G.2.5 Schwartz (1995)

Studies have reported associations between short term changes in air pollution and respiratory hospital admissions. This relationship was examined in two cities with substantially different levels of sulphur dioxide (SO2) but similar levels of airborne particles in an attempt to separate the effects of the two pollutants. Significant differences

in weather between the two cities allowed the evaluation of that potential confounder also. Daily counts of admissions to all hospitals for respiratory disease (ICD 9 460-519) were constructed for persons aged 65 years and older in two cities - New Haven, Connecticut and Tacoma, Washington.

Each city was analysed separately. Average daily concentrations of SO2, inhalable particles (PM10), and ozone were computed from all monitors in each city, and daily average temperature and humidity were obtained from the US weather service. Daily respiratory admission counts were regressed on temperature, humidity, day of the week indicators, and air pollution. A 19 day weighted moving regression filter was used to remove all seasonal and subseasonal patterns from the data. Possible U- shaped dependence of admissions on temperature was dealt with using indicator variables for eight categories each of temperature and humidity.

Each pollutant was first examined individually and then multiple pollutant models were fitted. All three pollutants were associated with respiratory hospital admissions of the elderly. The PM10 associations were little changed by control for either ozone or SO2. The ozone association was likewise independent of the other pollutants. The SO2 association was substantially attenuated by control for ozone in both cities, and by control for PM10 in Tacoma. The magnitude of the effect was small (relative risk 1.06 in New Haven and 1.10 in Tacoma for a 50 micrograms/m3 increase in PM10, for example) but, given the ubiquitous exposure, this has some public health significance. The authors concluded that air pollution concentrations within current guidelines were associated with increased respiratory hospital admissions of the elderly. The strongest evidence for an independent association was for PM10, followed by ozone.

Hospital Admissions, All Respiratory (ICD-9 codes 460-519) -- Tacoma

In a model with PM_{10}, the coefficient and standard error are estimated from the relative risk (1.20) and 95% confidence interval (1.06-1.37) for a 50 $\mu g/m^3$ increase in average daily ozone levels (Schwartz, 1995, Table 6, p. 535). To calculate the coefficient, a conversion of 1.96 $\mu g/m^3$ per ppb was used, based on a density of ozone of 1.96 grams per liter (at 25 degrees Celsius).

Hospital Admissions, All Respiratory (ICD-9 codes 460-519) -- New Haven

In a model with PM_{10}, the coefficient and standard error are estimated from the relative risk (1.07) and 95% confidence interval (1.00-1.15) for a 50 $\mu g/m^3$ increase in average daily ozone levels (Schwartz, 1995, Table 3, p. 534). To calculate the coefficient, a conversion of 1.96 $\mu g/m^3$ per ppb was used, based on a density of ozone of 1.96 grams per liter (at 25 degrees Celsius).

G.3 Emergency Room Visits

Table G-3 summarizes the health impacts functions used to estimate the relationship between ozone and emergency room (ER) visits. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table G-3. Health Impact Functions for Ozone and Emergency Room Visits

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma	Jaffe et al.	2003	Ohio cities	5-34		D8HourMax	0.003000	0.001531	Log-linear	
Asthma	Peel et al.	2005	Atlanta, GA	0-99		D8HourMax	0.000870	0.000529	Log-linear	
Asthma	Stieb et al.	1996	New Brunswick, CAN	0-99		D1HourMax	0.000040	0.000020	Quadratic	Warm season.
Asthma	Stieb et al.	1996	New Brunswick, CAN	0-99		D24HourMean	0.000100	0.000040	Quadratic	Warm season.
Asthma	Wilson et al.	2005	Portland, ME	0-99		D8HourMax	0.003000	0.001000	Log-linear	
Asthma	Wilson et al.	2005	Manchester, NH	0-99		D8HourMax	-0.001000	0.002000	Log-linear	

G.3.1 Jaffe et al (2003)

Jaffe et al. (2003) examined the relationship between ER visits and air pollution for persons ages 5-34 in Cleveland, Columbus, and Cincinnati, Ohio, from 1991 through 1996. In single-pollutant Poisson regression models, ozone and SO_2 were linked to asthma visits, and no significant effect was seen for NO_2 and PM_{10}.

Emergency Room Visits, Asthma

Assuming a 10 ppb increase in the daily 8-hour maximum ozone level, Jaffe et al (2003, Table 3) reported a 3.0% change in asthma ER visits with a 95% confidence interval of 0.0% to 6.0%.

G.3.2 Peel et al (2005)

A number of emergency department studies have corroborated findings from mortality and hospital admission studies regarding an association of ambient air pollution and respiratory outcomes. More refined assessment has been limited by study size and available air quality data. Measurements of 5 pollutants PM10, ozone, NO2, CO, and SO2 were available for the entire study period (1 January 1993 to 31 August 2000); detailed measurements of particulate matter were available for 25 months. The authors obtained data on 4 million emergency department visits from 31 hospitals in Atlanta. Visits for asthma, chronic obstructive pulmonary disease, upper respiratory infection, and pneumonia were assessed in relation to air pollutants using Poisson generalized estimating equations. In single-pollutant models examining 3-day moving averages of pollutants (lags 0, 1, and 2): standard deviation increases of ozone, NO2, CO, and PM10 were associated with 1-3%

increases in URI visits; a 2 microg/m increase of PM2.5 organic carbon was associated with a 3% increase in pneumonia visits; and standard deviation increases of NO2 and CO were associated with 2-3% increases in chronic obstructive pulmonary disease visits. Positive associations persisted beyond 3 days for several of the outcomes, and over a week for asthma. The results of this study contribute to the evidence of an association of several correlated gaseous and particulate pollutants, including ozone, NO2, CO, PM, and organic carbon, with specific respiratory conditions.

Emergency Room Visits, Asthma

The ozone coefficient and standard error are reported per 25 ppb increment of the maximum daily 8-hour average ozone level (Peel et al., 2003, Table 4). We used the results from the three cities combined. The relative risk is 1.022, with a 95 percent confidence interval of 0.996 to 1.049.

G.3.3 Stieb (1996)

Stieb et al. (1996) examined the relationship between ER visits and air pollution for persons of all ages in St. John, New Brunswick, Canada, from May through September in 1984-1992. Ozone was significantly linked to ER visits, especially when ozone levels exceeded 75 ppb. The authors reported results from a linear model, quadratic model, and linear-quadratic model using daily average and 1-hour maximum ozone. In the linear model, ozone was borderline significant. In the quadratic and linear-quadratic models, ozone was highly significant. This is consistent with the author's conclusion that "only ozone appeared to have a nonlinear relationship with visit rates" (p. 1356) and that "quadratic, linear-quadratic, and indicator models consistently fit the data better than the linear model ..." (p. 1358). The linear term in the linear-quadratic model is negative, implying that at low ozone levels, increases in ozone are associated with decreases in risk. Since this does not seem biologically plausible, the ozone health impact functions described here are based on the results of the quadratic regression models presented in Table 2 (Stieb et al., 1996, p. 1356).

Emergency Room Visits, Asthma

<u>One-hour Max Function</u>

The coefficient and standard error of the quadratic model are reported in Table 2 (Stieb et al., 1996, p. 1356) for a 1 ppb increase in 1-hour daily maximum ozone levels. The C-R function to estimate avoided emergency visits derived from a quadratic regression model is shown below:

$$\triangle Asthma\,ER\,Visits = \frac{\beta}{BasePop} \cdot [(O_{3,baseline})^2 - (O_{3,control})^2]\,pop,$$

Baseline Population: baseline population of St. John, New Brunswick (Stieb et al., 1996, p. 1354) = 125,000

Population: population of all ages

Daily Average Function

The coefficient and standard error of the quadratic model are reported in Table 2 (p. 1356) for a 1 ppb increase in daily average ozone levels. The C-R function to estimate avoided emergency visits derived from a quadratic regression model is shown below:

$$\triangle Asthma\,ER\,Visits = \frac{\beta}{BasePop} [(O_{3,baseline})^2 - (O_{3,control})^2]\,pop,$$

Baseline Population: baseline population of St. John, New Brunswick (Stieb et al., 1996, p. 1354) = 125,000

Population: population of all ages

G.3.4   Wilson et al (2005)

Daily emergency room (ER) visits for all respiratory (ICD-9 460-519) and asthma (ICD-9 493) were compared with daily SO2, ozone, and weather variables over the period 1998-2000 in Portland, Maine (population 248,000), and 1996-2000 in Manchester, New Hampshire (population 176,000). Seasonal variability was removed from all variables using nonparametric smoothed function (LOESS) of day of study. Generalized additive models were used to estimate the effect of elevated levels of pollutants on ER visits. Relative risks of pollutants were reported over their interquartile range (IQR, the 75th -25th percentile pollutant values). I n Portland, an IQR increase in SO2 was associated with a 5% (95% CI 2-7%) increase in all respiratory ER visits and a 6% (95% CI 1-12%) increase in asthma visits. An IQR increase in O3 was associated with a 5% (95% CI 1-10%) increase in Portland asthmatic ER visits. No significant associations were found in Manchester, New Hampshire, possibly due to statistical limitations of analyzing a smaller population. The absence of statistical evidence for a relationship should not be used as evidence of no relationship. This analysis reveals that, on a daily basis, elevated SO2 and O3 have a significant impact on public health in Portland, Maine.

Emergency Room Visits, Asthma

The coefficient and standard error are taken from Wilson et al. (2005, Table 5).

G.4   Minor Effects

Table G-4 summarizes the health impacts functions used to estimate the relationship between ozone and minor effects. Below, we present a brief summary of each of the studies and any items that are unique to the study.

Table G-4. Health Impact Functions for Ozone and Minor Effects

Effect	Author	Year	Lcoation	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
School Loss Days, All Cause	Chen et al.	2000	Washoe Co, NV	5-17	PM10, CO	D1HourMax	0.013247	0.004985	Linear	
School Loss Days, All Cause	Chen et al.	2000	Washoe Co, NV	5-17	PM10, CO	D8HourMax	0.015763	0.004985	Linear	All year. 8-hour max from 1-hour max.
Worker Productivity	Crocker and Horst	1981	Nationwide	18-64		D24HourMean	0.142700		Linear	All year.
Worker Productivity	Crocker and Horst	1981	Nationwide	18-64		D8HourMax	0.092755		Linear	All year. 8-hour max from 24-hour mean.
School Loss Days, All Cause	Gilliland et al.	2001	Southern California	5-17		D8HourMax	0.007824	0.004445	Log-linear	All year. 8-hour max from 8-hour mean.
School Loss Days, All Cause	Gilliland et al.	2001	Southern California	5-17		D8HourMean	0.008150	0.004630	Log-linear	
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	PM2.5	D1HourMax	0.002200	0.000658	Log-linear	
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	PM2.5	D8HourMax	0.002596	0.000776	Log-linear	8-hour max from 1-hour max.

G.4.1   Chen et al (2000)

Chen et al. (2000) studied the association between air pollution and elementary school absenteeism (grades 1-6) in Washoe County, Nevada. Assuming that most children start kindergarten at age 5, the corresponding ages for grades 1 through 6 would be 6 through 11. Daily absence data were available for all elementary schools in the Washoe Country School District. The authors regressed daily total absence rate on the three air pollutants, meteorological variables, and indicators for day of the week, month, and holidays. They reported statistically significant associations between both ozone and CO and daily total absence rate for grades one through six. PM_{10} was negatively associated with absence rate, after adjustment for ozone, CO, and meteorological and temporal variables. The C-R function for ozone is based on the results from a multiple linear regression model with CO, ozone, and PM_{10}.

School Loss Days, All Cause

The coefficient and standard error are presented in Table 3 (Chen et al., 2000, p. 1008) for a unit ppm increase in the two-week average of daily one-hour maximum ozone

concentration. This is converted to unit ppb increase by dividing by 1,000.

The reported coefficient represents an *absolute* increase in absenteeism rate for a unit increase in ozone. If we apply this study to other locations, we assume that the same absolute increase will occur for a unit increase in ozone, regardless of the baseline rate. If the study location has a particularly high baseline rate, we may be overestimating decreases in absenteeism nationally, and vice-versa. As an example, consider if the baseline absenteeism rate were 10% in the study and 5% nationally. An absolute increase in absence rate of 2% associated with a given increase in ozone reflects a relative increase in absence rate of 20% for the study population. However, in the national estimate, we would assume the same absolute increase of 2%, but this would reflect a relative increase in the absenteeism rate of 40%.

An alternative approach is to estimate apply the *relative* increase in absenteeism rate in the C-R function by adjusting the results by the ratio of the national absenteeism rate to the study-specific rate. As a result, the percent increase in absenteeism rate associated with an increase in ozone is extrapolated nationally rather than the absolute increase in absenteeism rate. The incidence derivation section above describes the data used to estimate national and study-specific absence rates.

In addition to this scaling factor, there are two other scaling factors which are applied to the function. A scaling factor of 0.01 is used to convert the beta from a percentage (x 100) per unit increase of ozone to a proportion per unit increase of ozone. As a result it can be applied directly to the national population of school children ages 6 through 11 to estimate the number of absences avoided.

The final scaling factor adjusts for the number of school days in the ozone season. In the modeling program, the function is applied to every day in the ozone season (May 1 - September 30), however, in reality, school absences will be avoided only on school days. We assume that children are in school during weekdays for all of May, two weeks in June, one week in August, and all of September. This corresponds to approximately 2.75 months out of the 5 month season, resulting in an estimate of 39.3% of days (2.75/5*5/7). The C-R function parameters are shown below.

Population: population of children ages 6-11

Scaling Factor 1: Ratio of national school absence rate to study-specific school absence rate = 1.081. (National school absence rate of 5.5% obtained from the U.S. Department of Education (1996, Table 42-1). Study-specific school absence rate of 5.09% obtained from Chen et al. (2000, Table 1).)

Scaling Factor 2: Convert beta in percentage terms to a proportion = 0.01

Scaling Factor 3: Proportion of days that are school days in the ozone season = 0.393. (Ozone is modeled for the 5 months from May 1 through September 30. We assume that children are in school during weekdays for all of May, 2 weeks in June, 1 week in August, and all of September. This corresponds to approximately 2.75 months out of the 5 month season, resulting in an estimate of 39.3% of days (2.75/5*5/7).)

G.4.2   Crocker and Horst (1981)

To monetize benefits associated with increased worker productivity resulting from improved ozone air quality, we used information reported in Crocker and Horst (1981) and summarized in EPA (1994). Crocker and Horst examined the impacts of ozone exposure on the productivity of outdoor citrus workers. The study measured productivity impacts as the change in income associated with a change in ozone exposure, given as the elasticity of income with respect to ozone concentration (-0.1427). The relationship estimated by Crocker and Horst between wages and ozone is a log-log relationship. Therefore the elasticity of wages with respect to ozone is a constant, equal to the coefficient of the log of ozone in the model. The reported elasticity translates a ten percent reduction in ozone to a 1.4 percent increase in income. Given the national median daily income for outdoor workers engaged in strenuous activity reported by the U.S. Census Bureau (2002), $68 per day (2000$), a ten percent reduction in ozone yields about $0.97 in increased daily wages. (The national median daily income for workers engaged in "farming, forestry, and fishing" from the U.S. Census Bureau (2002, Table 621, p. 403) is used as a surrogate for outdoor workers engaged in strenuous activity.) We adjust the national median daily income estimate to reflect regional variations in income using a factor based on the ratio of county median household income to national median household income. No information was available for quantifying the uncertainty associated with the central valuation estimate. Therefore, no uncertainty analysis was conducted for this endpoint.

Worker Productivity

The C-R function for estimating changes in worker productivity is shown below:

$$\Delta productivity = \beta \frac{Q_1 - Q_0}{Q_1} \cdot dailyincome \cdot pop,$$

Daily Income: median daily income for outdoor workers. (The national median daily income for workers engaged in "farming, forestry, and fishing" was obtained from the U.S. Census Bureau (2002, Table 621, p. 403) and is used as a surrogate for outdoor workers engaged in strenuous activity. This national median daily income ($68) is then scaled by the ratio of national median income to county median income to estimate county median daily income for outdoor workers.)

Population: population of adults 18 to 64 employed as farm workers.

G.4.3   Gilliland et al (2001)

Gilliland et al. (2001) examined the association between air pollution and school absenteeism among 4th grade school children (ages 9-10) in 12 southern Californian communities. The study was conducted from January through June 1996. The authors used school records to collect daily absence data and parental telephone interviews to identify causes. They defined illness-related absences as respiratory or non-respiratory. A

respiratory illness was defined as an illness that included at least one of the following: runny nose/sneezing, sore throat, cough, earache, wheezing, or asthma attack. The authors used 15 and 30 day distributed lag models to quantify the association between ozone, PM_{10}, and NO_2 and incident school absences. Ozone levels were positively associated with all school absence measures and significantly associated with all illness-related school absences (non-respiratory illness, respiratory illness, URI and LRI). Neither PM_{10} nor NO_2 was significantly associated with illness-related school absences, but PM_{10} was associated with non-illness related absences. The health impact function for ozone is based on the results of the single pollutant model.

School Loss Days

Gilliland et al. (2001) defines an incident absence as an absence that followed attendance on the previous day and the incidence rate as the number of incident absences on a given day over the population at risk for an absence on a given day (i.e. those children who were not absent on the previous day). Since school absences due to air pollution may last longer than one day, an estimate of the average duration of school absences could be used to calculated the total avoided school loss days from an estimate of avoided new absences. A simple ratio of the total absence rate divided by the new absence rate would provide an estimate of the average duration of school absences, which could be applied to the estimate of avoided new absences as follows:

$$Duration = \frac{totalAbsences}{newAbsences}$$

$$\Delta TotalAbsences = -\left[incidence(e^{-\beta \Delta O_3} - 1) \right] \cdot duration \cdot pop$$

Since the function is log-linear, the baseline incidence rate (in this case, the rate of new absences) is multiplied by duration, which reduces to the total school absence rate. Therefore, the same result would be obtained by using a single estimate of the total school absence rate in the C-R function. Using this approach, we assume that the same relationship observed between pollutant and new school absences in the study would be observed for total absences on a given day. As a result, the total school absence rate is used in the function below. The derivation of this rate is described in the section on baseline incidence rate estimation.

For all absences, the coefficient and standard error are based on a percent increase of 16.3 percent (95% CI -2.6 percent, 38.9 percent) associated with a 20 ppb increase in 8-hour average ozone concentration (2001, Table 6, p. 52).

A scaling factor is used to adjust for the number of school days in the ozone season. In the modeling program, the function is applied to every day in the ozone season (May 1 -

September 30), however, in reality, school absences will be avoided only on school days. We assume that children are in school during weekdays for all of May, two weeks in June, one week in August, and all of September. This corresponds to approximately 2.75 months out of the 5 month season, resulting in an estimate of 39.3% of days (2.75/5*5/7).

In addition, not all children are at-risk for a new school absence, as defined by the study. On average, 5.5% of school children are absent from school on a given day (U.S. Department of Education, 1996, Table 42-1). Only those who are in school on the previous day are at risk for a new absence (1-0.055 = 94.5%). As a result, a factor of 94.5% is used in the function to estimate the population of school children at-risk for a new absence.

Incidence Rate: daily school absence rate = 0.055 (U.S. Department of Education, 1996, Table 42-1)

Population: population of children ages 9-10 not absent from school on a given day = 94.5% of children ages 9-10 (The proportion of children not absent from school on a given day (5.5%) is based on 1996 data from the U.S. Department of Education (1996, Table 42-1).)

Scaling Factor: Proportion of days that are school days in the ozone season = 0.393

(Ozone is modeled for the 5 months from May 1 through September 30. We assume that children are in school during weekdays for all of May, 2 weeks in June, 1 week in August, and all of September. This corresponds to approximately 2.75 months out of the 5 month season, resulting in an estimate of 39.3% of days (2.75/5*5/7).)

G.4.4   Ostro and Rothschild (1989)

Ostro and Rothschild (1989) estimated the impact of $PM_{2.5}$ and ozone on the incidence of minor restricted activity days (MRADs) and respiratory-related restricted activity days (RRADs) in a national sample of the adult working population, ages 18 to 65, living in metropolitan areas. The study population is based on the Health Interview Survey (HIS), conducted by the National Center for Health Statistics. In publications from this ongoing survey, non-elderly adult populations are generally reported as ages 18-64. From the study, it is not clear if the age range stops at 65 or includes 65 year olds. We apply the C-R function to individuals ages 18-64 for consistency with other studies estimating impacts to non-elderly adult populations. The annual national survey results used in this analysis were conducted in 1976-1981. Controlling for $PM_{2.5}$, two-week average ozone had a highly variable association with RRADs and MRADs. Controlling for ozone, two-week average $PM_{2.5}$ was significantly linked to both health endpoints in most years. The C-R function for ozone is based on the co-pollutant model with $PM_{2.5}$.

The study is based on a "convenience" sample of non-elderly individuals. Applying the C-R function to this age group is likely a slight underestimate, as it seems likely that elderly are at least as susceptible to ozone as individuals under 65. A number of studies have found that hospital admissions for the elderly are related to ozone exposures (e.g., Schwartz, 1994b; Schwartz, 1995).

Minor Restricted Activity Days

The coefficient and standard error used in the C-R function are based on a weighted average of the coefficients in Ostro and Rothschild (1989, Table 4). The derivation of these estimates is described below.

Incidence Rate: daily incidence rate for minor restricted activity days (MRAD) = 0.02137 (Ostro and Rothschild, 1989, p. 243)

Population: adult population ages 18 to 64

The coefficient used in the C-R function is a weighted average of the coefficients in Ostro and Rothschild (1989, Table 4) using the inverse of the variance as the weight. The calculation of the MRAD coefficient and its standard error is exactly analogous to the calculation done for the work-loss days coefficient based on Ostro (1987).

$$\beta = \left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \frac{1}{\sigma_{\beta_i}^2}} \right) = 0.00220.$$

The standard error of the coefficient is calculated as follows, assuming that the estimated year-specific coefficients are independent:

$$\sigma_\beta^2 = \text{var}\left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \frac{1}{\sigma_{\beta_i}^2}} \right) = \left(\frac{\sum\limits_{i=1976}^{1981} \frac{\beta_i}{\sigma_{\beta_i}^2}}{\gamma} \right) = \sum\limits_{i=1976}^{1981} \text{var}\left(\frac{\beta_i}{\sigma_{\beta_i}^2 \cdot \gamma} \right).$$

This reduces down to:

$$\sigma_\beta^2 = \frac{1}{\gamma} \Rightarrow \sigma_\beta = \sqrt{\frac{1}{\gamma}} = 0.000658.$$

G.5   Converting Functions to 8-Hour Daily Maximum Metric

A number of health impact functions were converted from 1-hour maximum, 24-hour average, and 8-hour average to the 8-hour maximum metric. To convert, say, a 1-hour maximum function, we multiplied the 1-hour maximum coefficient with the ratio of the

typical 1-hour maximum value to the typical 8-hour maximum value. We calculated ozone metric ratios for each quarter and year in the period 2000-2007. We calculated ratios by monitor, and by county, core business statistical area (CBSA), state, and nation.

For each monitor, a day was considered valid if it had at least 18 hourly values out of 24. A quarter was considered valid if it had at least 85 percent valid days. Ratios are calculated for the year, only if that year had four quarterly values. The CBSA codes, which were defined by OMB on 6-6-03, were obtained from: http://www.census.gov/population/estimates/metro-city/03msa.txt.

We chose the time period for the ratio calculation (e.g., spring and summer quarters) and the locations based on the data used in each epidemiological study. Table G-5 presents the 8-hour adjustment used for each study. Tables G-6 through G-8 present supporting documentation for some of the multi-city 8-hour adjustments.

Table G-5. Eight-Hour Adjustments by Study

Effect	Author	Year	Lcoation	Adjustment Factor Location	Quarters	Metric	8-Hour Adj	Notes
Mortality, Non-Accidental	Bell et al.	2004	95 US cities	Nation	2-3	24HourMean	0.67	
Mortality, All Cause	Bell et al.	2005	Meta-analysis	From study. See comment.	--	24HourMean	0.53	
HA, All Respiratory	Burnett et al.	2001	Toronto, CAN	Buffalo-Cheektowaga-Tonawanda, NY MSA	2-3	1HourMax	1.12	
School Loss Days, All Cause	Chen et al.	2000	Washoe Co, NV	Washoe County	1-4	1HourMax	1.19	
Worker Productivity	Crocker and Horst	1981	Florida	FL	1-4	24HourMean	0.65	
School Loss Days, All Cause	Gilliland et al.	2001	Southern California	Los Angeles-Long Beach-Santa Ana, CA MSA	1-4	8HourMean	0.96	The statewide avg is 0.96.
Mortality, Cardiopulmonary	Huang et al.	2005	19 US cities	See below	See below	24HourMean	0.66	
Mortality, Non-Accidental	Ito et al.	2005	6 US cities	See below	See below	24HourMean	0.65	
Mortality, Non-Accidental	Ito et al.	2005	Meta-analysis	From study. See comment.	--	24HourMean	0.67	
Mortality, Non-Accidental	Ito et al.	2006	Meta-analysis	From study. See comment.	--	1HourMax	1.33	
Mortality, All Cause	Levy et al.	2005	Meta-analysis	From study. See comment.	--	1HourMax	1.33	
HA, Chronic Lung Disease	Moolgavkar et al.	1997	Minneapolis, MN	Minneapolis-St. Paul-Bloomington, MN-WI MSA	1-4	24HourMean	0.70	Data 2004-2007 only.
HA, Pneumonia	Moolgavkar et al.	1997	Minneapolis, MN	Minneapolis-St. Paul-Bloomington, MN-WI MSA	1-4	24HourMean	0.70	Data 2004-2007 only.
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	Nation	1-4	1HourMax	1.18	
HA, Chronic Lung Disease (less Asthma)	Schwartz	1994	Detroit, MI	Detroit-Warren-Livonia, MI MSA	1-4	24HourMean	0.62	Data 2006 only.

HA, Pneumonia	Schwartz	1994	Detroit, MI	Detroit-Warren-Livonia, MI MSA	1-4	24HourMean	0.62	Data 2006 only.
HA, Pneumonia	Schwartz	1994	Minneapolis, MN	Minneapolis-St. Paul-Bloomington, MN-WI MSA	1-4	24HourMean	0.70	Data 2004-2007 only.
HA, All Respiratory	Schwartz	1995	New Haven, CT	New Haven-Milford, CT MSA	2-3	24HourMean	0.67	
HA, All Respiratory	Schwartz	1995	Tacoma, WA	Seattle-Tacoma-Bellevue, WA MSA	2-3	24HourMean	0.69	
Mortality, Non-Accidental	Schwartz	2004	14 US cities	See below	See below	1HourMax	0.67	

Table G-6. Eight-Hour Adjustment Details -- 6-City Study

City/County	CBSAs or Counties Used in Ratio Average	Quarters Used	Study Metric	8-Hour Adj
Detroit	Detroit-Warren-Livonia, MI MSA	2-3	24HourMean	0.63
Cook County	Cook County	2-3	24HourMean	0.65
Houston	Houston-Baytown-Sugar Land, TX MSA	2-3	24HourMean	0.59
Minneapolis	Minneapolis-St. Paul-Bloomington, MN-WI MSA	2-3	24HourMean	0.70
Philadelphia	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD MSA	2-3	24HourMean	0.65
St. Louis	St. Louis, MO-IL MSA	2-3	24HourMean	0.64
	Average			0.64

Table G-7. Eight-Hour Adjustment Details -- 14-City Study

City/County	CBSAs or Counties Used in Ratio Average	Quarters Used	Study Metric	8-Hour Adj	Notes
Birmingham	Birmingham-Hoover, AL MSA	2-3	1HourMax	1.15	
Boulder	Boulder, CO MSA	2-3	1HourMax	1.15	
Canton	Canton-Massillon, OH MSA	2-3	1HourMax	1.12	
Chicago	Chicago-Naperville-Joliet, IL-IN-WI MSA	2-3	1HourMax	1.16	
Cincinnati	Cincinnati-Middletown, OH-KY-IN MSA	2-3	1HourMax	1.15	
Colorado Springs	Colorado Springs, CO MSA	2-3	1HourMax	1.11	
Columbus	Columbus, OH MSA	2-3	1HourMax	1.13	
Detroit	Detroit-Warren-Livonia, MI MSA	2-3	1HourMax	1.17	
Houston	Houston-Baytown-Sugar Land, TX MSA	2-3	1HourMax	1.22	
New Haven	New Haven-Milford, CT MSA	2-3	1HourMax	1.17	
Pittsburgh	Pittsburgh, PA MSA	2-3	1HourMax	1.15	
Provo	Provo-Orem, UT MSA	3	1HourMax	1.13	Only quarter 3 available

Seattle	Seattle-Tacoma-Bellevue, WA MSA	2-3	1HourMax	1.15	
Spokane	Spokane, WA MSA	3	1HourMax	1.08	Only quarter 3 available
	Average			1.15	

Table G-8. Eight-Hour Adjustment Details -- 19-City Study

City/County	CBSAs or Counties Used in Ratio Average	Quarters Used	Study Metric	8-Hour Adj	Notes
Atlanta	Atlanta-Sandy Springs-Marietta, GA MSA	2-3	24HourMean	0.59	
Chicago	Chicago-Naperville-Joliet, IL-IN-WI MSA	2-3	24HourMean	0.65	
Cleveland	Cleveland-Elyria-Mentor, OH MSA	2-3	24HourMean	0.69	
Dallas/Ft Worth	Dallas-Fort Worth-Arlington, TX MSA	2-3	24HourMean	0.67	
Detroit	Detroit-Warren-Livonia, MI MSA	2-3	24HourMean	0.63	
Houston	Houston-Baytown-Sugar Land, TX MSA	2-3	24HourMean	0.59	
Los Angeles	Los Angeles-Long Beach-Santa Ana, CA MSA	2-3	24HourMean	0.59	Los Angeles and Santa Ana/Anaheim have same CBSA.
Santa Ana/Anaheim	Los Angeles-Long Beach-Santa Ana, CA MSA	2-3	24HourMean	0.59	
Miami	Miami-Fort Lauderdale-Miami Beach, FL MSA	2-3	24HourMean	0.71	
New York	New York-Newark-Edison, NY-NJ-PA MSA	2-3	24HourMean	0.66	
Philadelphia	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD MSA	2-3	24HourMean	0.65	
Phoenix	Phoenix-Mesa-Scottsdale, AZ MSA	2-3	24HourMean	0.66	
Pittsburgh	Pittsburgh, PA MSA	2-3	24HourMean	0.61	
San Bernardino	Riverside-San Bernardino-Ontario, CA MSA	2-3	24HourMean	0.65	
San Antonio	San Antonio, TX MSA	2-3	24HourMean	0.66	
San Diego	San Diego-Carlsbad-San Marcos, CA MSA	2-3	24HourMean	0.70	
Oakland	San Francisco-Oakland-Fremont, CA MSA	2-3	24HourMean	0.70	
San Jose	San Jose-Sunnyvale-Santa Clara, CA MSA	2-3	24HourMean	0.64	
Seattle	Seattle-Tacoma-Bellevue, WA MSA	2-3	24HourMean	0.69	
	Average			0.65	Keeping 1 Los Angeles keeps 8-hour adj at 0.65.

Appendix H: Health Valuation Functions in U.S. Setup

This appendix presents the unit values that are available in BenMAP for each of the health endpoints included in the current suite of health impact functions. Wherever possible, we present a distribution of the unit value, characterizing the uncertainty surrounding any point estimate. The mean of the distribution is taken as the point estimate of the unit value, and the distribution itself is used to characterize the uncertainty surrounding the unit value, which feeds into the uncertainty surrounding the monetary benefits associated with reducing the incidence of the health endpoint. Below we give detailed descriptions of the derivations of unit values and their distributions, as well as tables listing the unit values and their distributions, available for each health endpoint. The definitions of the distributions and their parameters is given in Table H-1.

Table H-1. Unit Value Uncertainty Distributions and Their Parameters

Distribution *	Parameter 1 (P1)	Parameter 2 (P2)
Normal	standard deviation	–
Triangular	minimum value	maximum value
Lognormal **	mean of corresponding normal distribution	standard deviation of corresponding normal distribution
Uniform	minimum value	maximum vaue
Weibull ***	α	β

* In all cases, BenMAP calculates the mean of the distribution, which is used as the "point estimate" of the unit value.

** If Y is a normal random variable, and Y = logeX, then X is lognormally distributed. Equivalently, X is lognormally distributed if X = eY, where Y is normally distributed.

*** The Weibull distribution has the following probability density function: $\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^{\beta}}$

H.1 Mortality

The economics literature concerning the appropriate method for valuing reductions in premature mortality risk is still developing. The adoption of a value for the projected reduction in the risk of premature mortality is the subject of continuing discussion within the economics and public policy analysis communities. Issues such as the appropriate discount rate and whether there are factors, such as age or the quality of life, that should be taken into consideration when estimating the value of avoided premature mortality are still

under discussion. BenMAP currently offers a variety of options reflecting the uncertainty surrounding the unit value for premature mortality.

H.1.1   Value of a Statistical Life Based on 26 Studies

One unit value available in BenMAP is $6.3 million. This estimate is the mean of a distribution fitted to 26 "value of statistical life" (VSL) estimates that appear in the economics literature and that have been identified in the Section 812 Reports to Congress as "applicable to policy analysis." This represents an intermediate value from a variety of estimates, and it is a value EPA has frequently used in Regulatory Impact Analyses (RIAs) as well as in the Section 812 Retrospective and Prospective Analyses of the Clean Air Act.

The VSL approach and the set of selected studies mirrors that of Viscusi (1992) (with the addition of two studies), and uses the same criteria as Viscusi in his review of value-of-life studies. The $6.3 million estimate is consistent with Viscusi's conclusion (updated to 2000$) that "most of the reasonable estimates of the value of life are clustered in the $3.8 to $8.9 million range." Five of the 26 studies are contingent valuation (CV) studies, which directly solicit WTP information from subjects; the rest are wage-risk studies, which base WTP estimates on estimates of the additional compensation demanded in the labor market for riskier jobs. Because this VSL-based unit value does not distinguish among people based on the age at their death or the quality of their lives, it can be applied to all premature deaths.

H.1.2   Value of a Statistical Life Based on Selected Studies

In addition to the value of a statistical life based on the results of 26 studies, we have included three alternatives based loosely on the results of recent work by Mrozek and Taylor (2002) and Viscusi and Aldy (2003). Each of the four alternatives has a mean value of $5.5 million (2000$), but with a different distributions: normal, uniform, triangular, and beta. Table H-2 presents the distribution parameters for the suite of mortality valuations currently available in BenMAP.

Table H-2. Unit Values Available for Mortality

Basis for Estimate *	Age Range at Death		Unit Value (VSL) (2000$)	Distribution of Unit Value	Parameters of Distribution	
	min.	*max.*			*P1*	*P2*
VSL, based on 26 value-of-life studies.	0	99	$6,324,101	Weibull	5.32E-6	1.509588
VSL based on range from $1 million to $10 million – 95% CI of assumed normal distribution.	0	99	$5,500,000	Normal	2,295,960.54	--
VSL based on range from $1 million to $10 million – assumed uniform distribution.	0	99	$5,500,000	Uniform	1,000,000	10,000,000

Basis for Estimate *	Age Range at Death		Unit Value (VSL) (2000$)	Distribution of Unit Value	Parameters of Distribution	
	min.	*max.*			*P1*	*P2*
VSL based on range from $1 million to $10 million – assumed triangular distribution.	0	99	$5,500,000	Triangular	1,000,000	10,000,000

*The original value of a statistical life was calculated in 1990 $. We have used a factor of 1.3175, based on the All-Items CPI-U.

H.2   Chronic Illness

This sub-section presents the unit values developed for chronic bronchitis, chronic asthma, and non-fatal myocardial infarctions.

H.2.1   Chronic Bronchitis

PM-related chronic bronchitis is expected to last from the initial onset of the illness throughout the rest of the individual's life. WTP to avoid chronic bronchitis would therefore be expected to incorporate the present discounted value of a potentially long stream of costs (e.g., medical expenditures and lost earnings) as well as WTP to avoid the pain and suffering associated with the illness. Both WTP and COI estimates are currently available in BenMAP.

H.2.1.1 Unit Value Based on Two Studies of WTP

Two contingent valuation studies, Viscusi et al. (1991) and Krupnick and Cropper (1992), provide estimates of WTP to avoid a case of chronic bronchitis. Viscusi et al. (1991) and Krupnick and Cropper (1992) were experimental studies intended to examine new methodologies for eliciting values for morbidity endpoints. Although these studies were not specifically designed for policy analysis, they can be used to provide reasonable estimates of WTP to avoid a case of chronic bronchitis. As with other contingent valuation studies, the reliability of the WTP estimates depends on the methods used to obtain the WTP values. The Viscusi et al. and the Krupnick and Cropper studies are broadly consistent with current contingent valuation practices, although specific attributes of the studies may not be.

The study by Viscusi et al. (1991) uses a sample that is larger and more representative of the general population than the study by Krupnick and Cropper (1992), which selects people who have a relative with the disease. However, the chronic bronchitis described to study subjects in the Viscusi study is severe, whereas a pollution-related case may be less severe.

The relationship between the severity of a case of chronic bronchitis and WTP to avoid it was estimated by Krupnick and Cropper (1992). We used that estimated relationship to derive a relationship between WTP to avoid a severe case of chronic bronchitis, as described in the Viscusis study, and WTP to avoid a less severe case. The estimated relationship (see Table 4 in Krupnick and Cropper) can be written as:

$$\ln(WTP) = \alpha + \beta^* sev$$

where a denotes all the other variables in the regression model and their coefficients, ß is the coefficient of sev, estimated to be 0.18, and sev denotes the severity level (a number from 1 to 13). Let x (< 13) denote the severity level of a pollution-related case of chronic bronchitis, and 13 denote the highest severity level (as described in Viscusi et al., 1991). Then

$$\ln(WTP_{13}) = \alpha + \beta^* 13$$

and

$$\ln(WTP_x) = \alpha + \beta^* x.$$

Subtracting one equation from the other,

$$\ln(WTP_{13}) - \ln(WTP_x) = \beta^*(13 - x)$$

or

$$\ln\left(\frac{WTP_{13}}{WTP_x}\right) = \beta^*(13 - x)$$

.

Exponentiating and rearranging terms,

$$WTP_x = WTP_{13} * e^{-\beta^*(13-x)}.$$

There is uncertainty surrounding the exact values of WTP13; x, and ß, and this uncertainty can be incorporated in the equation, if you request that the analysis be carried out in "uncertainty mode." The distribution of WTP to avoid a severe case of chronic bronchitis, WTP13 ,is based on the distribution of WTP responses in the Viscusi et al. (1991) study. The distribution of x, the severity level of an average case of pollution-related chronic bronchitis, is modeled as a triangular distribution centered at 6.5, with endpoints at 1.0 and 12.0. And the distribution of ß is normal with mean = 0.18 and std. dev.= 0.0669 (the estimate of b and standard error reported in Krupnick and Cropper, 1992).

In uncertainty mode, BenMAP uses a Monte Carlo approach. On each Monte Carlo iteration, random draws for these three variables are made, and the resulting WTPx is

calculated from the equation above. Because this function is non-linear, the expected value of WTP for a pollution-related case of CB cannot be obtained by using the expected values of the three uncertain inputs in the function (doing that will substantially understate mean WTP). A Monte Carlo analysis suggests, however, that the mean WTP to avoid a case of pollution-related chronic bronchitis is about $340,000. Therefore, if you request that the analysis be carried out in "point estimate" mode, that is the unit value that is used.

H.2.1.2 Alternative Cost of Illness Estimates

Cost of illness estimates for chronic bronchitis were derived from estimates of annual medical costs and annual lost earnings by Cropper and Krupnick (1990). This study estimated annual lost earnings resulting from chronic bronchitis as a function of age at onset of the illness, for the following age categories: 25-43, 35-44, 45-54, and 55-65 (see Cropper and Krupnick, Table 8). Annual medical expenses were estimated for 10-years age groups (0-9, 10-19, 20-29, ..., 80-89). We derived estimates of the present discounted value of the stream of medical and opportunity costs for people whose age of onset is 30, 40, 50, 60, 70, and 80. Medical costs (which are in 1977$ in the Cropper and Krupnick study) were inflated to 2000$ using the CPI-U for medical care; lost earnings (opportunity costs) were inflated to 2000$ using the Employment Cost Index for Wages and Salaries. Life expectancies were assumed to be unaffected by the illness. For example, an individual at age 70 has a life expectancy of 14.3 more years, and we assumed that someone whose age of onset of chronic bronchitis is 70 will also live for 14.3 more years. (Source of life expectancies: National Center for Health Statistics, 1999, Table 5.) We also assumed that opportunity costs at ages 66 and over were zero. Present discounted values were calculated using three and seven percent discount rates.

For each of the two discount rates, there are three cost of illness unit values for chronic bronchitis available in BenMAP, for the following age categories: 27-44, 45-64, and 65+. These are the age categories that were used in the epidemiological study that estimated a concentration-response function for chronic bronchitis (Abbey et al., 1995b). The estimate for the 27-44 age group is an average of the present discounted values calculated for ages 30 and 40; the estimate for the 45-64 age category is an average of the present discounted values calculated for ages 50 and 60; and the estimate for the 65+ age category is an average of the present discounted values calculated for ages 70 and 80. The suite of unit values available for use in BenMAP are shown in Table H-3 below.

Table H-3. Unit Values Available for Chronic Bronchitis

Basis for Estimate	Age of Onset		Present Discounted Value of Medical Costs	Present Discounted Value of Opportunity Costs	Unit Value	Distribution
	min.	max.				
WTP: average severity	30	99	N/A	N/A	$340,482	custom
COI: med costs + wage loss, 3% DR	27	44	$18,960	$135,463	$154,422	none

Basis for Estimate	Age of Onset		Present Discounted Value of Medical Costs	Present Discounted Value of Opportunity Costs	Unit Value	Distribution
	min.	max.				
	45	64	$23,759	$76,029	$99,788	none
	65	99	$11,088	$0	$11,088	none
COI: med costs + wage loss, 7% DR	27	44	$7,886	$80,444	$88,331	none
	45	64	$14,390	$59,577	$73,967	none
	65	99	$9,030	$0	$9,030	none

H.2.2 Chronic Bronchitis Reversals

The unit value for chronic bronchitis reversals assumes that this is chronic bronchitis with a severity level of 1. The method for generating a distribution of unit values in BenMAP is therefore the same as the WTP-based unit value method for chronic bronchitis (see above), with x=1. The mean of this distribution is $150,221.

H.2.3 Chronic Asthma

Two studies have estimated WTP to avoid chronic asthma in adults. Blumenschein and Johannesson (1998) used two different contingent valuation (CV) methods, the dichotomous choice method and a bidding game, to estimate mean willingness to pay for a cure for asthma. The mean WTP elicited from the bidding game was $189 per month, or $2,268 per year (in 1996$). The mean WTP elicited from the dichotomous choice approach was $343 per month, or $4,116 per year (in 1996$). Using $2,268 per year, a three percent discount rate, and 1997 life expectancies for males in the United States (National Center for Health Statistics, 1999, Table 5), the present discounted value of the stream of annual WTPs is $47,637 (in 2000$).

O'Conor and Blomquist (1997) estimated WTP to avoid chronic asthma from estimates of risk-risk tradeoffs. Combining the risk-risk tradeoffs with a statistical value of life, the annual value of avoiding asthma can be derived. Assuming a value of a statistical life of $6 million, they derived an annual WTP to avoid asthma of $1500 (O'Connor and Blomquist, 1997, p. 677). For a value of a statistical life of $5,894,400 (in 1997 $), the corresponding implied annual value of avoiding chronic asthma, based on O'Conor and Blomquist would be $1,474. Assuming a three percent discount rate and 1997 life expectancies for males in the United States, the present discounted value of the stream of annual WTPs would be $30,257 (in 2000$). A unit value, based on a three percent discount rate, is the average of the two estimates, or $38,947. Following the method used for the §812 Prospective analysis, the uncertainty surrounding the WTP to avoid a case of chronic asthma among adult males was characterized by a triangular distribution on the range determined by the two study-specific WTP estimates.

A second unit value, using a seven percent discount rate, is also available for use in BenMAP. The method used to derive this unit value is the same as that described above

for the three percent discount rate unit value. The unit values available for use in BenMAP are summarized in Table H-4 below.

Table H-4. Unit Values Available for Chronic Asthma

Basis for Estimate	Age Range		Unit Value	Distribution of Unit Value	Parameters of Distribution	
	min.	max.			P1	P2
WTP: 3% DR (Discount Rate)	27	99	$38,947	triangular	$30,257	$47,637
WTP: 7% DR	27	99	$25,357	triangular	$19,699	$31,015

H.2.4   Non-Fatal Myocardial Infarctions (Heart Attacks)

In the absence of a suitable WTP value for reductions in the risk of non-fatal heart attacks, there are a variety of cost-of-illness unit values available for use in BenMAP. These cost-of-illness unit values incorporate two components: the direct medical costs and the opportunity cost (lost earnings) associated with the illness event. Because the costs associated with a heart attack extend beyond the initial event itself, the unit values include costs incurred over five years. Using age-specific annual lost earnings estimated by Cropper and Krupnick (1990), and a three percent discount rate, we estimated the following present discounted values in lost earnings over 5 years due to a heart attack: $8,774 for someone between the ages of 25 and 44, $12,932 for someone between the ages of 45 and 54, and $74,746 for someone between the ages of 55 and 65. The corresponding age-specific estimates of lost earnings using a seven percent discount rate are $7,855, $11,578, and $66,920, respectively. Cropper and Krupnick do not provide lost earnings estimates for populations under 25 or over 65. As such we do not include lost earnings in the cost estimates for these age groups.

We have found three possible sources of estimates of the direct medical costs of a myocardial infarction (MI) in the literature:

Wittels et al. (1990) estimated expected total medical costs of MI over 5 years to be $51,211 (in 1986$) for people who were admitted to the hospital and survived hospitalization. (There does not appear to be any discounting used.) Wittels et al. was used to value coronary heart disease in the 812 Retrospective Analysis of the Clean Air Act. Using the CPI-U for medical care, the Wittels estimate is $109,474 in year 2000$. This estimated cost is based on a medical cost model, which incorporated therapeutic options, projected outcomes and prices (using "knowledgeable cardiologists" as consultants). The model used medical data and medical decision algorithms to estimate the probabilities of certain events and/or medical procedures being used. The authors note that the average length of hospitalization for acute MI has decreased over time (from an average of 12.9 days in 1980 to an average of 11 days in 1983). Wittels et al. used 10 days as the average in their study. It is unclear how much further the length of stay (LOS) for MI may have decreased from 1983 to the present. The average LOS for ICD code 410 (MI) in the year-2000 AHQR HCUP database is 5.5 days. However, this

may include patients who died in the hospital (not included among our non-fatal MI cases), whose LOS was therefore substantially shorter than it would be if they hadn't died.

Eisenstein et al. (2001) estimated 10-year costs of $44,663, in 1997$ (using a three percent discount rate), or $49,651 in 2000$ for MI patients, using statistical prediction (regression) models to estimate inpatient costs. Only inpatient costs (physician fees and hospital costs) were included.

Russell et al. (1998) estimated first-year direct medical costs of treating nonfatal MI of $15,540 (in 1995$), and $1,051 annually thereafter. Converting to year 2000$, that would be $18,880 for a 5-year period, using a three percent discount rate, or $17,850, using a seven percent discount rate.

The age group-specific estimates of opportunity cost over a five-year period are combined with the medical cost estimates from each of the three studies listed above. Because opportunity costs are derived for each of five age groups, there are 3 x 5 = 15 unit values for each of 2 discount rates, or 30 unit values available for use in BenMAP. These are given in Table H-5 below.

Note that we were unable to achieve complete consistency, unfortunately, because of limitations in the input studies. For example, although we calculated opportunity costs over a five-year period using a 3 percent and a 7 percent discount rate, we were not able to do the same for medical costs, except for the medical costs estimated by Russell et al. (in which they estimate an annual cost). Wittels et al. appear to have used no discounting in their estimate; Eisenstein et al. used a 3 percent discount rate. Similarly, although almost all cost estimates (opportunity costs and medical costs) are for a 5-year period, the medical cost estimate reported by Eisenstein et al. is for a 10-year period. There was no reasonable method for inferring from that study what costs over a 5-year period would be.

Table H-5. Unit Values Available for Myocardial Infarction

Basis of Estimate	Age Range		Medical Cost *	Opportunity Cost **	Total Cost
	Min	*Max*			
COI: 5 yrs med, 5 yrs wages, 3% DR, Wittels (1990)	0	24	$109,474	$0	$109,474
	25	44	$109,474	$9,033	$118,507
	45	54	$109,474	$13,313	$122,787
	55	65	$109,474	$76,951	$186,425
	66	99	$109,474	$0	$109,474
COI: 10 yrs med, 5 yrs wages, 3% DR, Eisenstein (2001)	0	24	$49,651	$0	$49,651
	25	44	$49,651	$9,033	$58,683
	45	54	$49,651	$13,313	$62,964
	55	65	$49,651	$76,951	$126,602
	66	99	$49,651	$0	$49,651

COI: 5 yrs med, 5 yrs wages, 3% DR, Russell (1998)	0	24	$22,331	$0	$22,331
	25	44	$22,331	$9,033	$31,363
	45	54	$22,331	$13,313	$35,644
	55	65	$22,331	$76,951	$99,281
	66	99	$22,331	$0	$22,331
COI: 5 yrs med, 5 yrs wages, 7% DR, Wittels (1990)	0	24	$109,474	$0	$109,474
	25	44	$109,474	$8,087	$117,561
	45	54	$109,474	$11,919	$121,393
	55	65	$109,474	$68,894	$178,368
	66	99	$109,474	$0	$109,474
COI: 10 yrs med, 5 yrs wages, 7% DR, Eisenstein (2001)	0	24	$49,651	$0	$49,651
	25	44	$49,651	$8,087	$57,738
	45	54	$49,651	$11,919	$61,570
	55	65	$49,651	$68,894	$118,545
	66	99	$49,651	$0	$49,651
COI: 5 yrs med, 5 yrs wages, 7% DR, Russell (1998)	0	24	$21,113	$0	$21,113
	25	44	$21,113	$8,087	$29,200
	45	54	$21,113	$11,919	$33,032
	55	65	$21,113	$68,894	$90,007
	66	99	$21,113	$0	$21,113

* From Cropper and Krupnick (1990). Present discounted value of 5 yrs of lost earnings, at 3% and 7% discount rate, adjusted from 1977$ to 2000$ using CPI-U "all items".

** An average of the 5-year costs estimated by Wittels et al. (1990) and Russell et al. (1998). Note that Wittels et al. appears not to have used discounting in deriving a 5-year cost of $109,474; Russell et al. estimated first-year direct medical costs and annual costs thereafter. The resulting 5-year cost is $22,331, using a 3% discount rate, and $21,113, using a 7% discount rate. Medical costs were inflated to 2000$ using CPI-U for medical care.

H.3   Hospital Admissions & Emergency Room Visits

This section presents the values for avoided hospital admissions, as well as avoided emergency room visits. We assume that hospital admissions due to acute exposure to air pollution pass through the emergency room. However, the value of hospital admissions that we have calculated here does not account for the cost incurred in the emergency room visit.

H.3.1   Hospital Admissions

As suggested above, the total value to society of an individual's avoidance of a hospital admission can be thought of as having two components: (1) the cost of illness (COI) to society, including the total medical costs plus the value of the lost productivity, as well as

(2) the WTP of the individual, as well as that of others, to avoid the pain and suffering resulting from the illness.

In the absence of estimates of social WTP to avoid hospital admissions for specific illnesses (components 1 plus 2 above), estimates of total COI (component 1) are available for use in BenMAP as conservative (lower bound) estimates. Because these estimates do not include the value of avoiding the pain and suffering resulting from the illness (component 2), they are biased downward. Some analyses adjust COI estimates upward by multiplying by an estimate of the ratio of WTP to COI, to better approximate total WTP. Other analyses have avoided making this adjustment because of the possibility of over-adjusting -- that is, possibly replacing a known downward bias with an upward bias. Based on Science Advisory Board (SAB) advice, the COI values currently available for use in BenMAP are not adjusted.

Unit values are based on ICD-code-specific estimated hospital charges and opportunity cost of time spent in the hospital (based on the average length of a hospital stay for the illness). The opportunity cost of a day spent in the hospital is estimated as the value of the lost daily wage, regardless of whether or not the individual is in the workforce.

For all hospital admissions endpoints available in BenMAP, estimates of hospital charges and lengths of hospital stays were based on discharge statistics provided by the Agency for Healthcare Research and Quality's Healthcare Utilization Project (2000). The total COI for an ICD-code-specific hospital stay lasting n days is estimated as the mean hospital charge plus n times the daily lost wage. Year 2000 county-specific median annual wages divided by (52*5) were used to estimate county-specific median daily wages. (The source for median is Geolytics, 2001.) Because wage data used in BenMAP are county-specific, the unit value for a hospital admission varies from one county to another.

Most hospital admissions categories considered in epidemiological studies consisted of sets of ICD codes. The unit value for the set of ICD codes was estimated as the weighted average of the ICD-code-specific COI estimates. The weights were the relative frequencies of the ICD codes among hospital discharges in the United States, as estimated by the National Hospital Discharge Survey (Owings and Lawrence, 1999, Table 1). The hospital admissions for which unit values are available in BenMAP are given in Table H-6. Although unit values available for use in BenMAP are county-specific, the national median daily wage was used to calculate opportunity costs and total costs for the table below, to give a general idea of the cost of illness estimates for the different hospital admissions endpoints.

The mean hospital charges and mean lengths of stay provided by (AHRQ 2000) are based on a very large nationally representative sample of about seven million hospital discharges, and are therefore the best estimates of mean hospital charges and mean lengths of stay available, with negligible standard errors.

Table H-6. Unit Values Available for Hospital Admissions

EndPoint	ICD Codes	Age Range		Mean Hospital Charge *	Mean Length of Stay (days) *	Total Cost of Illness (Unit Value) **
		min.	*max.*			
HA, All Cardiovascular	390-429	65	99	$20,607	5.07	$21,191
HA, All Cardiovascular	390-429	0	99	$20,873	4.71	$21,415
HA, All Cardiovascular	390-429	20	64	$22,300	4.15	$22,778
HA, Congestive Heart Failure	428	65	99	$14,573	5.60	$15,218
HA, Dysrhythmia	427	0	99	$14,811	3.70	$15,237
HA, Ischemic Heart Disease	410-414	65	99	$25,322	4.81	$25,876
HA, All Respiratory	460-519	65	99	$17,600	6.88	$18,393
HA, All Respiratory	460-519	0	99	$14,999	5.63	$15,647
HA, All Respiratory	460-519	0	2	$7,416	2.97	$7,759
HA, Asthma	493	0	64	$7,448	2.95	$7,788
HA, Asthma	493	65	99	$11,417	4.99	$11,991
HA, Asthma	493	0	99	$8,098	3.30	$8,478
HA, Chronic Lung Disease	490-496	65	99	$12,781	5.59	$13,425
HA, Chronic Lung Disease	490-496	0	99	$10,882	4.59	$11,412
HA, Chronic Lung Disease	490-496	20	64	$10,194	4.04	$10,660
HA, Chronic Lung Disease (less Asthma)	490-492, 494-496	65	99	$12,993	5.69	$13,648
HA, Chronic Lung Disease (less Asthma)	490-492, 494-496	0	99	$12,742	5.45	$13,370
HA, Chronic Lung Disease (less Asthma)	490-492, 494-496	20	64	$11,820	4.48	$11,820
HA, Pneumonia	480-487	65	99	$17,030	7.07	$17,844
HA, Pneumonia	480-487	0	99	$14,693	5.92	$15,375

* Source of hospital charges and lengths of stay: Agency for Healthcare Research and Quality. 2000. HCUPnet, Healthcare Cost and Utilization Project. http://www.agrq.gov/data/hcup/hcupnet.htm .

** The opportunity cost of a day spent in the hospital was estimated, for this exhibit, at the median daily wage of all workers, $115.20, regardless of age. The median daily wage was calculated by dividing the median weekly wage ($576 in 2000$) by 5. The median weekly wage was obtained from U.S. Census Bureau, Statistical Abstract of the United States: 2001, Section 12, Table 621: "Full-Time Wage and Salary Workers – Numbers and Earnings: 1985 to 2000." Actual unit values used in BenMAP are based on county-specific wages, and are therefore county-specific.

H.3.2 Emergency Room Visits for Asthma

Two unit values are currently available for use in BenMAP for asthma emergency room (ER) visits. One is $311.55, from Smith et al., 1997, who reported that there were approximately 1.2 million asthma-related ER visits made in 1987, at a total cost of $186.5 million, in 1987$. The average cost per visit was therefore $155 in 1987$, or $311.55 in 2000 $ (using the CPI-U for medical care to adjust to 2000$). The uncertainty surrounding this estimate, based on the uncertainty surrounding the number of ER visits and the total cost of all visits reported by Smith et al. is characterized by a triangular distribution centered at $311.55, on the interval [$230.67, $430.93].

A second unit value is $260.67 from Stanford et al. (1999). This study considered asthmatics in 1996-1997, in comparison to the Smith et al. (1997) study, which used 1987 National Medical Expenditure Survey (NMES) data). In comparing their study, the authors note that the 1987 NMES, used by Smith et al., "may not reflect changes in treatment patterns during the 1990s." In addition, its costs are the costs to the hospital (or ER) for treating asthma rather than charges or payments by the patient and/or third party payer. Costs to the ER are probably a better measure of the value of the medical resources used up on an asthma ER visit (see above for a discussion of costs versus charges).

The unit values and the corresponding distributions available in BenMAP for asthma-related ER visits are summarized in Table H-7.

Table H-7. Unit Values Available for Asthma-Related ER Visits

Basis for Estimate	Age Range		Unit Value	Distribution of Unit Value	Parameters of Distribution	
	min.	max.			P1	P2
COI: Smith et al. (1997)	0	99	$312	triangular	$231	$431
COI: Standford et al. (1999)	0	99	$261	normal	5.22	–

H.4 Acute Symptoms and Illness Not Requiring Hospitalization

Several acute symptoms and illnesses have been associated with air pollution, including acute bronchitis in children, upper and lower respiratory symptoms, and exacerbation of asthma (as indicated by one of several symptoms whose occurrence in an asthmatic generally suggests the onset of an asthma episode). In addition, several more general health endpoints which are associated with one or more of these acute symptoms and illnesses, such as minor restricted activity days, school loss days, and work loss days, have also been associated with air pollution. We briefly discuss the derivation of the unit values for each of these acute symptoms and illnesses. Tables H-8 and H-9 summarize the values.

Table H-8. Unit Values Available for Acute Symptoms and Illnesses

Health Endpoint	Basis for Estimate *	Age Range min.	Age Range max.	Unit Value	Distribution of Unit Value	Parameters of Distribution P1	Parameters of Distribution P2
Acute Bronchitis	WTP: 1 day illness, CV studies	0	17	$59	uniform	17.51	101.11
	WTP: 6 day illness, CV studies	0	17	$356	uniform	105.06	606.64
	WTP: 28 symptom-days, Dickie and Ulery	0	17	$374	lognormal	5.947	0.0907
Any of 19 Respiratory Symptoms	WTP: 1 day illness, CV studies	18	65	$24	uniform	0	48.25
Minor Restricted Activity Days	WTP: 1 day, CV studies	18	99	$51	triangular	20.71	80.37
	WTP: 3 symptoms 1 day, Dickie and Ulery (2002).	18	99	$98	lognormal	4.6088	0.0649
Lower Respiratory Symptoms	WTP: 1 day, CV studies	0	17	$16	uniform	6.94	24.47
	WTP: 2 symptoms 1 day, Dickie and Ulery (2002).	0	17	$187	lognormal	5.2556	0.07048
	WTP: 2 x 1 day, CV studies	0	17	$31	uniform	13.89	48.93
School Loss Days	Described in text.	0	17	$75	none	N/A	N/A
Upper Respiratory Symptoms	WTP: 1 day, CV studies	0	17	$25	uniform	9.22	43.11
	WTP: 2 symptoms 1 day, Dickie and Ulery (2002)	0	17	$187	lognormal	5.2556	0.07048
	WTP: 2 x 1 day, CV studies	0	17	$49	uniform	18.45	86.22
Work Loss Days **	Median daily wage, county-specific	18	65	$115	none	N/A	N/A

* All unit values pulled from a lognormal distribution from Model 1, Table III in Dickie and Ulery are multiplied by 0.973811 to adjust for a difference in mean household income between the study participants and the general population. The unit values shown here have already been adjusted.

** Unit values for work loss days are county-specific, based on county-specific median wages. The unit value shown here is the national median daily wage, given for illustrative purposes only.

Table H-9. Unit Values Available for Asthma-related Acute Symptoms and Illnesses

Health Endpoint	Basis for WTP Estimate*	Age Range		Unit Value	Unit Value Distribution	Parameters of Distribution	
		min.	max.			P1	P2
Asthma Attacks; Cough; Moderate or Worse; One or more symptoms; Shortness of Breath; Wheeze	Bad asthma day, Rowe and Chestnut (1986)	18	99	$43	uniform	15.56	70.88
	1 symptom-day, Dickie and Ulery (2002)	18	99	$74	lognormal	4.321	0.0957
	Bbad asthma day, Rowe and Chestnut (1986)	0	17	$43	uniform	15.56	70.88
	2 x bad asthma day, Rowe and Chestnut (1986)	0	17	$86	uniform	31.12	141.77
	1 symptom-day, Dickie and Ulery (2002)	0	17	$156	lognormal	5.074	0.0925

*All unit values pulled from a lognormal distribution from Model 1, Table III in Dickie and Ulery, 2002, are multiplied by 0.973811 to adjust for a difference in mean household income between the study participants and the general population. The unit values shown here have already been adjusted.

H.4.1   Acute Bronchitis in Children

Estimating WTP to avoid a case of acute bronchitis is difficult for several reasons. First, WTP to avoid acute bronchitis itself has not been estimated. Estimation of WTP to avoid this health endpoint therefore must be based on estimates of WTP to avoid symptoms that occur with this illness. Second, a case of acute bronchitis may last more than one day, whereas it is a day of avoided symptoms that is typically valued. Finally, the C-R function used in the benefit analysis for acute bronchitis was estimated for children, whereas WTP estimates for those symptoms associated with acute bronchitis were obtained from adults.

Three unit values are available in BenMAP for acute bronchitis in children. In previous benefit analyses, EPA used a unit value of $59.31. This is the midpoint between a low estimate and a high estimate. The low estimate is the sum of the midrange values recommended by IEc (1994) for two symptoms believed to be associated with acute bronchitis: coughing and chest tightness. The high estimate was taken to be twice the value of a minor respiratory restricted activity day. For a more complete description of the derivation of this estimate, see Abt Associates (2000, p. 4-30).

The above unit value assumes that an episode of acute bronchitis lasts only one day. However, this is generally not the case. More typically, it can last for 6 or 7 days. A simple adjustment, then, would be to multiply the original unit value of $59.31 by 6 or 7. A second unit value of $356 (=$59.31 x 6) was therefore derived.

Finally, as noted above, the epidemiological study relating air pollution to the incidence of acute bronchitis referred to children specifically. The value of an avoided case should

therefore be WTP to avoid a case in a child, which may be different from WTP to avoid a case in an adult. Recent work by Dickie and Ulery (2002) suggests, in fact, that parents are generally willing to pay about twice as much to avoid sickness in their children as in themselves. In one of several models they estimated, the natural logarithm of parents' WTP was related both to the number of symptom-days avoided and to whether it was their child or themselves at issue. Dickie and Ulery noted that "experiencing all of the symptoms [considered in their study – cough and phlegm, shortness of breath/wheezing, chest pain, and fever] for 7 days, or 28 symptom-days altogether, is roughly equivalent to a case of acute bronchitis ..." Using this model, and assuming that a case of acute bronchitis can be reasonably modeled as consisting of 28 symptom-days, we estimated parents' WTP to avoid a case of acute bronchitis in a child to be $374. This is the third unit value available in BenMAP.

The mean household income among participants in the Dickie and Ulery CV survey was slightly higher than the national average. We therefore adjusted all WTP estimates that resulted from their models downward slightly, using an income elasticity of WTP of 0.147, the average of the income elasticities estimated in the four models in the study. The adjustment factor thus derived was 0.9738.

H.4.2 Upper Respiratory Symptoms (URS) in Children

In past benefit analyses, EPA based willingness to pay to avoid a day of URS on symptom-specific WTPs to avoid those symptoms identified as part of the URS complex of symptoms. Pope et al. (1991) defined a day of URS as consisting of one or more of the following symptoms: runny or stuffy nose; wet cough; and burning, aching, or red eyes. The three contingent valuation (CV) studies shown in Table H-10 have estimated WTP to avoid various morbidity symptoms that are either within the URS symptom complex defined by Pope et al., or are similar to those symptoms.

Table H-10. Median WTP Estimates and Derived Midrange Estimates (in 1999 $)

Symptom *	Dickie et al.	Tolley et al. (1986)	Loehman et al. (1979)	Mid-Range Estimate
Throat congestion	4.81	20.84	-	12.75
Head/sinus congestion	5.61	22.45	10.45	12.75
Coughing	1.61	17.65	6.35	8.93
Eye irritation	-	20.03	-	20.03
Headache	1.61	32.07	-	12.75
Shortness of breath	0.00	-	13.47	6.37
Pain upon deep inhalation (PDI)	5.63	-	-	5.63
Wheeze	3.21	-	-	3.21
Coughing up phlegm	3.51 **	-	-	3.51

Symptom *	Dickie et al.	Tolley et al. (1986)	Loehman et al. (1979)	Mid-Range Estimate
Chest tightness	8.03	-	-	8.03

* All estimates are WTP to avoid one day of symptom. Midrange estimates were derived by IEc (1993).

** 10% trimmed mean.

The three individual symptoms that were identified as most closely matching those listed by Pope et al. for URS are cough, head/sinus congestion, and eye irritation, corresponding to "wet cough," "runny or stuffy nose," and "burning, aching or red eyes," respectively. A day of URS could consist of any one of the seven possible "symptom complexes" consisting of at least one of these three symptoms. The original unit value for URS was based on the assumption that each of these seven URS complexes is equally likely. This unit value for URS, $24.64, is just an average of the seven estimates of mean WTP for the different URS complexes.

The WTP estimates on which the first unit value is based were elicited from adults, whereas the health endpoint associated with air pollution in the epidemiological study is in children. As noted above, recent research by Dickie and Ulery (2002) suggests that parental WTP to avoid symptoms and illnesses in their children is about twice what it is to avoid those symptoms and illnesses in themselves. We therefore derived a second unit value of $49.28 (=2 x $24.64) from the first unit value.

A third unit value was derived by using Model 1, Table III in Dickie and Ulery (2002) (the same model used for acute bronchitis), assuming that a day of URS consists of 2 symptoms.   As noted above, this model relates parental WTP to the number of symptom-days avoided and to whether it is the parent or the child at issue. The unit value derived from this model is $187.

A WTP estimate elicited from parents concerning their WTP to avoid symptoms in their children may well include some calculation of lost earnings resulting from having to lose a day of work. Estimates from the Dickie and Ulery model therefore (appropriately) probably include not only their WTP to have their children avoid the pain and suffering associated with their illness, but also the opportunity cost of a parent having to stay home with a sick child.

H.4.3   Lower Respiratory Symptoms (LRS) in Children

The three unit values for LRS in children currently available in BenMAP follow the same pattern as those for URS in children. In past benefit analyses, EPA based willingness to pay to avoid a day of LRS on symptom-specific WTPs to avoid those symptoms identified as part of the LRS complex of symptoms. Schwartz et al. (1994) defined a day of LRS as consisting of at least two of the following symptoms: cough, chest tightness, coughing up phlegm, and wheeze. Of the symptoms for which WTP estimates are available (listed in Table H-10), those that most closely match the symptoms listed by Schwartz et al. are

coughing, chest tightness, coughing up phlegm, and wheeze. A day of LRS, as defined by Schwartz et al., could consist of any one of 11 possible combinations of at least two of these four symptoms. In the absence of any further information, each of the 11 possible "symptom clusters" was considered equally likely. The original unit value for LRS, $15.57, is just an average of the eleven estimates of mean WTP for the different LRS symptom clusters.

A second unit value is twice the original unit value, or $31.15, based on the evidence from Dickie and Ulery (2002) that parents are willing to pay about twice as much to avoid symptoms and illness in their children as in themselves. The third unit value is based on Model 1, Table III in Dickie and Ulery, assuming that, as for URS, a day of LRS consists of 2 symptoms. As noted above, this model relates parental WTP to the number of symptom-days avoided and to whether it is the parent or the child at issue. The unit value derived from this model is $187.

H.4.4   Any of 19 Respiratory Symptoms

The presence of "any of 19 acute respiratory symptoms" is a somewhat subjective health effect used by Krupnick et al. (1990). Moreover, not all 19 symptoms are listed in the Krupnick et al. study. It is therefore not clear exactly what symptoms were included in the study. Even if all 19 symptoms were known, it is unlikely that WTP estimates could be obtained for all of the symptoms. Finally, even if all 19 symptoms were known and WTP estimates could be obtained for all 19 symptoms, the assumption of additivity of WTPs becomes tenuous with such a large number of symptoms. The likelihood that all 19 symptoms would occur simultaneously, moreover, is very small.

Acute respiratory symptoms must be either upper respiratory symptoms or lower respiratory symptoms. In the absence of further knowledge about which of the two types of symptoms is more likely to occur among the "any of 19 acute respiratory symptoms," we assumed that they occur with equal probability. Because this health endpoint may also consist of combinations of symptoms, it was also assumed that there is some (smaller) probability that upper and lower respiratory symptoms occur together. To value avoidance of a day of "the presence of any of 19 acute respiratory symptoms" we therefore assumed that this health endpoint consists either of URS, or LRS, or both. We also assumed that it is as likely to be URS as LRS and that it is half as likely to be both together. That is, it was assumed that "the presence of any of 19 acute respiratory symptoms" is a day of URS with 40 percent probability, a day of LRS with 40 percent probability, and a day of both URS and LRS with 20 percent probability. Using the point estimates of WTP to avoid a day of URS and LRS derived above, the point estimate of WTP to avoid a day of "the presence of any of 19 acute respiratory symptoms" is:

$(0.40)(\$24.64) + (0.40)(\$15.57) + (0.20)(\$24.64 + \$15.57) = \$24.12.$

Because this health endpoint is only vaguely defined, and because of the lack of information on the relative frequencies of the different combinations of acute respiratory symptoms that might qualify as "any of 19 acute respiratory symptoms," the unit dollar value derived for this health endpoint must be considered only a rough approximation.

H.4.5 Work Loss Days (WLDs)

Work loss days are valued at a day's wage. BenMAP calculates county-specific median daily wages from county-specific annual wages by dividing by (52*5), on the theory that a worker's vacation days are valued at the same daily rate as work days.

H.4.6 Minor Restricted Activity Days (MRADs)

Two unit values are currently available in BenMAP for MRADs. No studies are reported to have estimated WTP to avoid a minor restricted activity day (MRAD). However, IEc (1993) derived an estimate of WTP to avoid a minor respiratory restricted activity day (MRRAD), using WTP estimates from Tolley et al. (1986) for avoiding a three-symptom combination of coughing, throat congestion, and sinusitis. This estimate of WTP to avoid a MRRAD, so defined, is $38.37 (1990 $). Although Ostro and Rothschild (1989) estimated the relationship between $PM_{2.5}$ and MRADs, rather than MRRADs (a component of MRADs), it is likely that most of the MRADs associated with exposure to $PM_{2.5}$ are in fact MRRADs. The original unit value, then, assumes that MRADs associated with PM exposure may be more specifically defined as MRRADs, and uses the estimate of mean WTP to avoid a MRRAD.

Any estimate of mean WTP to avoid a MRRAD (or any other type of restricted activity day other than WLD) will be somewhat arbitrary because the endpoint itself is not precisely defined. Many different combinations of symptoms could presumably result in some minor or less minor restriction in activity. Krupnick and Kopp (1988) argued that mild symptoms will not be sufficient to result in a MRRAD, so that WTP to avoid a MRRAD should exceed WTP to avoid any single mild symptom. A single severe symptom or a combination of symptoms could, however, be sufficient to restrict activity. Therefore WTP to avoid a MRRAD should, these authors argue, not necessarily exceed WTP to avoid a single severe symptom or a combination of symptoms. The "severity" of a symptom, however, is similarly not precisely defined; moreover, one level of severity of a symptom could induce restriction of activity for one individual while not doing so for another. The same is true for any particular combination of symptoms.

Given that there is inherently a substantial degree of arbitrariness in any point estimate of WTP to avoid a MRRAD (or other kinds of restricted activity days), the reasonable bounds on such an estimate must be considered. By definition, a MRRAD does not result in loss of work. WTP to avoid a MRRAD should therefore be less than WTP to avoid a WLD. At the other extreme, WTP to avoid a MRRAD should exceed WTP to avoid a single mild symptom. The highest IEc midrange estimate of WTP to avoid a single symptom is $20.03 (1999 $), for eye irritation. The point estimate of WTP to avoid a WLD in the benefit analysis is $83 (1990 $). If all the single symptoms evaluated by the studies are not severe, then the estimate of WTP to avoid a MRRAD should be somewhere between $16 and $83. Because the IEc estimate of $38 falls within this range (and acknowledging the degree of arbitrariness associated with any estimate within this range), the IEc estimate is used as the mean of a triangular distribution centered at $38, ranging from $16 to $61. Adjusting to

2000 $, this is a triangular distribution centered at $50.55, ranging from $21 to $80.

A second unit value is based on Model 1, Table III in Dickie and Ulery (2002). This model estimates the natural logarithm of parents' WTP to avoid symptoms as a linear function of the natural logarithm of the number of symptom-days avoided and whether or not the person avoiding the symptoms is the parent or the child. The unit value derived from this model, assuming that an MRAD consists of one day of 3 symptoms in an adult, is $98.

H.4.7   Asthma Exacerbation

Several respiratory symptoms in asthmatics or characterizations of an asthma episode have been associated with exposure to air pollutants. All of these can generally be taken as indications of an asthma exacerbation ("asthma attack") when they occur in an asthmatic. BenMAP therefore uses the same set of unit values for all of the variations of "asthma exacerbation" that appear in the epidemiological literature.

Two unit values are currently available in BenMAP for asthma exacerbation in adults, and three are currently available for asthma exacerbation in children. In past benefit analyses, EPA based willingness to pay to avoid an asthma exacerbation on four WTP estimates from Rowe and Chestnut (1986) for avoiding a "bad asthma day." The mean of the four average WTPs is $32 (1990 $), or $43 in 2000$. The uncertainty surrounding this estimate was characterized by a continuous uniform distribution on the range defined by the lowest and highest of the four average WTP estimates from Rowe and Chestnut, [$12, $54] in 1990$, or [$16, $71] in 2000 $. This unit value is available for both adults and children.

A second unit value for adults was derived by using Model 1, Table III in Dickie and Ulery (2002) -- the same model used for acute bronchitis, LRS, and URS -- assuming that an asthma exacerbation consists of 1 symptom-day. As noted above, this model relates parental WTP to the number of symptom-days avoided and to whether it is the parent or the child at issue. The unit value derived from this model for adults is $74.

Two additional unit values are available for children. One of these is twice the original unit value, or $86, based on the evidence from Dickie and Ulery (2002) that parents are willing to pay about twice as much to avoid symptoms and illness in their children as in themselves. The third unit value is based on Model 1, Table III in Dickie and Ulery (the same model used for asthma exacerbation in adults, only now with the "adult or child" variable set to 1 rather than 0). The unit value derived from this model is $156.

H.4.8   School Loss Days

There is currently one unit value available in BenMAP for school loss days, based on (1) the probability that, if a school child stays home from school, a parent will have to stay home from work to care for the child, and (2) the value of the parent's lost productivity. We first estimated the proportion of families with school-age children in which both parents work, and then valued a school loss day as the probability of a work loss day resulting from a school loss day (i.e., the proportion of households with school-age

children in which both parents work) times a measure of lost wages.

From the U.S. Bureau of the Census (2002) we obtained (1) the numbers of single, married, and "other" (i.e., widowed, divorced, or separated) women with children in the workforce, and (2) the rates of participation in the workforce of single, married, and "other" women with children. From these two sets of statistics, we calculated a weighted average participation rate of 72.85 percent, as shown in Table H-11.

Our estimated daily lost wage (if a mother must stay at home with a sick child) is based on the median weekly wage among women age 25 and older in 2000 . This median weekly wage is $551. Dividing by 5 gives an estimated median daily wage of $103. The expected loss in wages due to a day of school absence in which the mother would have to stay home with her child is estimated as the probability that the mother is in the workforce times the daily wage she would lose if she missed a day = 72.85% of $103, or $75. We currently have insufficient information to characterize the uncertainty surrounding this estimate.

Table H-11. Women with Children: Number and Percent in the Labor Force, 2000, and Weighted Average Participation Rate

Category	Women in Labor Force (millions) *	Participation Rate (%) *	Implied Total Number in Population (in millions)	Implied Percent in Population	Population-Weighted Average Participation Rate
	(1)	(2)	(3) = (1)/(2)	(4)	[=sum (2)*(4) over rows]
Single	3.1	73.9%	4.19	11.84%	--
Married	18.2	70.6%	25.78	72.79%	--
Other **	4.5	82.7%	5.44	15.36%	--
Total	--	--	35.42	--	72.85%

* Source: U.S. Bureau of the Census (2002).

** Widowed, divorced, or separated.

A unit value based on the approach described above is likely to understate the value of a school loss day in two ways. First, it omits WTP to avoid the symptoms/illness which resulted in the school absence. Second, it effectively gives zero value to school absences which do not result in a work loss day. The unit value of $75 is therefore considered an "interim" value until such time as alternative means of estimating this unit value become available.

Appendix I: Population & Other Data in U.S. Setup

This section describes the population and monitor data in the United States setup.

Population Data. This describes how BenMAP forecasts population; the block-level and county-level data underlying the forecasts; and the PopGrid software application, which aggregates block-level population data to whatever grid definition might be needed.

Monitor Data. The default United States setup has ozone, PM2.5, PM10 and lead monitor data for the years 2000-2007. Data for CO, NO2, and SO2 are available at the BenMAP website: http://www.epa.gov/air/benmap/.

I.1   Population Data in U.S. Setup

The U.S. setup in BenMAP calculates health impacts for any desired grid definition, so long as you have a shapefile for that grid definition and population data for that grid definition. In this description, we use the term "population grid cell" to refer to a cell (e.g., county) within a grid definition. The foundation for calculating the population level in the population grid-cells is 2000 Census block data. A separate application developed by Abt Associates, called "PopGrid," described below, combines the Census block data with any user-specified set of population grid-cells, so long as they are defined by a GIS shape file. Unfortunately, PopGrid relies on extremely large census files that are too large to include with BenMAP -- hence the need for the separate application. If you are interested in PopGrid, please email: benmap@epa.gov.

Within any given population grid-cell, BenMAP has 304 unique race-ethnicity-gender-age groups: 19 age groups by 2 ethnic groups by gender by 4 racial groups (19*2*2*4=304). Exhibit B-1 presents the 304 population variables available in BenMAP. As discussed below, these variables are available for use in developing age estimates in whatever grouping desired by you.

Exhibit I-1 Demographic Groups and Variables Available in BenMAP

Racial/Ethnic Group	Ethnicity	Age	Gender
White, African American, Asian, American Indian, Other, Hispanic	Hispanic, Non-Hispanic	<1, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+	Male, Female

In this section on population data in the U.S. setup, we describe:

Forecasting Population. This describes how BenMAP forecasts population.

Data Needed. This section describes the block-level and county-level data underlying the forecasts.

PopGrid. This section reviews the PopGrid software application, which aggregates block-level population data to whatever grid definition might be needed.

I.1.1 How BenMAP Forecasts Population

In calculating the population in age groups that may include a portion of one of the pre-specified demographic groups in Exhibit I-1, BenMAP assumes the population is uniformly distributed in the age group. For example, to calculate the number of children ages 3 through 12, BenMAP calculates:

$$age_{3-12} = \frac{1}{2} \, age_{1-4} + age_{5-9} + \frac{3}{5} \, age_{10-14} \quad .$$

To estimate population levels for the years after the last Census in 2000, BenMAP scales the 2000 Census-based estimate with the ratio of the county-level forecast for the future year of interest over the 2000 county-level population level. Woods & Poole (2007) provides the county-level population forecasts used to calculate the scaling ratios; these data are discussed in detail below.

In the simplest case, where one is forecasting a single population variable, say, children ages 4 to 9 in the year 2010, CAMPS calculates:

$$age_{4-9, \, g, \, 2010} = age_{4-9, \, g, \, 2000} \, \frac{age_{4-9, \, county \, 2010}}{age_{4-9, \, county \, 2000}}$$

where the gth population grid-cell is wholly located within a given county.

In the case, where the gth grid-cell includes "n" counties in its boundary, the situation is somewhat more complicated. BenMAP first estimates the fraction of individuals in a given age group (e.g., ages 4 to 9) that reside in the part of each county within the gth grid-cell. BenMAP calculates this fraction by simply dividing the population all ages of a given county within the gth grid-cell by the total population in the gth grid-cell:

$$fraction \; of \; age_{4-9, \, g \, in \; county_c} = \frac{age_{all, \, g \, in \; county_c}}{age_{all, \, g}}$$

Multiplying this fraction with the number of individuals ages 4 to 9 in the year 2000 gives an estimate of the number of individuals ages 4 to 9 that reside in the fraction of the county within the gth grid-cell in the year 2000:

$$age_{4-9,\,g\,in\,county_c,\,2000} = age_{4-9,\,g,\,2000}\ fraction\ age_{4-9,\,g\,in\,county_c}$$

To then forecast the population in 2010, we scale the 2000 estimate with the ratio of the county projection for 2010 to the county projection for 2000:

$$age_{4-9,\,g\,in\,county_c,\,2010} = age_{4-9,\,g\,in\,county_c,\,2000}\ \frac{age_{4-9,\,county_c,\,2010}}{age_{4-9,\,county_c,\,2000}}\ .$$

Combining all these steps for "n" counties within the gth grid-cell, we forecast the population of persons ages 4 to 9 in the year 2010 as follows:

$$age_{4-9,\,g,\,2010} = \sum_{c=1}^{n} age_{4-9,\,g,\,2000}\ \frac{total\ pop_{g\,in\,county_c}}{total\ pop_g}\ \frac{age_{4-9,\,county_c,\,2010}}{age_{4-9,\,county_c,\,2000}}$$

In the case where there are multiple age groups and multiple counties, BenMAP first calculates the forecasted population level for individual age groups, and then combines the forecasted age groups. In calculating the number of children ages 4 to 12, BenMAP calculates:

$$age_{4-9,\,g,\,2010} = \sum_{c=1}^{n} age_{4-9,\,g,\,2000}\ \frac{total\ pop_{g\,in\,county_c}}{total\ pop_g}\ \frac{age_{4-9,\,county_c,\,2010}}{age_{4-9,\,county_c,\,2000}}$$

$$age_{10-14,\,g,\,2010} = \sum_{c=1}^{n} age_{10-14,\,g,\,2000}\ \frac{total\ pop_{g\,in\,county_c}}{total\ pop_g}\ \frac{age_{10-14,\,county_c,\,2010}}{age_{10-14,\,county_c,\,2000}}$$

$$age_{4-12,\,g,\,2010} = age_{4-9,\,g,\,2010} + \frac{3}{5}\ age_{10-14,\,g,\,2010}\ .$$

Since the Woods and Poole (2007) projections only extend through 2030, we used the existing projections and constant growth factors to provide additional projections. To estimate population levels beyond 2030, CAPMS linearly extrapolates from the final two years of data. For example, to forecast population in 2035, CAPMS calculates:

$$age_{4-9,\,2035} = age_{4-9,\,2030} + 5\left(age_{4-9,\,2030} - age_{4-9,\,2029}\right) \ .$$

I.1.2 Data Needed for Forecasting

Underlying the population forecasts in BenMAP there are block-level databases used to provide year 2000 population estimates and a county-level database of forecast ratios. Both files have the same set of 304 race-ethnicity-gender-age population groups.

The block-level data is typically not used directly in BenMAP, and instead is used with the PopGrid software (described below) to provide year 2000 estimates for a grid definition of interest (e.g., 12 kilometer CMAQ grid). The output from PopGrid with the year 2000 population estimates can then be loaded into BenMAP.

The county-level data comes pre-installed in the U.S. setup, and is not something that the user needs to load herself. These data are simply county-level ratios of a "future" year (2000-2030) and year 2000 population data for each county and each of the 304 race-ethnicity-gender-age population groups.

We describe the development of each databases below.

I.1.2.1 Block-Level Census 2000

There are about five million "blocks" in the United States, and for each block we have 304 race-ethnicity-gender-age groups. The block-level population database is created separately for each state, in order to make the data more manageable. (A single national file of block data would be about six gigabytes.)

The initial block file from the U.S. Census Bureau is not in the form needed. The block data has 7 racial categories and 23 age groups, as opposed to the 4 and 19 used in BenMAP. Table I-4 summarizes the initial set of variables and the final desired set of variables.

Table I-4 Race, Ethnicity and Age Variables in 2000 Census Block Data

Type	Race	Ethnicity	Gender	Age
Initial Variables (SF1 file)	White Alone, Black Alone, Native American Alone, Asian Alone, Pacific Islander / Hawaiian Alone, Other Alone, Two or More Alone	--	Male, Female	0-4, 5-9, 10-14, 15-17, 18-19, 20, 21, 22-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-61, 62-64, 65-66, 67-69, 70-74, 75-79, 80-84 85+
Final Desired Variables	White, African-American, Asian-American, Native-American	Hispanic, Non-Hispanic	Female, Male	<1, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80-84, 85+

The initial set of input files are as follows.

Census 2000 block-level file (Summary File 1)
Data: ftp.census.gov/census_2000/datasets/Summary_File_1
Docs: http://www.census.gov/prod/cen2000/doc/sf1.pdf

Census 2000 tract-level file (Summary File 4)
Data: ftp.census.gov/census_2000/datasets/Summary_File_4
Docs: http://www.census.gov/prod/cen2000/doc/sf4.pdf

Census 2000 MARS national-level summary
Docs: http://www.census.gov/popest/archives/files/MRSF-01-US1.pdf

The SF4 and MARS data, as described below, are needed to reorganize the variables that come initially in the SF1 file. (For the sake of completeness, we note that there exists a county-level Census 2000 MARS file, however, due to major population count discrepancies between the county-level MARS file and block-level SF1 file, we used only the nation-level summary table. Tables in MARS documentation file did not have the discrepancies that the county-level file had. We were unable to get an adequate explanation of this from the US Census.)

The steps in preparing the data are as follows:

1. Adjust Age-classifications:

We combined some age groups in the SF1 data to match the age groups wanted for BenMAP. For example, we combined age groups 15-17 and 18-19 to create the 15-19 age group used in BenMAP. Then, in the case of the 0-4 age group, we split it into <1 and 1-4 using the county-level SF4 data, which gave us the fraction of 0-4 year-olds who are <1.

2. Fill in Missing Racial-Ethnic Interactions:

We used the county-level SF4 data to calculate the fraction of Hispanics in each ethnically-aggregated subpopulation from the block-level data, by age and sex. We used these fractions to distribute each age-sex-race-block-level datum into Hispanics and non-Hispanics.

This process tended to underestimate the number of Hispanics. Therefore, we made a correction. We calculated the ratio of total Hispanics in a state to the estimated Hispanics in a state. In each age-sex-race-block-level datum, if the resulting total of Hispanics is not greater than the total number of people in the datum, we increased the number of people according to this ratio.

3. Assign "Other" and "Multi-Racial" to the Remaining Four Racial Categories:

We assign the "Other" race category in two steps. First, based on the national MARS data, we estimated how many people in the "multi-racial" category checked off "some other race" as one of their races, for Hispanics and non-Hispanics separately. In each age-sex-race-block-level datum, we added those people to "other race" category to create the re-distribution pool, analogously to the method implemented by Census while creating MARS data (*see* U.S. Census Bureau, 2002a, Table 1, below). Second, based on the national re-allocation fractions for Hispanics and non-Hispanics (derived from the MARS data), we assigned the "Other" race into the four races of interest and "multi-race".

After the assignment of the "Other" race category, we then assigned "multi-racial" category to the four racial categories categories, using state fractions of these races in each age-sex-race-block-level datum.

Table 1. Summary of Modified Race and Census 2000 Race Distributions for the United States

Subject	Modified Race Number	Modified Race Percent	Census 2000 Number	Census 2000 Percent
TOTAL POPULATION	281,421,906	100.00	281,421,906	100.00
One race	277,524,226	98.62	274,595,678	97.57
Specified race only	277,524,226	98.62	259,236,605	92.12
White	228,104,485	81.05	211,460,626	75.14
Black or African American	35,704,124	12.69	34,658,190	12.32
American Indian and Alaska Native	2,663,818	0.95	2,475,956	0.88
Asian	10,589,265	3.76	10,242,998	3.64
Native Hawaiian and Other Pacific Islander	462,534	0.16	398,835	0.14
Non-specified race only	(X)	(X)	15,359,073	5.46
Two races	3,578,053	1.27	6,368,075	2.26
Specified race only	3,578,053	1.27	3,366,517	1.20
Specified and non-specified races	(X)	(X)	3,001,558	1.07
Three or more races	319,627	0.11	458,153	0.16
Specified race only	319,627	0.11	297,298	0.11
Specified and non-specified races	(X)	(X)	160,855	0.06
HISPANIC OR LATINO AND RACE	35,305,818	100.00	35,305,818	100.00
One race	34,814,386	98.61	33,081,736	93.70
Specified race only	34,814,386	98.61	18,190,433	51.52
White	32,529,000	92.13	16,907,852	47.89
Black or African American	1,391,117	3.94	710,353	2.01
American Indian and Alaska Native	566,378	1.60	407,073	1.15
Asian	232,461	0.66	119,829	0.34
Native Hawaiian and Other Pacific Islander	95,430	0.27	45,326	0.13
Non-specified race only	(X)	(X)	14,891,303	42.18
Two races	433,726	1.23	2,110,965	5.98
Specified race only	433,726	1.23	315,611	0.89
Specified and non-specified races			1,795,354	5.09
Three or more races	57,706	0.16	113,117	0.32
Specified race only	57,706	0.16	48,933	0.14
Specified and non-specified races	(X)	(X)	64,184	0.18
NOT HISPANIC OR LATINO AND RACE	246,116,088	100.00	246,116,088	100.00
One race	242,709,840	98.62	241,513,942	98.13
Specified race only	242,709,840	98.62	241,046,172	97.94
White	195,575,485	79.46	194,552,774	79.05
Black or African American	34,313,007	13.94	33,947,837	13.79
American Indian and Alaska Native	2,097,440	0.85	2,068,883	0.84
Asian	10,356,804	4.21	10,123,169	4.11
Native Hawaiian and Other Pacific Islander	367,104	0.15	353,509	0.14
Non-specified race only	(X)	(X)	467,770	0.19
Two races	3,144,327	1.28	4,257,110	1.73
Specified race only	3,144,327	1.28	3,050,906	1.24
Specified and non-specified races	(X)	(X)	1,206,204	0.49
Three or more races	261,921	0.11	345,036	0.14
Specified race only	261,921	0.11	248,365	0.10
Specified and non-specified races	(X)	(X)	96,671	0.04

(X) Not applicable.

U.S. Census Bureau, Census 2000   Procedure   3-3

I.1.2.2 County-Level Forecasts

Woods & Poole (2006) developed county-level forecasts for each year from 2000 through 2030, by age and gender for non-Hispanic White, African-American, Asian-American, and Native-American and for all Hispanics. As discussed below, the adjustments necessary to prepare the data for use in BenMAP are relatively straightforward. The starting data is the following:

> Woods & Poole county-level files
> Data: L:\project data\Benmap\Database Development\Population\Woods and Poole from 2007-2008\Data\
> Docs: http://www.woodsandpoole.com/pdfs/CED07.pdf

For each non-Hispanic subset of the population and each year from 2000-2030, we divided the Woods and Poole population for that year by the Woods and Poole population for that subset in 2000. These serve as the growth coefficients for the non-Hispanic subsets of each race. We used a similar calculation to determine the growth rates for the Hispanic population. We assume that each Hispanic race grows at the same rate, and use these growth rates for the Hispanic subsets of each race.

Matching Age Groups Used in BenMAP

There are 86 age groups, so it is a simple matter of aggregating age groups to match the 19 used in BenMAP.

Matching Counties Used in U.S. Census

The county geographic boundaries used by Woods & Poole are somewhat more aggregated than the county definitions used in the 2000 Census (and BenMAP), and the FIPS codes used by Woods and Poole are not always the standard codes used in the Census. To make the Woods and Poole data consistent with the county definitions in BenMAP, we disaggregated the Woods and Poole data and changed some of the FIPS codes to match the U.S. Census.

Calculating Growth Ratios with Zero Population in 2000

There are a small number of cases were the 2000 county population for a specific demographic group is zero, so the ratio of any future year to the year 2000 data is undefined. In these relatively rare cases, we set the year 2000 ratio and all subsequent ratios to 1, assuming no growth.

I.1.3   PopGrid

If the geographic center of a Census block falls within a population grid-cell, PopGrid assigns the block population to this particular population grid-cell. Note that the grid-cells in an air quality model, such as CMAQ, may cross multiple county boundaries. PopGrid keeps track of the total number of people in each race-ethnic group by county within a particular population grid-cell. Of course, when the population grid-cell is for U.S. counties, then there is only a single county associated with the population grid-cell. However, with air quality models, there can clearly be multiple counties in a population grid-cell.

Keeping track of the total number of people in a county is necessary when forecasting population, as the population forecast for a given grid cell is equal to the year 2000 population estimate from the Census Bureau multiplied by the ratio of future-year to year 2000 county population estimates from Woods & Poole. BenMAP assumes that all age-gender groups within a given race-ethnic group have the same geographic distribution.

I.1.3.1   How to Use PopGrid

After installing PopGrid, double-click on the PopGrid executable "PopGrid4.exe." The following screen will appear:
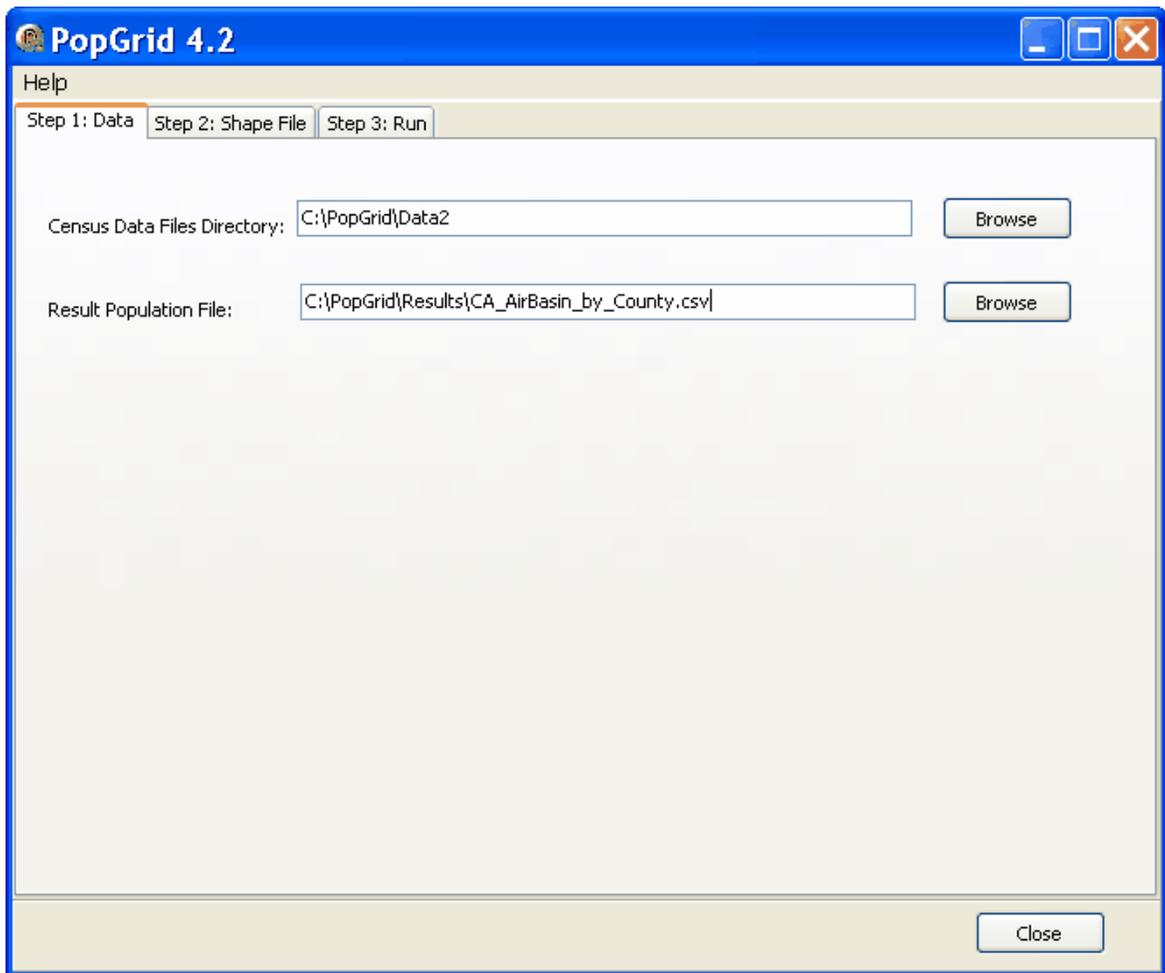
The **Census Data Files Directory** box points PopGrid to where the block data are located that PopGrid uses. Make sure that the files in this directory are unzipped. This data folder should look something like the following:
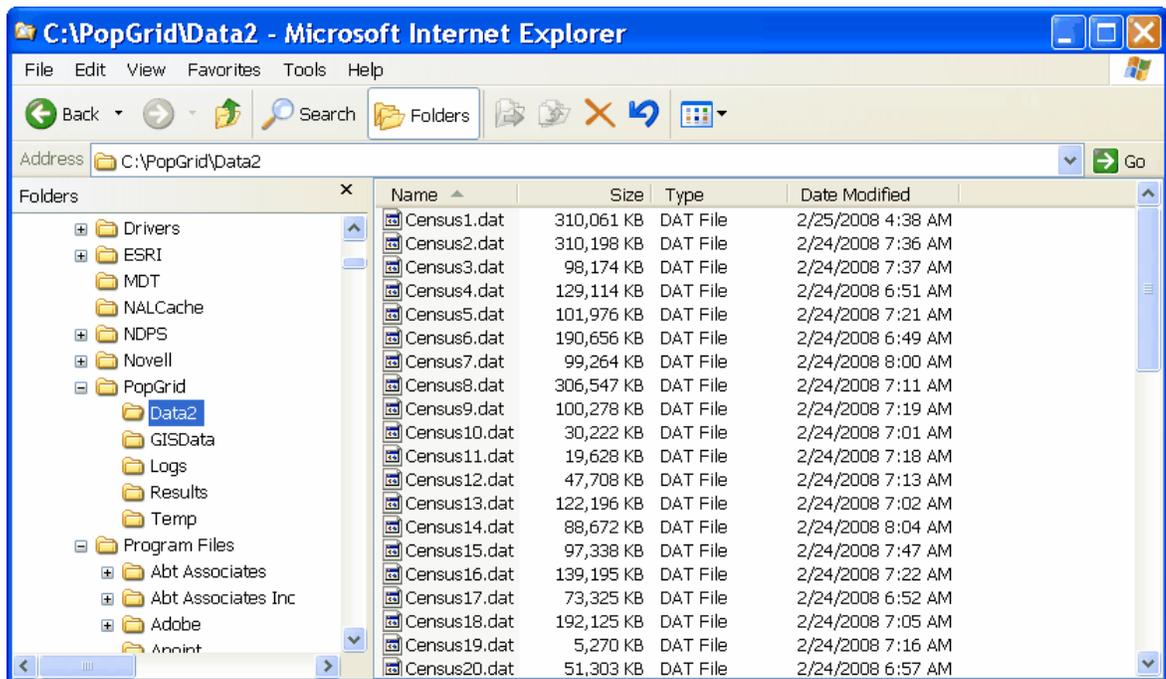
The **Result Population File** box provides the path and the name of the file that you want to create. In the example above, PopGrid is being used to estimate population for the intersection of air basins and counties in California (CA_AirBasin_by_County).

Click on the **Step 2: Shape File** tab. Choose the shapefile that you want to use. The example for air basins and counties in California looks as follows:
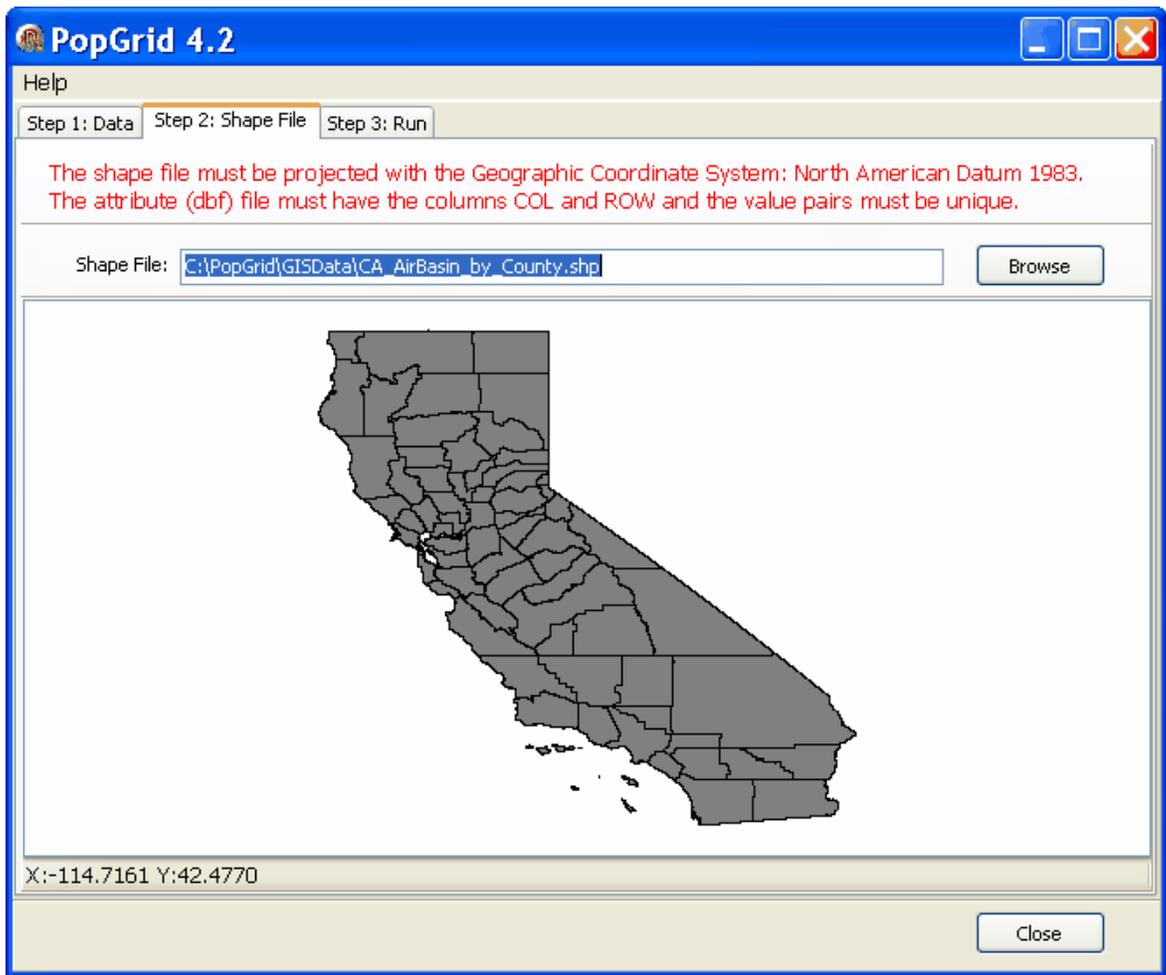
After choosing your shapefile, go to the **Step 3: Run** tab, which should look as follows:
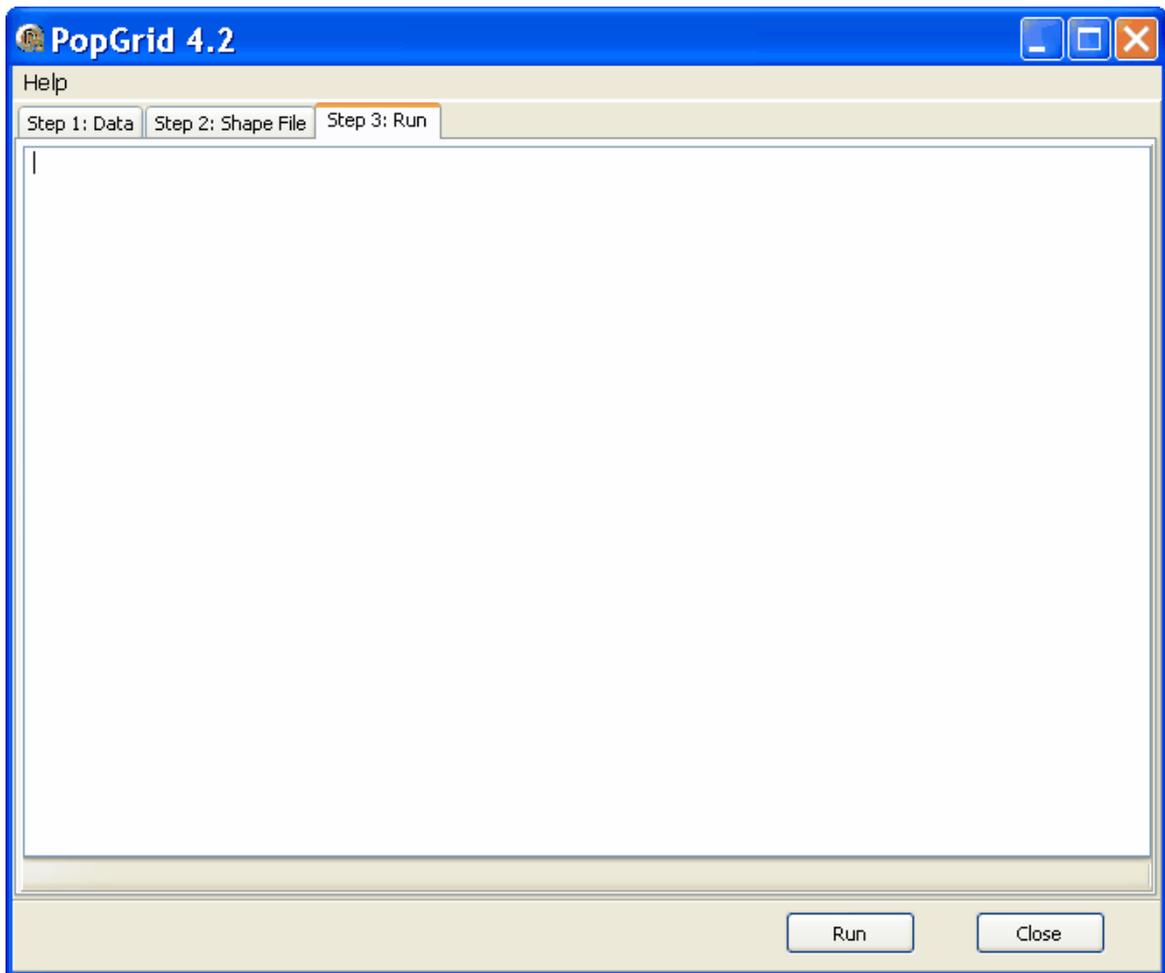
Click **Run**. PopGrid will now begin processing. It can take a very long time to run. When PopGrid has finished running, check the log file. The log file notes the start time, the files that PopGrid used, and the end time. Also, at the very end of the log file, PopGrid notes the number of people that PopGrid assigned to your grid definition ("Population covered by grid") and the number of people that PopGrid determined are outside of your grid definition ("Population outside grid").

I.1.3.2 PopGrid Output

PopGrid generates two files. One file has the number of people in each grid cell for each of the 304 race-ethnicity-gender-age demographic groups available in PopGrid. Table I-1 presents an example of what the population file looks like from PopGrid. The Row and Column uniquely identify each grid cell. Note that the Race, Ethnicity, Gender and AgeRange variables are precisely defined (see section on loading population data LoadData_Setups_Population).

Table I-1. Population File Fragment from PopGrid

Row	Column	Year	Population	Race	Ethnicity	Gender	AgeRange
58	81	2000	1.54	WHITE	HISPANIC	MALE	0TO0
58	81	2000	0.03	BLACK	HISPANIC	MALE	0TO0
58	81	2000	0.01	NATAMER	HISPANIC	MALE	0TO0
58	81	2000	0.01	ASIAN	HISPANIC	MALE	0TO0
58	81	2000	4.86	WHITE	HISPANIC	MALE	1TO4
58	81	2000	0.12	BLACK	HISPANIC	MALE	1TO4
58	81	2000	0.03	NATAMER	HISPANIC	MALE	1TO4
58	81	2000	0.03	ASIAN	HISPANIC	MALE	1TO4

58	81	2000	6.79	WHITE	HISPANIC	MALE	5TO9
58	81	2000	0.21	BLACK	HISPANIC	MALE	5TO9
58	81	2000	0.05	NATAMER	HISPANIC	MALE	5TO9
58	81	2000	0.05	ASIAN	HISPANIC	MALE	5TO9
58	81	2000	0.90	WHITE	HISPANIC	MALE	10TO14
58	81	2000	0.04	BLACK	HISPANIC	MALE	10TO14
58	81	2000	0.01	NATAMER	HISPANIC	MALE	10TO14
58	81	2000	0.01	ASIAN	HISPANIC	MALE	10TO14
58	81	2000	3.44	WHITE	HISPANIC	MALE	15TO19
58	81	2000	0.15	BLACK	HISPANIC	MALE	15TO19
58	81	2000	0.04	NATAMER	HISPANIC	MALE	15TO19
58	81	2000	0.04	ASIAN	HISPANIC	MALE	15TO19
58	81	2000	1.49	WHITE	HISPANIC	MALE	20TO24
58	81	2000	0.06	BLACK	HISPANIC	MALE	20TO24
58	81	2000	0.02	NATAMER	HISPANIC	MALE	20TO24
58	81	2000	0.03	ASIAN	HISPANIC	MALE	20TO24
58	81	2000	1.93	WHITE	HISPANIC	MALE	25TO29
58	81	2000	0.04	BLACK	HISPANIC	MALE	25TO29
58	81	2000	0.01	NATAMER	HISPANIC	MALE	25TO29
58	81	2000	0.01	ASIAN	HISPANIC	MALE	25TO29
58	81	2000	1.87	WHITE	HISPANIC	MALE	30TO34
58	81	2000	0.08	BLACK	HISPANIC	MALE	30TO34

PopGrid generates a second file that keeps track of the fraction of the total population in each of the eight race-ethnic groups that comes from each county in the United States. Table I-2 presents a sample. The SourceCol and SourceRow uniquely identify each county, and the TargetCol and TargetRow uniquely identify each grid cell. The Value variable gives the fraction of the total population in the grid cell for a given race-ethnic group that comes from the "source" county.

When a grid cell lies completely within a county, then the fraction will be 1. When a grid cell is in more than county, then the sum of the fractions across the counties for a given race-ethnic group must sum to one. In Table I-2, you can see that for grid cell (TargetCol=123, TargetRow=18) that the fraction of Asian Non-Hispanic coming from county (SourceCol=16, SourceRow=71) is 0.49 and for county (SourceCol=49, SourceRow=3) the fraction is 0.51. In this case, about half the population of Asian Non-Hispanics comes from each of the two counties. In the case of Black Hispanics, the fraction from county (SourceCol=16, SourceRow=71) is only 0.12, with most Black Hispanics in this grid cell coming from county (SourceCol=49, SourceRow=3).

Table I-2. Population-Weight File Fragment from PopGrid

SourceCol	SourceRow	TargetCol	TargetRow	Race	Ethnicity	Value	Year
16	71	123	18	ASIAN	NON-HISPANIC	0.49	2000
16	71	123	18	ASIAN	HISPANIC	0.21	2000
16	71	123	18	BLACK	NON-HISPANIC	0.49	2000
16	71	123	18	BLACK	HISPANIC	0.12	2000
16	71	123	18	NATAMER	NON-HISPANIC	0.98	2000
16	71	123	18	NATAMER	HISPANIC	0.43	2000
16	71	123	18	WHITE	NON-HISPANIC	0.23	2000
16	71	123	18	WHITE	HISPANIC	0.06	2000
49	3	123	18	ASIAN	NON-HISPANIC	0.51	2000
49	3	123	18	ASIAN	HISPANIC	0.79	2000
49	3	123	18	BLACK	NON-HISPANIC	0.51	2000
49	3	123	18	BLACK	HISPANIC	0.88	2000
49	3	123	18	NATAMER	NON-HISPANIC	0.02	2000
49	3	123	18	NATAMER	HISPANIC	0.57	2000
49	3	123	18	WHITE	NON-HISPANIC	0.77	2000
49	3	123	18	WHITE	HISPANIC	0.94	2000
6	23	45	1	ASIAN	NON-HISPANIC	0.00	2000
6	23	45	1	ASIAN	HISPANIC	0.00	2000
6	23	45	1	BLACK	NON-HISPANIC	0.00	2000
6	23	45	1	BLACK	HISPANIC	0.00	2000
6	23	45	1	NATAMER	NON-HISPANIC	1.00	2000
6	23	45	1	NATAMER	HISPANIC	1.00	2000
6	23	45	1	WHITE	NON-HISPANIC	1.00	2000
6	23	45	1	WHITE	HISPANIC	1.00	2000
6	23	45	2	ASIAN	NON-HISPANIC	1.00	2000
6	23	45	2	ASIAN	HISPANIC	1.00	2000
6	23	45	2	BLACK	NON-HISPANIC	1.00	2000
6	23	45	2	BLACK	HISPANIC	1.00	2000
6	23	45	2	NATAMER	NON-HISPANIC	1.00	2000
6	23	45	2	NATAMER	HISPANIC	1.00	2000
6	23	45	2	WHITE	NON-HISPANIC	1.00	2000
6	23	45	2	WHITE	HISPANIC	1.00	2000

I.2   Monitor Data in U.S. Setup

BenMAP-ready data files were created from 2000 through 2007 data, as reported to the U. S. Environmental Protection Agency's (EPA) Air Quality System (AQS), for PM2.5, PM10 STP and LC, lead TSP, ozone, NO2, SO2, and CO. Table I-5 summarizes the data sources and vintage of the processed data.

Table I-5. Underlying data sources for BenMAP air quality data files.

Pollutant	AQS Parameter Code	Year	Data Source	Date Acquired	Documented Vintage
PM$_{2.5}$	88101	2000–2006	http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	5/22/2008	4/10/2008
		2007	Requested from AQS representative	6/4/2008	
Ozone	44201	2000–2006	http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	5/22/2008	4/14/2008–4/15/2008
		2007	Requested from AQS representative	6/4/2008	
Lead TSP	12128	2000	http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	5/29/2008	6/20/2007
		2001–2006	http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	5/29/2008	4/9/2008
		2007	Requested from AQS representative	6/4/2008	
PM$_{10}$ STP	81102	2000–2006	http://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm	5/22/2008	4/10/2008
		2007	Requested from AQS representative	6/4/2008	
PM$_{10}$ LC	85101	2000–2007	Requested from AQS representative	6/4/2008	

The AQS data were uploaded to the STI Air Quality Archive (AQA) Oracle database. The AQA database performs additional quality control (QC) checks against the AQS data, such as uniqueness by AQS site, method, parameter occurrence code (POC), and duration codes; checks of minimum and maximum values; and maximum rate of change between consecutive data values (where appropriate). The specific QC checks imposed on the BenMAP data are outlined in Table I-6. No maximum value filters were applied to the concentration data. High aerosol concentration values caused by dust storms or other exceptional events are included in the BenMAP-ready data files.

Table I-6. Pollutant-specific QC checks performed in AQA

Pollutant	AQS Parameter Code	Acceptable Concentration Range	Maximum Rate of Change
PM$_{2.5}$	88101	>= 0 µg/m^3	–
Ozone	44201	>= 0 ppb	60 ppb
Lead	12128	>= 0 µg/m^3	–
PM$_{10}$	81102 and 85101	>= 0 µg/m^3	–
CO	42101	>= 0 ppm	–
NO$_2$	42602	>= 0 ppb	50 ppb
SO$_2$	42401	>= 0 ppb	–

I.2.1   Data Processing

STI developed data processing procedures consistent with those used in the past by Abt Associates to create air quality data for files for use in the BenMAP model. Critical data processing rules implemented in the deliverable data are listed below:

Data delivered by STI are reported with consistent units: µg/m3 for aerosols; ppb for ozone, NO2, and SO2; and ppm for CO.

The "monitor name" field is populated by concatenating the AQS site, parameter, and POC codes.

The "monitor description" field is populated with the following metadata: method code, land use, location setting, POC, and AQS parameter code. The AQS probe location and monitoring objective code fields are left blank in STI-processed data.

The data were formatted with one record per site, pollutant, POC, and year for use in the BenMAP program. Data for 365 days, or 8,760 hourly values, are expected per record. This format is satisfied regardless of leap years; an average of February 28 and 29 data are reported.

The monitoring method is allowed to change over the course of a year. To provide a more complete record, data with multiple method codes for a given site, parameter, POC, and year were combined and the first reported method code was reported in the BenMAP-ready data files.

Aerosol data collected with 24-hr sample durations were used before data collected with underlying 1-hr sample durations.One-hour sampling duration data are used for ozone, NO2, SO2,, and CO.

I.2.2   Output Files

Table I-7 lists the number of monitors by pollutant and year, represented in the resulting BenMAP-ready data files.

Table I-7. Number of monitors by pollutant, AQS parameter code, and year included in the BenMAP-ready data files

Pollutant	AQS Parameter Code	Number of Monitors by Year							
		2000	2001	2002	2003	2004	2005	2006	2007
$PM_{2.5}$	88101	1,311	1,339	1,328	1,316	1,226	1,260	1,197	1,144
PM_{10} STP	81102	1,415	1,388	1,320	1,240	1,211	1,199	1,164	1,111
PM_{10} LC	85101	868	757	714	694	688	748	558	502
Lead TSP	12128	209	232	253	250	231	247	204	175
Ozone	44201	1,138	1,184	1,191	1,210	1,206	1,194	1,199	1,217
NO_2	42602	444	458	445	444	446	437	428	423
CO	42101	523	518	498	482	452	429	413	389
SO2	42401	613	605	580	561	557	534	518	520

Appendix J: Uncertainty & Pooling

This Appendix discusses the treatment of uncertainty in BenMAP, both for incidence changes and associated dollar benefits. Some background is then given on pooling methodology. Finally, the mechanics of the various *Pooling Methods* available in BenMAP are discussed in detail, including *Subjective Weight* based pooling, *Fixed Effects* pooling, *Random / Fixed Effects* pooling, and independent and dependent *Sum* and *Subtraction*.

J.1   Uncertainty

Although there are several sources of uncertainty affecting estimates of incidence changes and associated benefits, the sources of uncertainty that are most readily quantifiable in benefit analyses are uncertainty surrounding the health impact functions and uncertainty surrounding unit dollar values. The total dollar benefit associated with a given endpoint group depends on how much the endpoint group will change in the control scenario (e.g., how many premature deaths will be avoided) and how much each unit of change is worth (e.g., how much a statistical death avoided is worth).

Both the uncertainty about the incidence changes and uncertainty about unit dollar values can be characterized by *distributions*. Each "uncertainty distribution" characterizes our beliefs about what the true value of an unknown (e.g., the true change in incidence of a given health effect) is likely to be, based on the available information from relevant studies. Although such an "uncertainty distribution" is not formally a Bayesian posterior distribution, it is very similar in concept and function (see, for example, the discussion of the Bayesian approach in Kennedy 1990, pp. 168-172). Unlike a sampling distribution (which describes the possible values that an *estimator* of an unknown value might take on), this uncertainty distribution describes our beliefs about what values the unknown value itself might be.

Such uncertainty distributions can be constructed for each underlying unknown (such as a particular pollutant coefficient for a particular location) or for a function of several underlying unknowns (such as the total dollar benefit of a regulation). In either case, an uncertainty distribution is a characterization of our beliefs about what the unknown (or the function of unknowns) is likely to be, based on all the available relevant information. Uncertainty statements based on such distributions are typically expressed as 90 percent credible intervals. This is the interval from the fifth percentile point of the uncertainty distribution to the ninety-fifth percentile point. The 90 percent credible interval is a " credible range" within which, according to the available information (embodied in the uncertainty distribution of possible values), we believe the true value to lie with 90 percent probability. The uncertainty surrounding both incidence estimates and dollar benefits estimates can be characterized quantitatively in BenMAP. Each is described separately below.

J.1.1   Characterization of Uncertainty Surrounding Incidence Changes

To calculate point estimates of the changes in incidence of a given adverse health effect

associated with a given set of air quality changes, BenMAP performs a series of calculations at each grid-cell. First, it accesses the health impact functions needed for the analysis, and then it accesses any data needed by the health impact functions. Typically, these include the grid-cell population, the change in population exposure at the grid-cell, and the appropriate baseline incidence rate. BenMAP then calculates the change in incidence of adverse health effects for each selected health impact function. The resulting incidence change is stored, and BenMAP proceeds to the next grid-cell, where the above process is repeated.

In *Latin Hypercube* mode, BenMAP reflects the uncertainty surrounding estimated incidence changes (resulting from the sampling uncertainty surrounding the pollutant coefficients in the health impact functions used) by producing a *distribution* of possible incidence changes rather than a single point estimate. To do this, it uses the distribution (*Dist Beta*) associated with the pollutant coefficient (*Beta*, or β), and potentially the point estimate (*Beta*) and two parameters (*P1Beta*, *P2Beta*). Typically, pollutant coefficients are normally distributed, with mean *Beta* and standard deviation *P1Beta*.

BenMAP uses an N-point Latin Hypercube to represent the underlying distribution of β and to create a corresponding distribution of incidence changes in each population grid cell, where N is specified by you. The Latin Hypercube method represents an underlying distribution by N percentile points of the distribution, where the nth percentile point is equal to:

$$(n-1)\cdot\frac{100}{N}+\frac{100}{2N}$$

The Latin Hypercube method is used to enhance computer processing efficiency. It is a sampling method that divides a probability distribution into intervals of equal probability, with an assumption value for each interval assigned according to the interval's probability distribution. Compared with conventional Monte Carlo sampling, the Latin Hypercube approach is more precise over a fewer number of trials because the distribution is sampled in a more even, consistent manner (Decisioneering, 1996, pp. 104-105).

Suppose, for example, that you elect to use a 20-point Latin Hypercube. BenMAP would then represent the distribution of β by 20 percentile points, specifically the 2.5^{th}, 7.5^{th}, ..., 97.5^{th}. To do this, the inverse cumulative distribution function specified by the distribution of β is called with the input probability equal to each the 20 percentile points. BenMAP then generates an estimate of the incidence change in a grid-cell for each of these values of β, resulting in a distribution of N incidence changes. This distribution is stored, and BenMAP proceeds to the next population grid-cell, where the process is repeated.

J.1.2   Characterization of Uncertainty Surrounding Dollar Benefits

The uncertainty distribution of the dollar benefits associated with a given health or welfare effect is derived from the two underlying uncertainty distributions – the distribution of the change in incidence of the effect (number of cases avoided) and the distribution of the

value of a case avoided (the "unit value"). The derivation of the uncertainty distribution for incidence change is described above. The distributions used to characterize the uncertainty surrounding unit values are described in detail in the appendix on the Economic Value of Health Effects. As noted in that Appendix, a variety of distributions have been used to characterize the uncertainty of unit values, including uniform, triangular, normal, and Weibull.

To represent the underlying distribution of uncertainty surrounding unit values, a 100-point Latin Hypercube is generated in the same way described in the previous section for the distribution of β. That is, the unit value distribution is represented using the 0.5^{th}, 1.5^{th}, ..., and 99.5^{th} percentile values of its distribution.

A distribution of the uncertainty surrounding the dollar benefits associated with a given endpoint is then derived from Latin Hypercube values generated to represent the change in incidence and the Latin Hypercube values generated to represent the unit value distribution. To derive this new distribution, each of the 100 unit values is multiplied by each of the N incidence change values, yielding a set of 100 * N dollar benefits. These values are sorted low to high and binned down to a final distribution of N dollar benefit values.

J.1.3   Characterization of Uncertainty Surrounding QALY Estimates

The uncertainty distribution of the QALY estimates associated with a given health effect is similar to that for dollar benefits. That is, it is derived from the two underlying uncertainty distributions – the distribution of the change in incidence of the effect (number of cases avoided) and the distribution of the QALYs per case avoided. The derivation of the uncertainty distribution for incidence change is described above. The distributions used to characterize the uncertainty surrounding QALYs are described in detail in the appendix on the Economic Value of Health Effects. As noted in that Appendix, a variety of distributions have been used to characterize the uncertainty of unit values, including uniform, triangular, normal, and Weibull.

To represent the underlying distribution of uncertainty surrounding unit values, a 100-point Latin Hypercube is generated in the same way described in the previous section for the distribution of β. That is, the unit value distribution is represented using the 0.5^{th}, 1.5^{th}, ..., and 99.5^{th} percentile values of its distribution.

A distribution of the uncertainty surrounding the QALYs associated with a given endpoint is then derived from Latin Hypercube values generated to represent the change in incidence and the Latin Hypercube values generated to represent the QALY distribution. To derive this new distribution, each of the 100 QALY weights is multiplied by each of the N incidence change values. These values are sorted low to high and binned down to a final distribution of QALY values.

J.2   Pooling

There is often more than one study that has estimated a health impact function for a given

pollutant-health endpoint combination. Each study provides an estimate of the pollutant coefficient, β, along with a measure of the uncertainty of the estimate. Because uncertainty decreases as sample size increases, combining data sets is expected to yield more reliable estimates of β, and therefore more reliable estimates of the incidence change predicted using β. Combining data from several comparable studies in order to analyze them together is often referred to as meta-analysis.

For a number of reasons, including data confidentiality, it is often impractical or impossible to combine the original data sets. Combining the *results* of studies in order to produce better estimates of β provides a second-best but still valuable way to synthesize information . This is referred to as pooling. Pooling β's requires that all of the studies contributing estimates of β use the same functional form for the health impact function. That is, the β's must be measuring the same thing.

It is also possible to pool the study-specific estimates of incidence change derived from the health impact functions, instead of pooling the underlying β's themselves. For a variety of reasons, this is often possible when it is not feasible to pool the underlying β's. For example, if one study is log-linear and another is linear, we could not pool the β's because they are not different estimates of a coefficient in the same health impact function, but are instead estimates of coefficients in different health impact functions. We can, however, calculate the incidence change predicted by each health impact function (for a given change in pollutant concentration and, for the log-linear function, a given baseline incidence rate), and pool these incidence changes. BenMAP allows the pooling of incidence changes predicted by several studies for the same pollutant-health endpoint group combination. It also allows the pooling of the corresponding study-specific estimates of monetary benefits.

As with estimates based on only a single study, BenMAP allows you to characterize the uncertainty surrounding pooled estimates of incidence change and/or monetary benefit. To do this, BenMAP pools the study-specific distributions of incidence changes (or monetary benefit or QALYs) to derive a pooled distribution. This pooled distribution incorporates information from all the studies used in the pooling procedure.

J.2.1   Weights Used for Pooling

The relative contribution of any one study in the pooling process depends on the weight assigned to that study. A key component of the pooling process, then, is the determination of the weight given to each study. There are various methods that can be used to assign weights to studies. Below we discuss the possible weighting schemes that are available in BenMAP.

J.2.1.1 Subjective Weights

BenMAP allows you the option of specifying the weights to be used. Suppose, for example, you want to simply average all study-specific results. You would then assign a weight of 1/N to each of the N study-specific distributions that are to be pooled. Note that subjective weights are limited to two decimal places, and are normalized to sum to one, if

they do not already sum to one.

J.2.1.2 Automatically Generated Weights

A simple average has the advantage of simplicity but the disadvantage of not taking into account the uncertainty of each of the estimates. Estimates with great uncertainty surrounding them are given the same weight as estimates with very little uncertainty. 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. BenMAP has two methods of automatically generating pooling weights using the variances of the input distributions - *Fixed Effects Pooling* and *Random / Fixed Effects Pooling*.

The discussion of these two weighting schemes is first presented in terms of pooling the pollutant coefficients (the β's), because that most closely matches the discussion of the method for pooling study results as it was originally presented by DerSimonian and Laird. We then give an overview of the analogous weighting process used within BenMAP to generate weights for incidence changes rather than β's.

J.2.1.3 Fixed-Effect Weights

The fixed effects model assumes that there is a single true concentration-response relationship and therefore a single true value for the parameter β that applies everywhere. 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). Fixed effects pooling 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

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

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:

$$\beta_{fe} = \sum w_i * \beta_i .$$

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

$$v_{fe} = \frac{1}{\sum 1/v_i} .$$

Exhibit J-2 shows the relevant calculations for this pooling for three sample studies.

Exhibit J-2. Example of Fixed Effects Model Calculations

Study	β_i	v_i	$1/v_i$	w_i	$w_i*\beta_i$
1	0.75	0.1225	8.16	0.016	0.012
2	1.25	0.0025	400	0.787	0.984
3	1.00	0.0100	100	0.197	0.197
Sum			$\sum = 508.16$	$\sum = 1.000$	$\sum = 1.193$

The sum of weighted contributions in the last column is the pooled estimate of β based on the fixed effects model. This estimate (1.193) is considerably closer to the estimate from study 2 (1.25) than is the estimate (1.0) that simply averages the study estimates. This reflects the fact that the estimate from study 2 has a much smaller variance than the estimates from the other two studies and is therefore more heavily weighted in the pooling.

The variance of the pooled estimate, v_{fe}, is the inverse of the sum of the variances, or 0.00197. (The sums of the β_i and v_i are not shown, since they are of no importance. The sum of the $1/v_i$ is S, used to calculate the weights. The sum of the weights, w_i , i=1, ..., n, is 1.0, as expected.)

J.2.1.4 Random- / Fixed-Effect Weights

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 PM10 on mortality, for example, if the composition of PM10 varies among study locations the underlying relationship between mortality and PM10 may be different from one study location to another. For example, fine particles make up a greater fraction of PM10 in Philadelphia than in El Paso. If fine particles are disproportionately responsible for mortality relative to coarse particles, then one would expect the true value of β in Philadelphia to be greater than the true value of β in El Paso. 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, Qw , the weighted sum of squared differences of the separate study estimates from the pooled estimate based on the fixed effects model, is calculated as:

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

Under the null hypothesis that there is a single underlying parameter, β, of which all the β_i 's are estimates, Qw has a chi-squared distribution with n-1 degrees of freedom. (Recall that n is the number of studies in the meta-analysis.) If Qw 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. (BenMAP uses a five percent one-tailed test).

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 Qw , the between-study variance, η2, 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 Qw < n-1), then η2 is a negative number, and it is not possible to calculate a random effects estimate. In this case, however, the small value of Qw 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 η2 is positive.

Given a value for η2 , 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 (vi) and the between-study variance (η2). Whereas the weights implied by the fixed effects model used only vi, the within-study variance, the weights implied by the random effects model use vi +η 2.

Let vi* = vi +η2. Then:

$$S^* = \sum \frac{1}{v_i^*} \; ,$$

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

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:

$$\beta_{rand} = \sum w_i^* * \beta_i \; .$$

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:

$$v_{rand} = \frac{1}{\sum 1/v_i^*} \; .$$

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.

Fixed Effects and Random / Fixed Effects Weighting to Pool Incidence Change Distributions and Dollar Benefit Distributions

Weights can be derived for pooling incidence changes predicted by different studies, using either the fixed effects or the fixed / random effects model, in a way that is analogous to the derivation of weights for pooling the â's in the C-R functions. As described above, BenMAP generates a Latin Hypercube representation of the distribution of incidence change corresponding to each health impact function selected. The means of those

study-specific Latin Hypercube distributions of incidence change are used in exactly the same way as the reported â's are used in the calculation of fixed effects and random effects weights described above. The variances of incidence change are used in the same way as the variances of the â'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 mean incidence change for the ith health impact function for âi and the variance of incidence change for the ith health impact function for vi.200

Similarly, weights can be derived for dollar benefit distributions. As described above, BenMAP generates a Latin Hypercube representation of the distribution of dollar benefits . The means of those Latin Hypercube distributions are used in exactly the same way as the reported â's are used in the calculation of fixed effects and random effects weights described above. The variances of dollar benefits are used in the same way as the variances of the â'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 mean dollar benefit change for the ith valuation for âi and the variance of dollar benefits for the ith valuation for vi.

BenMAP always derives Fixed Effects and Random / Fixed Effects weights using nationally aggregated results, and uses those weights for pooling at each grid cell (or county, etc. if you choose to aggregate results prior to pooling). This is done because BenMAP does not include any regionally based uncertainty – that is, all uncertainty is at the national level in BenMAP, and all regional differences (population, for example) are treated as certain.

J.2.2 Mechanics of Pooling in BenMAP

Once weights are generated for each input distribution, BenMAP has three options for using these weights to combine the input distributions into a single new distribution. These options are referred to as Advanced Pooling Methods.

Round Weights to Two Digits

This is BenMAP's default Advanced Pooling Method, and is always the method used when Subjective Weights are used. The first step is converting the weights to two digit integers by multiplying them by 100 and rounding to the nearest integer. If all the integral weights thus generated are divisible by the smallest weight, they are each divided by that smallest weight. For example, if the original weights were 0.1, 0.2, 0.3, and 0.4, the resulting integral weights would be 10/10, 20/10, 30/10, and 40/10 (or 1, 2, 3, and 4).

BenMAP then creates a new distribution by sampling each entire input distribution according to its weight. That is, in the above example the first distribution would be sampled once, the second distribution twice, and so forth. The advantage of sampling whole distributions is that it preserves the characteristics (i.e., the moments - the mean, the variance, etc.) of the underlying distributions. Assuming n latin hypercube points, the resulting distribution will contain a maximum of 100 * n values, which are then sorted low to high and binned down to n values, which will represent the new, pooled distribution.

Round Weights to Three Digits

This Advanced Pooling Method is essentially the same as rounding weights to two digits, except that the weights are converted to three digit integers, and so forth. That is, the weights are multiplied by 1000 and rounded to the nearest integer. Again, if all the integral weights thus generated are divisible by the smallest weight, they are each divided by that smallest weight. Assuming n Latin Hypercube points, the resulting distribution with this Advanced Pooling Method can contain a maximum of 1000 * n values, which are sorted low to high and binned down to n values, which represent the new, pooled distribution.

Exact Weights for Monte Carlo

This Advanced Pooling Method uses a Monte Carlo method to combine the input distributions. Using this method, on each of many iterations, (1) an input distribution is selected (with the probability of selection equal to the weight assigned to the distribution), and (2) a value is randomly drawn from that distribution. Values chosen in this way are placed into a temporary pooled distribution, which will have one point per iteration of the Monte Carlo method. The number of iterations is specified by the user, and defaults to 5,000. After the temporary distribution is fully generated, it is sorted low to high and binned down to n values (where n is the number of Latin Hypercube Points chosen for the analysis).

J.2.3   Summing Distributions

Sometimes rather than pooling distributions we want to add them. For example, some studies have estimated a health impact function for hospital admissions for COPD and another health impact function for hospital admissions for pneumonia. From each of these health impact functions, BenMAP can derive the corresponding distributions for incidence change. Hospital admissions for COPD and pneumonia are two of the most important components of respiratory hospital admissions, and we may want to estimate the number of cases of "respiratory hospital admissions," as characterized by being either COPD or pneumonia. To do this we would add the two distributions.

Summing across distributions can be done in one of two ways: We can assume the two distributions are independent of each other or dependent. Which is the more reasonable assumption depends on the particulars of the distributions being summed.

Assuming Independence

This is the Sum (Independent) Pooling Method. To sum two distributions that are independent, on each of many iterations of a Monte Carlo procedure, BenMAP (1) randomly selects a value from the first input distribution, (2) randomly selects a value from the second input distribution, and (3) adds the two values together. To sum N distributions that are independent, BenMAP follows an analogous procedure in which, on each iteration it makes a random selection from each of the input distributions and then adds the results together. When the Monte Carlo procedure is completed, all such generated results are sorted low to high and binned down to the appropriate number of latin hypercube points. The number of iterations is determined by the Monte Carlo Iterations setting.

Assuming Dependence

This is the Sum (Dependent) Pooling Method. Recall that the uncertainty distributions in BenMAP are latin hypercube representations, consisting of N percentile points. To sum two distributions assumed to be dependent, BenMAP simply generates a new N point latin hypercube where each point is the sum of the corresponding points from the input latin hypercubes. That is, the first point in the new latin hypercube is the sum of the first points in the two input latin hypercubes, and so forth. To sum n distributions that are assumed to be dependent, BenMAP follows an analogous procedure in which each point in the new latin hypercube is the sum of the corresponding points from each of the input latin hypercubes.

J.2.4   Subtracting Distributions

In some cases, you may want to subtract one or more distribution(s) from another. For example, one study may have estimated a health impact function for minor restricted activity days (MRADs), and another study may have estimated a health impact function for asthma "episodes." You may want to subtract the change in incidence of asthma episodes from the change in incidence from MRADs before estimating the monetary value of the MRADs, so that the monetary value of asthma episodes avoided will not be included.

Subtracting across distributions can be done in one of two ways: we can assume the two distributions are independent of each other or dependent. Which is the more reasonable assumption depends on the particulars of the distributions being subtracted.

Assuming Independence

This is the Subtraction (Independent) Pooling Method. To subtract one distribution from another, assuming independence, on each of many iterations of a Monte Carlo procedure, BenMAP (1) randomly selects a value from the first input distribution, (2) randomly selects a value from the second input distribution, and (3) subtracts the second value from the first. To subtract N distributions from another distribution, assuming independence, BenMAP follows an analogous procedure in which, on each iteration it makes a random selection from each of the input distributions and then subtracts the second through the Nth from the first. When the Monte Carlo procedure is completed, all such generated results are sorted low to high and binned down to the appropriate number of Latin Hypercube points. The number of iterations is determined by the Monte Carlo Iterations setting.

Assuming Dependence

This is the Subtraction (Dependent) Pooling Method (see Chapter 6 for details). Recall that the uncertainty distributions in BenMAP are Latin Hypercube representations, consisting of N percentile points. To subtract one distribution from another, assuming them to be dependent, BenMAP simply generates a new N point Latin Hypercube where each point is the result of subtracting the corresponding point of the second input Latin Hypercube from the corresponding point of the first input Latin Hypercube. That is, the first point in the new Latin Hypercube is the result of subtracting the first point in the second Latin Hypercube from the first point of the first Latin Hypercube, and so forth. To

subtract n distributions from another distribution, assuming dependence, BenMAP follows an analogous procedure in which each point in the new Latin Hypercube is the result of subtracting the corresponding points of the second through the Nth input Latin Hypercubes from the corresponding point of the first.

Appendix K: Command Line BenMAP

The command line version of BenMAP is capable of performing all of the functions of the GUI-based version. It is most useful for large, complex analyses that require generation of a substantial number of files. This appendix describes the syntax and use of the command line version.

K.1   Overview

The overall format of the file is a variable definitions section followed by a commands section.

Comment statements are supported at any point in the file. Lines beginning with a pound character (#) are considered comment lines and will be ignored during file parsing.

Additionally, LOAD <filename> statements are supported at any point in the file. These work as string replacements – the contents of the file specified by <filename> are simply inserted into the main file. Multi-level LOAD statements are supported, but no attempt is made to detect cycles (two files referencing each other with LOAD statements, for example).

The control file is, in general, not case sensitive. In the case of user-defined strings, (variable values, etc.), it is preserved.

K.2   Variables

The variable definitions section is optional, and if present will consist of a single line with the word "Variables" on it, followed by one or more lines that define variables. A variable definition consists of a variable name and a variable value. When parsing lines in the commands section of the control file, all occurrences of the variable name will be replaced by the variable value.

All variable names must begin and end with the percent character (%).

Variable Name/Value replacement will be done in multiple passes (until no variable names remain), so variable values may contain other variable names. No attempt will be made to detect cycles, however, so be careful not to introduce them. For example, avoid variable definitions like the following:

%BENMAPDIR%   %AQGDIR%\

%AQGDIR%   %BENMAPDIR%\Air Quality Grids

Variable values must be contained in a single line, and will consist of the first

non-whitespace character after the variable name through the newline character. Watch out for undesired trailing whitespaces!

K.3 Commands

The commands section is required, and will consist of one or more command sections. There are five types of command sections:

SETACTIVESETUP

CREATE AQG

RUN CFG

RUN APV

GENERATE REPORT

This section will discuss each one in turn.

In general, in command sections, there must be at least one white space between each token (where a token is either a command, a parameter name, or a parameter value). Additional white space is ignored, including newline characters. To include white space in a parameter value, you must enclose the parameter in double quotes. The double quotes will not be included in the parameter value in this case (If you wish to include beginning and trailing double quotes in a parameter value, put two in a row at the beginning and end – e.g. ""Look at all those double quotes."").

K.3.1 Set Active Setup

For the US version of the BenMAP command line executable the only valid value is United States. The SETACTIVESETUP section is required.

Example

-ActiveSetup "United States"

K.3.2 Create AQG

This section initiates the creation of one or more air quality grids (normally one, potentially two in the case of monitor rollback grid creation – see below). It always starts with the words **CREATE AQG**. It must then include the following options, in any order:

-Filename <filename>
-Gridtype <gridtype>

-Pollutant <pollutant>

The **Filename** value is the name of the air quality grid that will be created.

The **GridType** value must be one found in the BenMAP database. The actual values for this parameter are found on the Modify Setup screen in the Grid Definitions list box.

Supported Pollutant values are:

-Ozone
-PM10
-PM2.5

These values are also found on the Modify Setup screen in the Pollutants list box.

After these required options, the type of grid creation must be identified, and then the parameters for that grid creation type must be specified. There are four air quality grid creation types:

-ModelDirect
-MonitorDirect
-MonitorModelRelative
-MonitorRollback

K.3.2.1 Model Direct

This section initiates the creation of a model direct air quality grid.

This creation type has two required parameters:

-ModelFilename <filename>

-DSNName   <ODBC DSN name>

and one optional parameter:

-TableName   <tablename>

Supported DSNName values are:

"Excel Files"   Excel Spreadsheet (.xls)

"Text Files"   Comma-delimited (.csv) files

"MS Access Database"   Access Database (.mdb)

If the DSNName is "Excel Files" and there is more than one worksheet in the workbook or "MS Access Database" and there is more than one table in the database then the TableName parameter must indicate the worksheet or table name.

K.3.2.2 Monitor Direct

This section initiates the creation of a monitor direct air quality grid.

The required parameters are:

-MonitorDataType <DataSource descriptor>

-InterpolationMethod <Interpolation Method>

Valid values for MonitorDataType are:

-Library

-DatabaseRows

-DatabaseColumns

-TextFile

Valid values for Interpolation method are:

-ClosestMonitor

-VNA

If MonitorDataType is Library then the following parameters are required:

-MonitorDataSet <Monitor Dataset Name>

MonitorDataSet is the Dataset name of Monitor data stored in the BenMAP database. These values can be found on the Modify Setup screen in Monitor Datasets list box.

-MonitorYear <Year>

MonitorYear specifies the year of interest in the monitor library.

If MonitorDataType is DatabaseRows then the following parameters are required:

-MonitorFile   <filename>

-DSNName   <ODBC DSN name>

and one optional parameter:

-TableName   <tablename>

Supported DSNName values are:

-"Excel Files"   Excel Spreadsheet (.xls)

-"Text Files"   Comma-delimited (.csv) files

-"MS Access Database" Access Database (.mdb)

If the DSNName is "Excel Files" and there is more than one worksheet in the workbook or "MS Access Database" and there is more than one table in the database then the TableName parameter must indicate the worksheet or table name.

If MonitorDataType is DatabaseColumns then the same parameters for MonitorDataType DatabaseRows are required along with the following:

-MonitorDefFilename

-DefDSNName

-DefTableName

These parameters behave the same as the corresponding DatabaseRows parameters.

If MonitorDataType is TextFile the following parameter is required:

-MonitorFile <filename>

MonitorFile specifies a comma separated values (*.csv, generally) file containing monitor data.

Optional Parameters:

-MaxDistance <real>

Specifies the maximum distance (in kilometers) to be used in ClosestMonitor interpolation or VNA interpolation. Monitors outside this distance will not be considered in the interpolation procedure.

-MaxRelativeDistance <real>

Specifies the maximum relative distance to be used in VNA interpolation, where relative distance is the multiple of the distance to the closest monitor used in the interpolation procedure.

-WeightingMethod <method>

Specifies the weighting procedure used for monitors in VNA interpolation. Supported values are InverseDistance and InverseDistanceSquared. If this parameter is not specified, InverseDistance weighting is used.

K.3.2.3 Monitor Model Relative

This section initiates the creation of a monitor model relative air quality grid. This creation type has all the same required and optional parameters as the MonitorDirect creation type. In addition, it has two/three new required parameters.

Required Parameters:

-ScalingMethod <scaling method>

Supported scaling methods are Spatial, Temporal, and Both.

-BaseYearFilename <filename>

Specifies the base year adjustment file to use in monitor scaling.

-BaseYearDSNName <ODBC DSN Name>

Supported –BaseYearDSNName values are

"Excel Files"	Excel Spreadsheet (.xls)
"Text Files"	Comma-delimited (.csv) files
"MS Access Database"	Access Database (.mdb)

When the ScalingMethod is Temporal or Both, the FutureYearFileName and FutureYearDSNName parameters are required. These specify the future year adjustment file to use in monitor scaling.

K.3.2.4 Monitor Rollback

// MonitorRollback

SpatialScaling   = '-SpatialScaling';

BaselineFilename   = '-BaselineFilename';

// RollbackOptions

Percentage   = '-Percentage';

Increment   = '-Increment';

// RollbackToStandardOptions

Standard   = '-Standard';

Metric   = '-Metric';

Ordinality   = '-Ordinality';

InterdayRollbackMethod = '-InterdayRollbackMethod';

IntradayRollbackMethod = '-IntradayRollbackMethod';

K.3.3   Run CFG

The command line version of BenMAP does not support creation of new .cfg files, both because this would be quite cumbersome to do in plain text, and because it probably is not needed. Slight modifications of existing .cfg files are supported, and it is thought that at this point this should be enough.

As such, the only required parameter to run a configuration is the configuration filename. Optional parameters allow the slight modifications mentioned above.

Required Parameters

-CFGFilename <filename>

 Specifies the .cfg file to run.

-ResultsFilename <filename>

 Specifies the .cfgr file to save the results in.

Optional Parameters

-BaselineAQG <filename>

 Specifies the baseline air quality grid file to use when running the configuration – overrides whatever value is   present in the .cfg file.

-ControlAQG <filename>

 Specifies the control air quality grid file to use when running the configuration – overrides whatever value is   present in the .cfg file.

-Year <Integer>

 Year in which to run the configuration (this will affect the population numbers used) – overrides whatever   value is present in the .cfg file. Supported values are 1990 and up.

-LatinHypercubePoints <integer>

 Number of latin hypercube points to generate when running the configuration (zero means run in point mode),   overrides whatever value is present in the .cfg file.

-Threshold <real>

Threshold to use when running the configuration – overrides whatever value is present in the .cfg file.

K.3.4   Run APV

The command line version of BenMAP does not support creation of new .apv files, both because this would be quite cumbersome to do in plain text, and because it probably is not needed. Slight modifications of existing .apv files are supported, and it is thought that at this point this should be enough.

As such, the only required parameter to run an APV configuration is the APV configuration filename. Optional parameters allow the slight modifications mentioned above.

Required Parameters

-APVFilename <filename>

 Specifies the .apv file to run.

-ResultsFilename <filename>

Specifies the .apvr file to save the results in.

Optional Parameters

-CFGRFilename <filename>

Specifies the .cfgr file to use when running the APV configuration – note that this file must contain the same set of results which the .cfgr file originally used to generate the .apv file contained. Overrides whatever value is   present in the .apv file.

-IncidenceAggregation <aggregation level>

Level to aggregate incidence results to before pooling them. Supported values are None, County, State, and   Nation. Overrides whatever value is present in the .apv file.

-ValuationAggregation <aggregation level>

Level to aggregate valuation results to before pooling them. Supported values are None, County, State, and   Nation (though the value must be greater than or equal to IncidenceAggregation). Overrides whatever value is present in the .apv file.

-RandomSeed <integer>

Random seed to use for all procedures requiring pseudo-random numbers (e.g. monte carlo procedures).   Overrides the default behavior, which is to generate a new random seed each time the APV configuration is   run.

-DollarYear <integer>

Year in which dollar figures should be reported. Supported values are 1980 – 2001. Overrides whatever   value   is present in the .apv file.

K.3.5   Generate Report

Reports come in three main varieties – Audit Trail Reports, which can be generated from any BenMAP file; Configuration Results Reports, which can be generated from .cfgr files; and APV Configuration Results Reports, which can be generated from .apvr files. All these report types need an input filename and an output filename. CFGR reports and APVR reports additionally take many optional parameters.

The format for each report type is:

GENERATE REPORT <ReportType>

-InputFile <filename>

-ReportFile <filename>

<optional parameters>

Supported ReportType values are: AuditTrail, CFGR, and APVR.

K.3.5.1 Audit Trail

Audit trail reports require only the parameters described in the "Generate Report" section.

K.3.5.2 CFGR Report

A CFGR report may be generating using only the parameters described in the "Generate Report" section. However, there are also a number of additional options, described below.

Optional Parameters

-GridFields <comma separated field names>

Specifies the set of grid fields to include in the report. Grid fields include Column and Row. If this parameter is   not present, all fields will be included in the report.

-CustomFields <comma separated field names>

Specifies the set of custom fields (C-R Function identifiers, in this case) to include in the report. If this   parameter is not present, all fields will be included in the report.

-ResultFields <comma separated field names>

Specifies the set of result fields to include in the report. Result fields include Point Estimate, Population, Delta,   Mean, Standard Deviation, Variance, and Latin Hypercube Points. If this

parameter is not present, all fields will   be included in the report.

> -Grouping <grouping method>

> Specifies the grouping for the results – Gridcell, then C-R Function, or C-R Function, then Gridcell. Supported values are GridcellFirst, GridcellLast. The default value is GridcellFirst.

> -DecimalDigits <integer>

> Specifies the number of digits after the decimal point to include in the report. Supported values are zero to eight.   The default value is four.

K.3.5.3 APVR Report

Required Parameters

APVR Reports require one additional parameter beyond those required for Audit Trail or CFGR Reports.

> -ResultType <result type>

Specifies the result type for which a report should be created. Supported result types are: IncidenceResults, AggregatedIncidence, PooledIncidence,Valuation, AggregatedValuation, PooledValuation, QALYValuation, AggregatedQALYValuation and PooledQALYValuation.

Optional Parameters

All of the CFGR report parameters are supported for APVR reports as well, except that Population and Delta are not supported ResultField elements.

> -Totals <total type>

Specifies the type of totals which should be included in the report. Supported types are Dependent and Independent. Totals can only be generated for valuation results (Valuation, AggregatedValuation, and PooledValuation result types).

K.4   Example 1

VARIABLES

%CFG%	C:\BenMAP\CommandLine\Configurations\PM25 Wizard.cfg
%APV%	C:\BenMAP\CommandLine\Configurations\PM25 Wizard.apv
%RESULTSDIR%	C:\BenMAP\Temp
%REPORTDIR%	C:\BenMAP\Temp
%AQG%	C:\BenMAP\CommandLine\Air Quality Grids

COMMANDS

SETACTIVESETUP

    -ActiveSetup   "United States"

CREATE AQG

    -Filename   %AQG%\PM25_2002Baseline_50km.aqg
    -GridType   "CMAQ 12km"
    -Pollutant   PM2.5

    MonitorDirect

    -InterpolationMethod   VNA_Alt
    -MonitorDataType   Library
    -MonitorDataSet   "EPA Standard Monitors"
    -MonitorYear   2002
    -MaxDistance   50

CREATE AQG

    -Filename   %AQG%\PM25_2002Control_50km.aqg
    -GridType   "CMAQ 12km"
    -Pollutant   PM2.5

    MonitorRollback

    -InterpolationMethod   VNA_Alt
    -MonitorDataType   Library
    -MonitorDataSet   "EPA Standard Monitors"
    -MonitorYear   2002
    -RollbackGridType   State
    -MaxDistance   50

    RollbackToStandardOptions

    -Standard   65
    -Metric   D24HourMean
    -InterdayRollbackMethod   Quadratic

RUN CFG

    -CFGFilename   %CFG%
    -ResultsFilename   %RESULTSDIR%\PM25_2002_50km.cfgr
    -BaselineAQG   %AQG%\PM25_2002Baseline_50km.aqg
    -ControlAQG   %AQG%\PM25_2002Control_50km.aqg

RUN APV

    -APVFilename   %APV%
    -ResultsFilename   %RESULTSDIR%\PM25_2002_50km.apvr
    -CFGRFilename   %RESULTSDIR%\PM25_2002_50km.cfgr
    -IncidenceAggregation Nation
    -ValuationAggregation Nation

GENERATE REPORT APVR

    -InputFile   %RESULTSDIR%\PM25_2002_50km.apvr
    -ReportFile   %REPORTDIR%\PM25_2002_50km_IncidenceNation.csv

```
    -ResultType           PooledIncidence
    -CustomFields         "Endpoint Group,Author,Start Age,Endpoint,Qualifier,Pooling
Window"
    -ResultFields         "Mean,Standard Deviation,Latin Hypercube Points"
    -DecimalDigits         0
```

GENERATE REPORT APVR

```
    -InputFile            %RESULTSDIR%\PM25_2002_50km.apvr
    -ReportFile           %REPORTDIR%\PM25_2002_50km_ValuationNation.csv
    -ResultType            PooledValuation
    -CustomFields         "Endpoint Group,Author,Start Age,Endpoint,Qualifier,Pooling
Window"
    -ResultFields         "Mean,Standard Deviation,Latin Hypercube Points"
    -DecimalDigits        0
```

K.5   Example 2

VARIABLES

```
%CFG%                 C:\BenMAP\CommandLine\Configurations\PM25 Wizard.cfg
%APV%                 C:\BenMAP\CommandLine\Configurations\PM25 Wizard.apv
%RESULTSDIR%          C:\BenMAP\Temp
%REPORTDIR%           C:\BenMAP\Temp
%AQG%                 C:\BenMAP\CommandLine\Air Quality Grids
```

COMMANDS

SETACTIVESETUP

```
    -ActiveSetup          "United States"
```

CREATE AQG

```
    -Filename             %AQG%\PM25_2004Baseline.aqg
    -GridType             "County"
    -Pollutant            PM2.5
```

MonitorDirect

```
    -InterpolationMethod  VNA_Alt
    -MonitorDataType      Library
    -MonitorDataSet              "EPA Standard Monitors"
    -MonitorYear          2004
```

CREATE AQG

```
    -Filename             %AQG%\PM25_2004_Control.aqg
    -GridType             "County"
    -Pollutant            PM2.5
```

MonitorRollback

```
    -InterpolationMethod  VNA_Alt
    -MonitorDataType      Library
```

```
-MonitorDataSet              "EPA Standard Monitors"
-MonitorYear           2004
-RollbackGridType      State
-MaxDistance           50
```

RollbackToStandardOptions

```
-Standard              35
-Metric                D24HourMean
-InterdayRollbackMethod       Quadratic
```

RUN CFG

```
-CFGFilename           %CFG%
-ResultsFilename       %RESULTSDIR%\PM25_2004.cfgr
-BaselineAQG           %AQG%\PM25_2004Baseline.aqg
-ControlAQG            %AQG%\PM25_2004Control.aqg
```

RUN APV

```
-APVFilename           %APV%
-ResultsFilename       %RESULTSDIR%\PM25_2004.apvr
-CFGRFilename          %RESULTSDIR%\PM25_2004.cfgr
-IncidenceAggregation  Nation
-ValuationAggregation  Nation
```

GENERATE REPORT APVR

```
-InputFile             %RESULTSDIR%\PM25_2004.apvr
-ReportFile            %REPORTDIR%\PM25_2004_IncidenceNation.csv
-ResultType            PooledIncidence
-CustomFields          "Endpoint Group,Author,Start Age,Endpoint,Qualifier,Pooling
Window"
-ResultFields          "Mean,Standard Deviation,Latin Hypercube Points"
-DecimalDigits          0
```

GENERATE REPORT APVR

```
-InputFile             %RESULTSDIR%\PM25_2004.apvr
-ReportFile            %REPORTDIR%\PM25_2004_ValuationNation.csv
-ResultType             PooledValuation
-CustomFields          "Endpoint Group,Author,Start Age,Endpoint,Qualifier,Pooling
Window"
-ResultFields          "Mean,Standard Deviation,Latin Hypercube Points"
-DecimalDigits         0
```

Appendix L: Function Editor

The function editor is used to develop both health impact functions and valuation functions. This appendix describes the syntax of this editor.

L.1   User Defined Variables

In addition to pre-defined variables that you can select from the **Available Variables** list, you can create your own variables in the C-R Function Editor.

A variable is an identifier whose value can change at runtime. Put differently, a variable is a name for a location in memory; you can use the name to read or write to the memory location. Variables are like containers for data, and, because they are typed, they tell the compiler how to interpret the data they hold.

The basic syntax for a variable declaration is

var identifierList: type;

where identifierList is a comma-delimited list of valid identifiers and type is any valid type. For example,

var I: Integer;

declares a variable I of type Integer, while

var X, Y: Real;

declares two variables--X and Y--of type Real.

Consecutive variable declarations do not have to repeat the reserved word var:

var

     X, Y, Z: Double;

     I, J, K: Integer;

     Digit: 0..9;

     IndicatorName: String;

     Okay: Boolean;

Variables can be initialized at the same time they are declared, using the syntax

var identifier: type = constantExpression;

where constantExpression is any constant expression representing a value of type type. Thus the declaration

var I: Integer = 7;

is equivalent to the declaration and statement

var I: Integer;

 ...

I := 7;

Multiple variable declarations (such as var X, Y, Z: Real;) cannot include initializations, nor can declarations of variant and file-type variables.

L.2　The Script Language

In the C-R Function Editor, you can evaluate complex block of statements.

You can use constructions like:

If...then...else;

for I:= ... to .. do ;

while... do ;

repeat .... until...;

break;

assignment (...:=....;)

try...finally...end;

try...except...end;

Each function you create can be a single statement or a block of statements.

When you specify it as a block of statements, your script must conform to the rules of the script language, as follows:

1. Each single statement must end with a semicolon (;)

2. You can use the following statements:

    variable := *expression*;
    If *logical expression* **then** *statement(s)* [**else** *statement(s)*];
    for variable := from_expression **to/downto** to_*expression* **do** statement(s);
    while logical_expression **do** *statement(s)*;
    repeat *statement(s)* **until** *logical_expression*;
    try *statement(s)* **finally** *statement(s)* **end**;
    try *statement(s)* **except** *statement(s)* **end**;
    inline comments: // *comment... until the end of the line*

nested comments: { *nested comment* }

Statement(s) in the above declarations states that you can specify either a single statement or a block of statements. The block of statements must be enclosed in **begin** ... **end** keywords. It is not necessary to enclose the body of the function in **begin** .. **end**. Cycle statements can use **break** keyword to break the cycle (break must also end with semicolon.)

L.3   Operands

Expressions may contain the following constant and variable types:

Integer numbers;

Floating point numbers;

Scientific numbers;

Decimal separator for all floating point and scientific-format numbers in expressions, is independent of the Regional Settings of Windows and always is a decimal point ('.').

Boolean values - TRUE or FALSE;

Date type values - values of that type must be put in quotes (' '), and also date separator character is independent of the Regional Settings of Windows and always is a slash - /, i.e. - '01/01/2005'

String values - values of that type must be put in double quotes (" "); If a string contains double quotes, you should double them(i.e., "this is a ""string"" ");

L.4   Operations

Arithmetical

+, -, *, /;

div - integer division;

mod - modulo;

^ - power of;

- - negate;

Logical

<, <=, >=, >, <>, =;

and, or, xor, not;

Bitwise

and, or, xor;

~ - negate;

L.5   Arithmetic Functions

ABS(X)	absolute value
SQR(X)	square = X^2 = X*X
SQRT(X)	square root
SIGN(X)	sign of X; =1 for X>0, =0 for X=0, =-1 for X<0
ZERO(X)	=0 for X=0, =1 for X<>0
TRUNC(X)=INT(X)	integer part
FRAC(X)	fractional part
ROUND(X)	rounds X to the nearest integer value
CEIL(X)	always returns "ceil" integer value
FLOOR(X)	always returns "floor" integer value
DEC(X)	decrements a value X by 1 and returns a new value
INC(X)	increments a value X by 1 and returns a new value
ARG(X,Y)	argument(phase) of X and Y
RADIUS(X,Y)	= sqrt(sqr(X)+sqr(Y))
POWER(X,Y)	raises X to a power of Y (Y is a floating point value)
IPOWER(X,Y)	raises X to a power of Y (Y is a integer value)
X ^ Y	raises X to a power of Y (same as above two functions)
EXP(X)	exponent

LN(X)   natural logarithm

LG(X)   decimal logarithm

LOG(X)   base 2 logarithm

SIN(X)   sine

COS(X)   cosine

TAN(X)   tangent

COTAN(X)   cotangent

ASIN(X)   arcsine

ACOS(X)   arccosine

ATAN(X)   arctangent

SINH(X)   hyperbolic sine

COSH(X)   hyperbolic cosine

TANH(X)   hyperbolic tangent

L.6   Aggregate Functions

AVG(X1,X2,...)   returns average value of (unlimited number of) arguments.

MAX(X1,X2,...)   maximum of (unlimited number of) arguments.

MIN(X1,X2,...)   minimum of (unlimited number of) arguments.

SUM(X1,X2,...)   sum of (unlimited number of) arguments.

PROD(X1,X2,..)   product of (unlimited number of) arguments.

References

Abbey, D. E., B. L. Hwang, R. J. Burchette, T. Vancuren and P. K. Mills. 1995a. Estimated long-term ambient concentrations of PM10 and development of respiratory symptoms in a nonsmoking population. Arch Environ Health. Vol. 50 (2): 139-52.

Abbey, D. E., B. E. Ostro, F. Petersen and R. J. Burchette. 1995b. Chronic Respiratory Symptoms Associated with Estimated Long-Term Ambient Concentrations of Fine Particulates Less Than 2.5 Microns in Aerodynamic Diameter (PM2.5) and Other Air Pollutants. J Expo Anal Environ Epidemiol. Vol. 5 (2): 137-159.

Abbey, D. E., F. Petersen, P. K. Mills and W. L. Beeson. 1993. Long-Term Ambient Concentrations of Total Suspended Particulates, Ozone, and Sulfur Dioxide and Respiratory Symptoms in a Nonsmoking Population. Archives of Environmental Health. Vol. 48 (1): 33-46.

Abt Associates Inc. 2000. Final Heavy Duty Engine/Diesel Fuel Rule: Air Quality Estimation, Selected Health and Welfare Benefits Methods, and Benefit Analysis Results. Prepared for U.S. EPA, Office of Air Quality Planning and Standards, Research Triangle Park, NC. Bethesda, MD. December.

Adams, P. F., G. E. Hendershot and M. A. Marano. 1999. Current Estimates from the National Health Interview Survey, 1996. Vital Health Stat. Vol. 10 (200): 1-212.

Agency for Healthcare Research and Quality. 2000. HCUPnet, Healthcare Cost and Utilization Project. Rockville, MD.

American Lung Association. 2002a. Trends in Asthma Morbidity and Mortality. American Lung Association, Best Practices and Program Services, Epidemiology and Statistics Unit.http://www.lungusa.org/data/asthma/ASTHMAdt.pdf.

American Lung Association. 2002b. Trends in Chronic Bronchitis and Emphysema: Morbidity and Mortality. American Lung Association, Best Practices and Program Services, Epidemiology and Statistics Unit. http://www.lungusa.org/data/copd/COPD1.pdf.

American Lung Association. 2002c. Trends in Morbidity and Mortality: Pneumonia, Influenza, and Acute Respiratory Conditions. American Lung Association, Best Practices and Program Services, Epidemiology and Statistics Unit. http://www.lungusa.org/data/pi/PI_1.pdf.

Bell, M. L., F. Dominici and J. M. Samet. 2005. A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. Epidemiology. Vol. 16 (4): 436-45.

Bell, M. L., A. McDermott, S. L. Zeger, J. M. Samet and F. Dominici. 2004. Ozone and short-term mortality in 95 US urban communities, 1987-2000. Jama. Vol. 292 (19): 2372-8.

Belova, A., E. Post and D. McCubbin. 2007. Using Expert Distributions in BenMAP. Memorandum prepared for Bryan Hubbell at U.S. EPA, Office of Air Quality Planning and Standards. Prepared by Abt Associates, Bethesda, MD. March 7.

Blumenschein, K. and M. Johannesson. 1998. Relationship between quality of life instruments, health state utilities, and willingness to pay in patients with asthma. Ann Allergy Asthma Immunol. Vol. 80 (2): 189-94.

Burnett, R. T., M. Smith-Doiron, D. Stieb, M. E. Raizenne, J. R. Brook, R. E. Dales, J. A. Leech, S. Cakmak and D. Krewski. 2001. Association between ozone and hospitalization for acute respiratory diseases in children less than 2 years of age. Am J Epidemiol. Vol. 153 (5): 444-52.

Chen, L., B. L. Jennison, W. Yang and S. T. Omaye. 2000. Elementary school absenteeism and air pollution. Inhal Toxicol. Vol. 12 (11): 997-1016.

Chiang, P. L. 1967. Variance and Covariance of Life Table Functions Estimated from a Sample of Deaths. National Center for Health Statistics. Washington, DC. March.

Collet, D. 1994. Modelling Survival Data in Medical Research. Chapman & Hall: New York.

Crocker, T. D. and R. L. Horst, Jr. 1981. Hours of Work, Labor Productivity, and Environmental Conditions: A Case Study. The Review of Economics and Statistics. Vol. 63: 361-368.

Cropper, M. L. and A. J. Krupnick. 1990. The Social Costs of Chronic Heart and Lung Disease. Resources for the Future. Washington, DC. Discussion Paper QE 89-16-REV.

Decisioneering. 1996. Crystal Ball: Forecasting and Risk Analysis for Spreadsheet Users: User Manual. Version 4.0. Denver, CO. www.decisioneering .com.

DerSimonian, R. and N. Laird. 1986. Meta-Analysis in Clinical Trials. Controlled Clinical Trials. Vol. 7: 177-188.

Dickie, M. and S. Gerking. 1987. Reconciling Averting Behavior and Contingent Valuation Benefit Estimates of Reducing Symptoms of Ozone Exposure (draft), as cited in Neumann, J.E., M. Dickie, and R.E. Unsworth. 1994. Prepared by Industrial Economics. Prepared for Jim DeMocker, U.S. EPA, Office of Air and Radiation. March 31.

Dickie, M. and V. L. Ulery. 2002. Parental Altruism and the Value of Avoiding Acute

Illness: Are Kids Worth More Than Parents? (Paper to be submitted for publication. Presented at Association of Environmental and Resource Economists 2001 Workshop, "Assessing and Managing Environmental and Public Health Risks."). December.

Dockery, D. W., J. Cunningham, A. I. Damokosh, L. M. Neas, J. D. Spengler, P. Koutrakis, J. H. Ware, M. Raizenne and F. E. Speizer. 1996. Health Effects of Acid Aerosols On North American Children - Respiratory Symptoms. Environmental Health Perspectives. Vol. 104 (5): 500-505.

Dockery, D. W., F. E. Speizer, D. O. Stram, J. H. Ware, J. D. Spengler and B. G. Ferris, Jr. 1989. Effects of Inhalable Particles on Respiratory Health of Children. Am Rev Respir Dis. Vol. 139: 587-594.

Eisenstein, E. L., L. K. Shaw, K. J. Anstrom, C. L. Nelson, Z. Hakim, V. Hasselblad and D. B. Mark. 2001. Assessing the clinical and economic burden of coronary artery disease: 1986-1998. Med Care. Vol. 39 (8): 824-35.

ESEERCO. 1994. New York State Environmental Externalities Cost Study. Report 2: Methodology. Empire State Electric Energy Research Corporation.

GeoLytics Inc. 2001. CensusCD® 1990 + Maps, Release 4.1. CD-ROM. GeoLytics, Inc.

Gilliland, F. D., K. Berhane, E. B. Rappaport, D. C. Thomas, E. Avol, W. J. Gauderman, S. J. London, H. G. Margolis, R. McConnell, K. T. Islam and J. M. Peters. 2001. The effects of ambient air pollution on school absenteeism due to respiratory illnesses. Epidemiology. Vol. 12 (1): 43-54.

Greene, W. H. 1997. Econometric Analysis. Prentice Hall: Upper Saddle River, NJ.

Haase, N. 2002. Phone conversation. American Heart Association, October.

Huang, Y., F. Dominici and M. L. Bell. 2005. Bayesian hierarchical distributed lag models for summer ozone exposure and cardio-respiratory mortality. Environmetrics. Vol. 16: 547–562.

Industrial Economics Incorporated (IEc). 1993. Memorandum to Jim DeMocker, U.S. Environmental Protection Agency, Office of Air and Radiation, Office of Policy Analysis and Review. September 30.

Industrial Economics Incorporated (IEc). 1994. Linkage Between Health Effects Estimation and Morbidity Valuation in the Section 812 Analysis -- Draft Valuation Document. Memorandum to Jim DeMocker, U.S. Environmental Protection Agency, Office of Air and Radiation, Office of Policy Analysis and Review. Prepared by J.E. Neumann, M.T. Dickie, and R.E. Unsworth. March 31.

Industrial Economics Incorporated (IEc). 2006. Expanded Expert Judgment Assessment of

the Concentration-Response Relationship Between PM2.5 Exposure and Mortality, Final Report. Prepared for U.S. Environmental Protection Agency, Office of Air and Radiation, Office of Policy Analysis and Review. Prepared by IEc. Cambridge, MA. September 21.

Ito, K. 2003. Associations of Particulate Matter Components with Daily Mortality and Morbidity in Detroit, Michigan. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Health Effects Institute. Boston, MA. May.

Ito, K., S. F. De Leon and M. Lippmann. 2005. Associations between ozone and daily mortality: analysis and meta-analysis. Epidemiology. Vol. 16 (4): 446-57.

Ito, K. and G. D. Thurston. 1996. Daily PM10/mortality associations: an investigations of at-risk subpopulations. Journal of Exposure Analysis and Environmental Epidemiology. Vol. 6 (1): 79-95.

Jaffe, D. H., M. E. Singer and A. A. Rimm. 2003. Air pollution and emergency department visits for asthma among Ohio Medicaid recipients, 1991-1996. Environ Res. Vol. 91 (1): 21-8.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lutkepohl and T.-C. Lee. 1985. The Theory and Practice of Econometrics. John Wiley and Sons: New York.

Kennedy. 1990. A Guide to Econometrics. MIT Press: Cambridge, MA.

Krupnick, A. J. and M. L. Cropper. 1992. The Effect of Information On Health Risk Valuations. Journal of Risk and Uncertainty. Vol. 5 (1): 29-48.

Laden, F., J. Schwartz, F. E. Speizer and D. W. Dockery. 2006. Reduction in Fine Particulate Air Pollution and Mortality: Extended follow-up of the Harvard Six Cities Study. Am J Respir Crit Care Med. Vol. 173 (6): 667-72.

Levy, J. I., S. M. Chemerynski and J. A. Sarnat. 2005. Ozone exposure and mortality: an empiric bayes metaregression analysis. Epidemiology. Vol. 16 (4): 458-68.

Lippmann, M., K. Ito, A. Nádas and R. Burnett. 2000. Association of Particulate Matter Components with Daily Mortality and Morbidity in Urban Populations. Health Effects Institute. Number 95. August.

Loehman, E. T., S. V. Berg, A. A. Arroyo, R. A. Hedinger, J. M. Schwartz, M. E. Shaw, R. W. Fahien, V. H. De, R. P. Fishe, D. E. Rio, W. F. Rossley and A. E. S. Green. 1979. Distributional Analysis of Regional Benefits and Cost of Air Quality Control. Journal of Environmental Economics and Management. Vol. 6: 222-243.

McConnell, R., K. Berhane, F. Gilliland, S. J. London, H. Vora, E. Avol, W. J. Gauderman, H. G. Margolis, F. Lurmann, D. C. Thomas and J. M. Peters. 1999. Air

pollution and bronchitic symptoms in Southern California children with asthma. Environ Health Perspect. Vol. 107 (9): 757-60.

Moolgavkar, S. H. 2000a. Air Pollution and Hospital Admissions for Chronic Obstructive Pulmonary Disease in Three Metropolitan Areas in the United States. Inhalation Toxicology. Vol. 12 (Supplement 4): 75-90.

Moolgavkar, S. H. 2000b. Air pollution and hospital admissions for diseases of the circulatory system in three U.S. metropolitan areas. J Air Waste Manag Assoc. Vol. 50 (7): 1199-206.

Moolgavkar, S. H. 2003. Air Pollution and Daily Deaths and Hospital Admissions in Los Angeles and Cook Counties. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Health Effects Institute. Boston, MA. May.

Moolgavkar, S. H., E. G. Luebeck and E. L. Anderson. 1997. Air pollution and hospital admissions for respiratory causes in Minneapolis St. Paul and Birmingham. Epidemiology. Vol. 8 (4): 364-370.

Moolgavkar, S. H., E. G. Luebeck, T. A. Hall and E. L. Anderson. 1995. Air Pollution and Daily Mortality in Philadelphia. Epidemiology. Vol. 6 (5): 476-484.

Mrozek, J. R. and L. O. Taylor. 2002. What Determines the Value of Life? A Meta-Analysis. Journal of Policy Analysis and Management. Vol. 21: 253-270.

National Center for Health Statistics. 1999. National Vital Statistics Reports. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics. Washington, DC. Volume 47, Number 19. June 30.

Nelder, J. A. and R. Mead. 1965. A simplex algorithm for function minimization. Computer Journal. Vol. 7: 308-313.

Norris, G., S. N. YoungPong, J. Q. Koenig, T. V. Larson, L. Sheppard and J. W. Stout. 1999. An association between fine particles and asthma emergency department visits for children in Seattle. Environ Health Perspect. Vol. 107 (6): 489-93.

O'Connor, R. M. and G. C. Blomquist. 1997. Measurement of Consumer-Patient Preferences Using a Hybrid Contingent Valuation Method. Journal of Health Economics. Vol. 16: 667-683.

Ostro, B., M. Lipsett, J. Mann, H. Braxton-Owens and M. White. 2001. Air pollution and exacerbation of asthma in African-American children in Los Angeles. Epidemiology. Vol. 12 (2): 200-8.

Ostro, B. D. 1987. Air Pollution and Morbidity Revisited: A Specification Test. Journal of

Environmental Economics and Management. Vol. 14: 87-98.

Ostro, B. D., M. J. Lipsett, M. B. Wiener and J. C. Selner. 1991. Asthmatic Responses to Airborne Acid Aerosols. Am J Public Health. Vol. 81 (6): 694-702.

Ostro, B. D. and S. Rothschild. 1989. Air Pollution and Acute Respiratory Morbidity - an Observational Study of Multiple Pollutants. Environ Res. Vol. 50 (2): 238-247.

Owings, M. F. and L. Lawrence. 1999. Detailed Diagnoses and Procedures, National Hospital Discharge Survey, 1997. National Center for Health Statistics. Hyattsville, MD. Vital Health Statistics, Series 13, No. 145. December.

Peel, J. L., P. E. Tolbert, M. Klein, K. B. Metzger, W. D. Flanders, K. Todd, J. A. Mulholland, P. B. Ryan and H. Frumkin. 2005. Ambient air pollution and respiratory emergency department visits. Epidemiology. Vol. 16 (2): 164-74.

Peters, A., D. W. Dockery, J. E. Muller and M. A. Mittleman. 2001. Increased particulate air pollution and the triggering of myocardial infarction. Circulation. Vol. 103 (23): 2810-5. http://www.circulationaha.org/cgi/content/full/103/23/2810.

Pope, C. A., 3rd, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito and G. D. Thurston. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Jama. Vol. 287 (9): 1132-41.

Pope, C. A., D. W. Dockery, J. D. Spengler and M. E. Raizenne. 1991. Respiratory Health and Pm10 Pollution - a Daily Time Series Analysis. American Review of Respiratory Disease. Vol. 144 (3): 668-674.

Popovic, J. R. 2001. 1999 National Hospital Discharge Survey: annual summary with detailed diagnosis and procedure data. Vital Health Stat 13. (151): i-v, 1-206.

Ransom, M. R. and C. A. Pope. 1992. Elementary school absences and PM10 pollution in Utah Valley. Environ Res. Vol. 58 (2): 204-19.

Rosamond, W., G. Broda, E. Kawalec, S. Rywik, A. Pajak, L. Cooper and L. Chambless. 1999. Comparison of medical care and survival of hospitalized patients with acute myocardial infarction in Poland and the United States. Am J Cardiol. Vol. 83 (8): 1180-5.

Rowe, R. D. and L. G. Chestnust. 1986. Oxidants and Asthmatics in Los Angeles: A Benefits Analysis -- Executive Summary. Prepared for U.S. Environmental Protection Agency, Office of Policy Analysis. Prepared by Energy and Resource Consultants, Inc. Washington, DC. EPA-230-09-86-018. March.

Russell, M. W., D. M. Huse, S. Drowns, E. C. Hamel and S. C. Hartz. 1998. Direct medical costs of coronary artery disease in the United States. Am J Cardiol. Vol. 81 (9):

1110-5.

Samet, J. M., S. L. Zeger, J. E. Kelsall, J. Xu and L. S. Kalkstein. 1997. Air Pollution, Weather, and Mortality in Philadelphia 1973-1988. Health Effects Institute. Cambridge, MA. March.

Schwartz, J. 1994a. Air Pollution and Hospital Admissions For the Elderly in Detroit, Michigan. American Journal of Respiratory and Critical Care Medicine. Vol. 150 (3): 648-655.

Schwartz, J. 1994b. PM(10) Ozone, and Hospital Admissions For the Elderly in Minneapolis St Paul, Minnesota. Archives of Environmental Health. Vol. 49 (5): 366-374.

Schwartz, J. 1995. Short term fluctuations in air pollution and hospital admissions of the elderly for respiratory disease. Thorax. Vol. 50 (5): 531-538.

Schwartz, J. 2005. How sensitive is the association between ozone and daily deaths to control for temperature? Am J Respir Crit Care Med. Vol. 171 (6): 627-31.

Schwartz, J., D. W. Dockery, L. M. Neas, D. Wypij, J. H. Ware, J. D. Spengler, P. Koutrakis, F. E. Speizer and B. G. Ferris. 1994. Acute Effects of Summer Air Pollution On Respiratory Symptom Reporting in Children. Am J Respir Crit Care Med. Vol. 150 (5): 1234-1242.

Schwartz, J. and L. M. Neas. 2000. Fine particles are more strongly associated than coarse particles with acute respiratory health effects in schoolchildren. Epidemiology. Vol. 11 (1): 6-10.

Sheppard, L. 2003. Ambient Air Pollution and Nonelderly Asthma Hospital Admissions in Seattle, Washington, 1987-1994. In: Revised Analyses of Time-Series Studies of Air Pollution and Health. Health Effects Institute. Boston, MA. May.

Sheppard, L., D. Levy, G. Norris, T. V. Larson and J. Q. Koenig. 1999. Effects of ambient air pollution on nonelderly asthma hospital admissions in Seattle, Washington, 1987-1994. Epidemiology. Vol. 10 (1): 23-30.

Smith, D. H., D. C. Malone, K. A. Lawson, L. J. Okamoto, C. Battista and W. B. Saunders. 1997. A national estimate of the economic costs of asthma. Am J Respir Crit Care Med. Vol. 156 (3 Pt 1): 787-93.

Stanford, R., T. McLaughlin and L. J. Okamoto. 1999. The cost of asthma in the emergency department and hospital. Am J Respir Crit Care Med. Vol. 160 (1): 211-5.

Stieb, D. M., R. T. Burnett, R. C. Beveridge and J. R. Brook. 1996. Association between ozone and asthma emergency department visits in Saint John, New Brunswick, Canada.

Environmental Health Perspectives. Vol. 104 (12): 1354-1360.

Tolley, G. S. and et al. 1986. Valuation of Reductions in Human Health Symptoms and Risks. Prepared for U.S. Environmental Protection Agency. January.

U.S. Bureau of the Census. 1997. Statistical Abstract of the United States: 1997. Washington, DC.

U.S. Bureau of the Census. 2002a. Modified Race Data Summary File 2000 Technical Documentation. Census of Population and Housing: Washington DC. http://www.census.gov/popest/archives/files/MRSF-01-US1.pdf.

U.S. Bureau of the Census. 2002b. Statistical Abstract of the United States: 2001. Washington DC.

U.S. Department of Education. 1996. The Condition of Education 1996, Indicator 42: Student Absenteeism and Tardiness. National Center for Education Statistics. Washington DC.

U.S. EPA. 1997. The Benefits and Costs of the Clean Air Act: 1970 to 1990. U.S. EPA, Office of Air and Radiation, Office of Policy, Planning and Evaluation. Washington, DC. EPA 410-R-97-002. October. http://www.epa.gov/airprogm/oar/sect812/index.html .

U.S. EPA. 1999a. The Benefits and Costs of the Clean Air Act: 1990 to 2010: EPA Report to Congress. U.S. EPA, Office of Air and Radiation, Office of Policy. Washington, DC. EPA 410-R-99-001. November. http://www.epa.gov/airprogm/oar/sect812/index.html.

U.S. EPA. 1999b. OAQPS Economic Analysis Resource Document. Prepared by U.S. EPA, Office of Air Quality Planning and Standards, Innovative Strategies and Economics Group. Durham, NC. April. http://www.epa.gov/ttn/ecas/analguid.html.

U.S. EPA. 2000. Guidelines for Preparing Economic Analyses. Office of the Administrator. Washington, DC. EPA #: 240-R-00-003. September. http://yosemite.epa.gov/ee/epa/eed.nsf/webpages/Guidelines.html/$file/Guidelines.pdf.

U.S. EPA. 2004. Final Regulatory Analysis: Control of Emissions from Nonroad Diesel Engine. U.S. EPA, Office of Transportation and Air Quality. Washington, DC. EPA 420-R-04-007. April.

U.S. EPA. 2006. Final Regulatory Impact Analysis: PM2.5 NAAQS. Office of Air and Radiation, Office of Air Quality Planning and Standards. Research Triangle Park, NC. http://www.epa.gov/ttn/ecas/ria.html.

U.S. EPA. 2008a. Final Ozone NAAQS Regulatory Impact Analysis. Prepared by Office of Air and Radiation, Office of Air Quality Planning and Standards. Research Triangle

Park, NC. http://www.epa.gov/ttn/ecas/ria.html.

U.S. EPA. 2008b. Regulatory Impact Analysis: Control of Emissions of Air Pollution from Locomotive Engines and Marine Compression Ignition Engines Less than 30 Liters Per Cylinder. Office of Transportation and Air Quality, Assesssment and Standards Division. Washington, DC. EPA420-R-08-001. March. http://www.epa.gov/otaq/locomotv.htm.

Vedal, S. 1998. Lung health and the environment in developing countries. International Journal of Tuberculosis and Lung Disease. Vol. 2 (9 SUPP1): S71-S76.

Viscusi, W. K. 1992. Fatal Tradeoffs: Public and Private Responsibilities for Risk. Oxford University Press: New York.

Viscusi, W. K. and J. E. Aldy. 2003. The Value of a Statistical LIfe: A Critical Review of Market Estimates throughout the World. AEI-Brookings Joint Center for Regulatory Studies. Washington, DC. January.

Viscusi, W. K., W. A. Magat and J. Huber. 1991. Pricing Environmental Health Risks - Survey Assessments of Risk - Risk and Risk - Dollar Trade-Offs For Chronic Bronchitis. Journal of Environmental Economics and Management. Vol. 21 (1): 32-51.

Wilson, A. M., C. P. Wake, T. Kelly and J. C. Salloway. 2005. Air pollution, weather, and respiratory emergency room visits in two northern New England cities: an ecological time-series study. Environ Res. Vol. 97 (3): 312-21.

Wittels, E. H., J. W. Hay and A. M. Gotto, Jr. 1990. Medical costs of coronary artery disease in the United States. Am J Cardiol. Vol. 65 (7): 432-40.

Woodruff, T. J., J. Grillo and K. C. Schoendorf. 1997. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. Environmental Health Perspectives. Vol. 105 (6): 608-612.

Woodruff, T. J., J. D. Parker and K. C. Schoendorf. 2006. Fine particulate matter (PM2.5) air pollution and selected causes of postneonatal infant mortality in California. Environmental Health Perspectives. Vol. 114: 786–790.

Woods & Poole Economics Inc. 2007. Complete Demographic Database. Washington, DC. http://www.woodsandpoole.com/index.php.