Environmental Benefits Mapping and Analysis Program – Community Edition

User's Manual

July 2018

Updated for BenMAP-CE Version 1.4.8

Master Table of Contents

Chapter 1
Welcome to BenMAP-CE

In this chapter, find...

- **An overview of the tool.**
- **Instructions for installing the tool.**
- **Contacts, sources for more information, and answers to frequently asked general questions.**

Chapter 1 Table of Contents

The environmental Benefits Mapping and Analysis Program—Community Edition (BenMAP-CE) is a powerful, yet easy-to-use program that estimates the number and economic value of health impacts resulting from changes in air pollution concentrations. The open-source BenMAP-CE tool replaces the proprietary version of the program (BenMAP) that the U.S. Environmental Protection Agency (U.S. EPA) first developed in 2003 to analyze national-scale air quality policies. These analyses include health benefits assessments for the National Ambient Air Quality Standards (NAAQs) for Particulate Matter (2006, 2012) and Ozone (2008, 2010) as well as the Locomotive Marine Engine Rule (2008).

U.S. EPA and its partners designed BenMAP-CE to serve the analytical needs of a range of users, including scientists, policy analysts, and decision makers. Most users apply the BenMAP-CE tool to answer one of two types of questions:

1. What are the human health and economic benefits associated with a policy improving air quality?

2. What is the human health burden attributable to total air pollution levels?

While the BenMAP-CE development team designed the program to be accessible to novice users, the tool includes a number of features that will appeal to advanced analysts as well. For example, analysts can add their own health impact and valuation functions, map results, and perform a suite of sensitivity analyses. Beginning users can take advantage of U.S. EPA's pre-programmed settings and reports in the core program.

1.1   Overview of BenMAP-CE & Benefits Assessment

The BenMAP-CE program estimates the human health impacts and economic value of air quality changes. That is — BenMAP-CE relates air quality changes to human health benefits. Such analyses are a critical component of air quality policy assessments. As such, a variety of Federal, State and Local air pollution officials have used BenMAP-CE to inform air quality management decisions.[1]

BenMAP-CE estimates benefits from improvements in human health, such as reductions in the risk of premature death, heart attacks, and other adverse health effects. Other benefits of reducing air pollution (i.e., visibility and ecosystem effects) are not quantified in the current version of BenMAP-CE. After estimating the reductions in the incidence of adverse health effects, BenMAP-CE calculates the monetary benefits associated with those reductions.

How does BenMAP-CE estimate human health effects?

First, BenMAP-CE determines the change in ambient air pollution using user-specified air quality data. Because BenMAP-CE does not model air quality changes, these data

---

[1] For a list of peer-reviewed articles that used the BenMAP and BenMAP-CE tools, see: www.epa.gov/benmap

must be input into BenMAP-CE as modeling data or generated from air pollution monitoring data (although some monitoring data is pre-loaded in BenMAP-CE, see Chapter 5 for details). Next, BenMAP-CE relates the change in pollution concentration with certain health effects (also known as health endpoints, see Chapter 6 for details). This relationship is often referred to as the health impact function (HIF) or the concentration-response (C-R) function. As shown in Figure 1-1, these HIFs are derived from epidemiology studies that relate pollutant concentrations with health outcomes. BenMAP-CE applies that relationship to the population experiencing the change in pollution exposure to calculate health impacts.
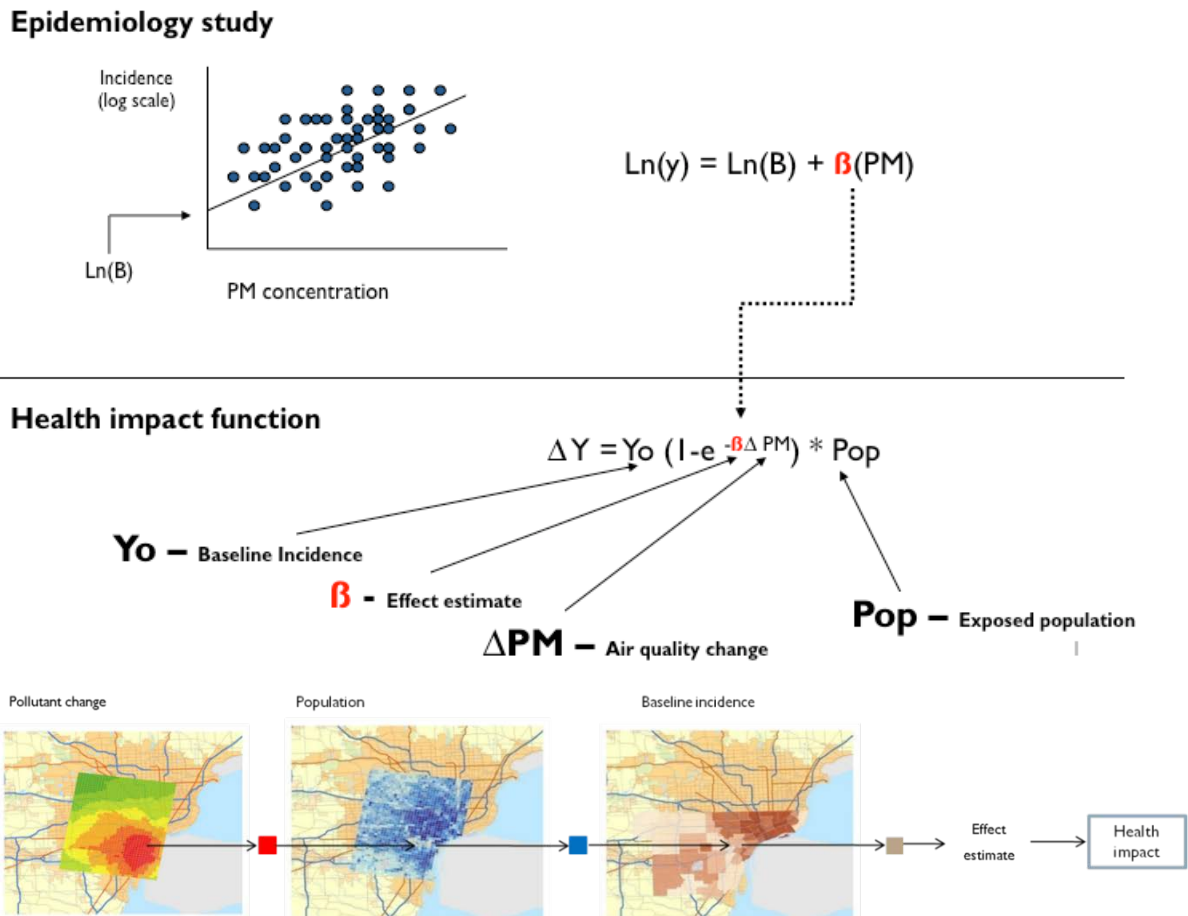
Figure 1-1. Deriving a Health Impact Function from the Epidemiology Literature

A simplified example is shown below.

Health Effect = Air Quality Change * Health Effect Estimate * Exposed Population * Health Baseline Incidence

- **Air Quality Change**. The air quality change is the difference between the starting air pollution level (i.e., the baseline) and the air pollution level after some change, such as a new regulation (i.e., the control).

- **Health Effect Estimate**. The health effect estimate is an estimate of the percentage change in the risk of an adverse health effect due to a one unit change in ambient air pollution. Epidemiological studies are a good source for effect estimates.

- **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.

- **Health Baseline Incidence**. The health incidence rate is an estimate of the average number of people who die (or suffer from some adverse health effect) in a given population over a given period of time. For example, the health incidence rate might be the probability that a person will die in a given year. Health incidence rates and other health data are typically collected by the government. In addition the World Health Organization is a good source for this.[2]

How does BenMAP-CE estimate the economic value of human health effects?

BenMAP-CE also calculates the economic value of avoided health effects (see Chapter 7 for details). After calculating the health changes, you can estimate the economic value by multiplying the reduction of the health effect by an estimate of the economic value per case (see Figure 1-2):

Economic Value = Health Effect * Value of Health Effect

---

[2] The World Health Organization is a good source for international health data, see: http://www.who.int.
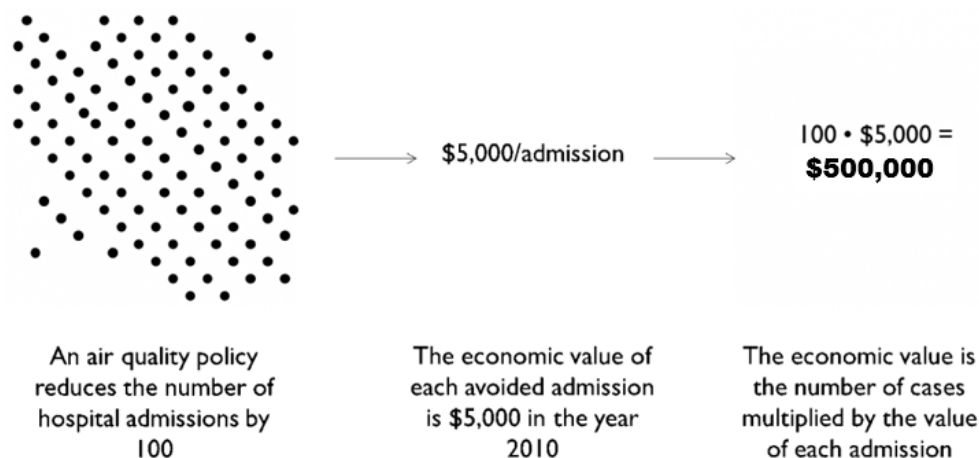
Figure 1-2. Estimating the Economic Value of Human Health Effects

There are several different ways of calculating the value of the health effect. For example, the value of an avoided premature mortality is generally calculated using the Value of Statistical Life (VSL). The value of a statistical life is the monetary value that a group of people are willing to pay to slightly reduce the risk of premature death in the population. For other health effects, the medical costs of the illness may be the only valuation data available. The BenMAP-CE database includes several different functions for VSL and valuation functions for other health effects for you to choose, or you can use the U.S. EPA's approach for quantifying and valuing air pollution effects[3].

Figure 1-3 summarizes the BenMAP-CE inputs and outputs. This figure shows the types of choices that you make regarding the modeling of population exposure, the types of health effects to model, and how to place an economic value on these health effects. Please note that BenMAP-CE does not have air quality modeling capabilities, and therefore the user must provide externally created data in order to work with modeled data. BenMAP-CE is preloaded with limited air quality monitoring data, but externally created monitoring data may also be needed.

What else can BenMAP-CE do?

BenMAP-CE incorporates a geographic information system (GIS), allowing users to create, utilize, visualize, and export maps of air pollution, population, incidence rates, incidence rate changes, economic valuations, and other types of data (Figure 1-4).

---

[3] See https://www.epa.gov/benmap/benmap-community-edition.
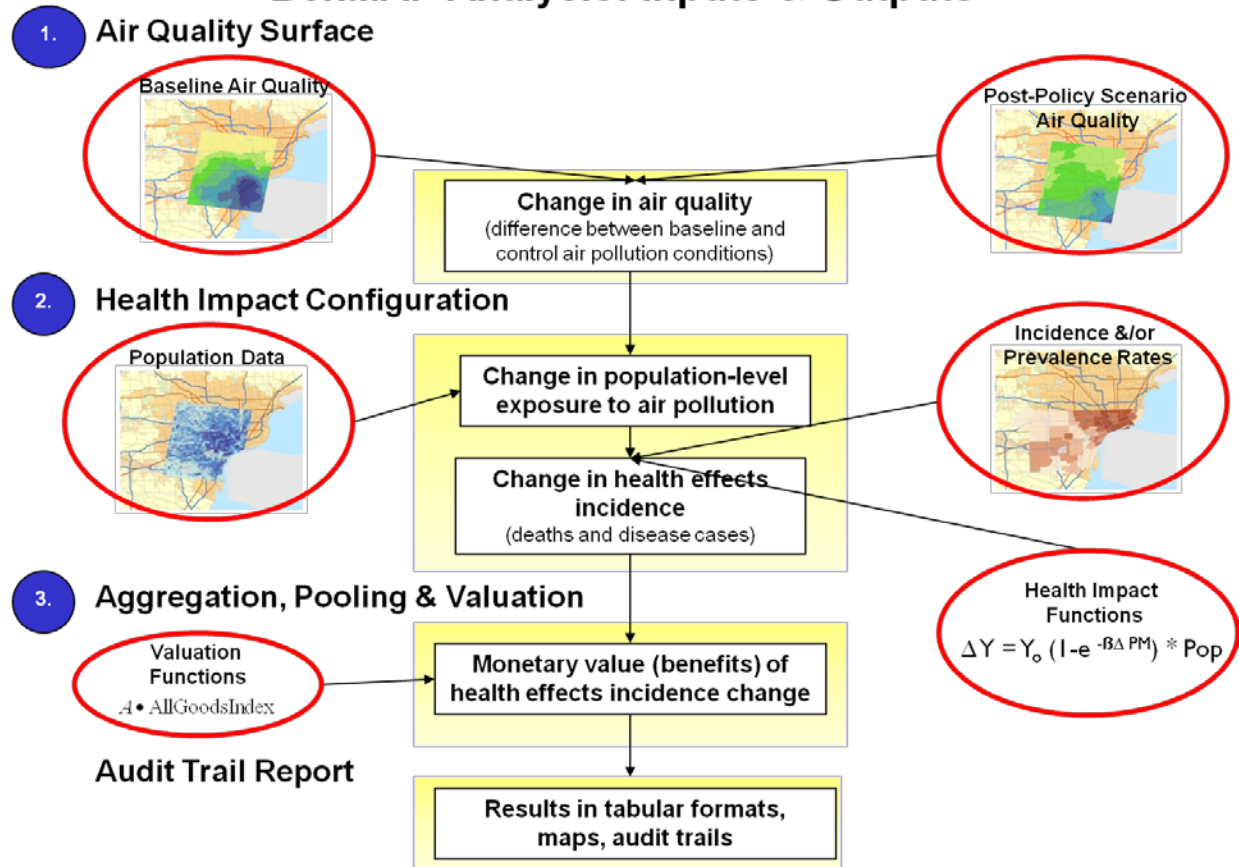
Figure 1-3. BenMAP-CE Flow Diagram
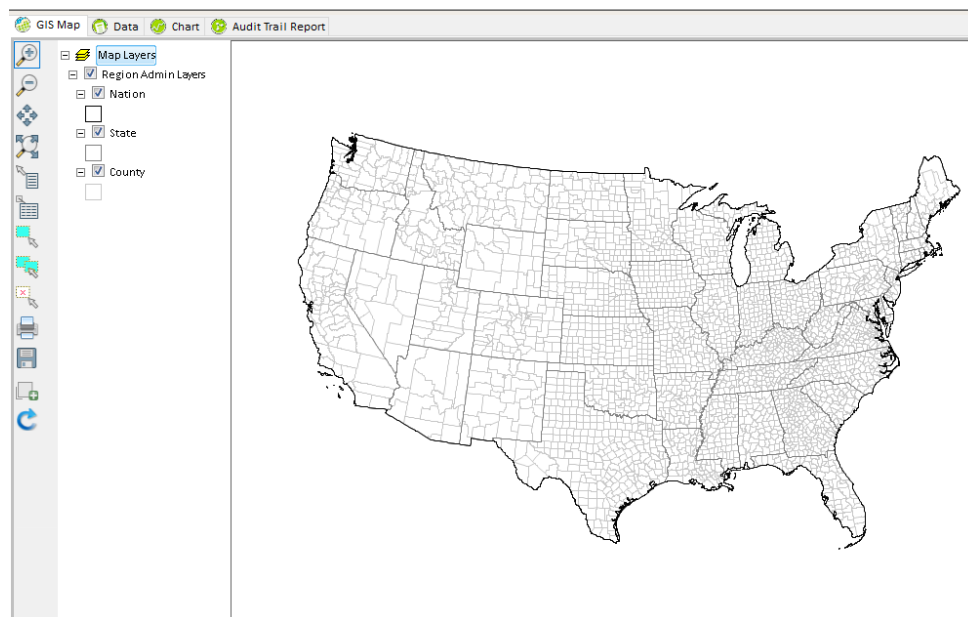
Figure 1-4. BenMAP-CE GIS Example

Analysts can use BenMAP-CE to:

- Create maps illustrating the population/community level ambient pollution levels;

- Compare benefits associated with various regulatory programs;

- Characterize the distribution of health impacts among population sub-groups;

- Estimate health impacts and economic values of existing air pollution concentrations;

- Estimate the health benefits of alternative ambient air quality standards; and

- Perform sensitivity analyses of health or valuation functions, or of other inputs.

1.2 How to Use this Manual

Chapters 2 through 9 of this manual provide step-by-step instructions on how to use BenMAP-CE. New users should start with Chapters 2 and 3, which are both relatively short. These chapters provide a basic overview of the tool and how it works, and explain some potentially confusing terminology.[4] Use the rest of the manual to answer any specific questions you may have, or to walk you step-by-step through the various components. Chapter 4 discusses how to enter data into BenMAP-CE, Chapters 5 through 7 cover each of the main steps in the Core Program, and Chapters 8 and 9 cover mapping, report options, and additional tools.

Each chapter is introduced by a short section that describes what you can find within the chapter and provides an outline of the chapter's contents. This is a good place to go if the Table of Contents does not provide enough detail for you to find the section you need. The end of most chapters has a series of "Frequently Asked Questions," which may also be helpful for answering specific questions. In chapters that provide instructions on navigating the tool, the following conventions are observed: tree menu items, buttons, tabs and selection box labels are in bold type; prompts and messages are enclosed in quotation marks; and drop-down menu items, options to click or check, and items that need to be filled in or selected by the user are italicized. Throughout the chapters you will also see boxes presenting common mistakes and important things to remember when working with BenMAP-CE.

There is also a set of Technical Appendices to provide more detailed information on model functions, data, and underlying assumptions.

Appendix A: Monitor Rollback Algorithms
Appendix B: Air Pollution Exposure Estimation Algorithms

---

[4] Another good reference is the BenMAP-CE Quick Start Guide, see: http://www.epa.gov/benmap.

1.3　Computer Requirements

The computer hardware requirements for BenMAP-CE are typically modest, though this will vary depending on the complexity of the analysis. BenMAP-CE requires a Windows platform and can be used on machines running Windows 7, Windows 8 or Windows 10. In particular, BenMAP-CE requires a computer with:

- Either a 64- or 32-bit operating system, although a 64-bit operating system is recommended

- Adobe Acrobat Reader

- Microsoft Excel or other spreadsheet program (in order to read exported .xlsx files)[5]

- Microsoft .NET Framework 4 (or 4.5)[6]

- At least 4 gigabytes of RAM[7]

- Intel or compatible processor, Core i5 (or better)

---

[5] OpenOffice and LibreOffice are two open-source options for spreadsheet tools.

[6] If .NET is not pre-installed, BenMAP-CE will provide a message advising you to install .NET. A standalone installer is available on the Microsoft website (URL: http://www.microsoft.com/en-us/download/details.aspx?id=17718). Install the runtime version and associated files (e.g., dotnetfx40_full_x86_x64.exe).

[7] BenMAP-CE works best in a 64-bit Windows environment. With a 32-bit installation there are limits on the memory available to the software application; it can utilize no more than 2 GB of RAM. This will impact performance when processing large spatial datasets or numerous health impact/valuation functions. To determine whether your computer is running a 32-bit or 64-bit version of Windows, refer to this article from Microsoft: https://support.microsoft.com/en-us/help/15056/windows-7-32-64-bit-faq

---

- At least 10 GB free disk space is necessary for installation of the BenMAP-CE database and ancillary files[8].

1.4   Installing BenMAP-CE

The installation of BenMAP-CE is very simple. Double click **Setup.exe** in your installation directory to bring up the setup wizard. Then follow the setup wizard by clicking 'Next' or 'OK' to complete the installation.

---
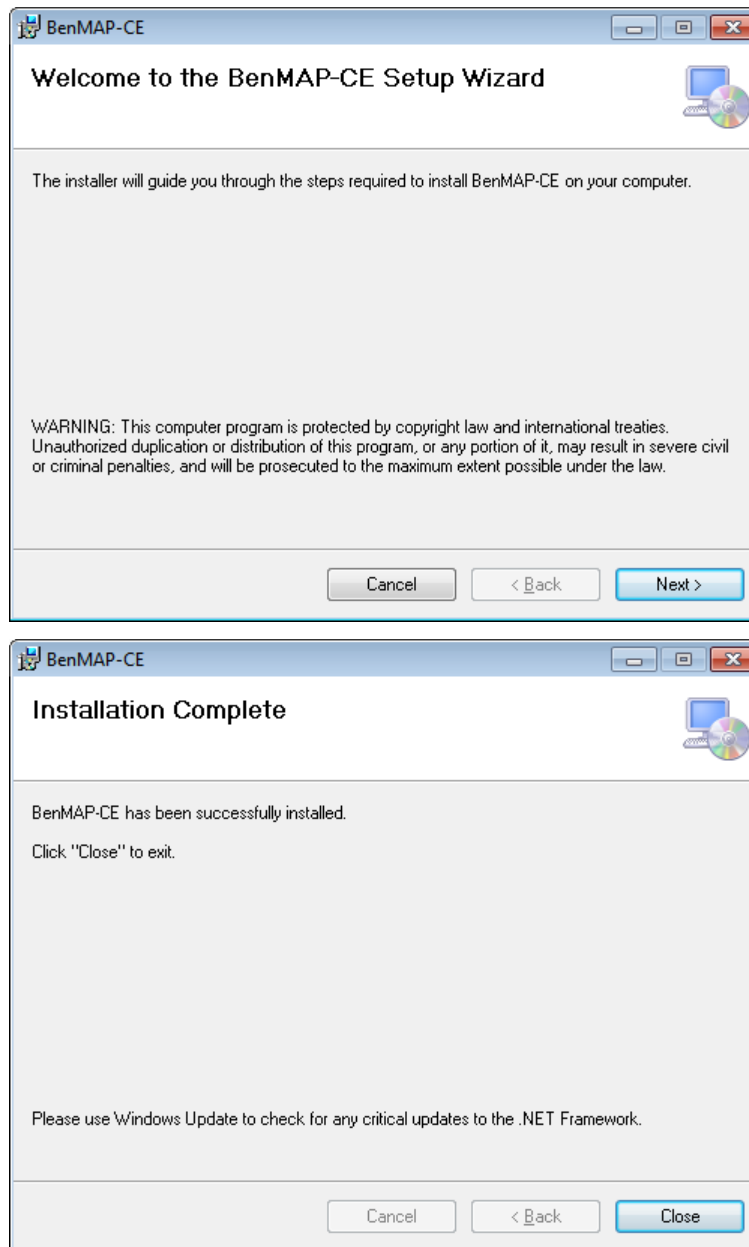
8 A solid state drive (SSD) has also shown improved performance over hard disk drives (HDD).

Upgrading to a new version of BenMAP-CE

Periodically, new versions of BenMAP-CE will be made available and posted to the BenMAP-CE website: http://www.epa.gov/benmap/. If you are upgrading, first uninstall the previous version of the software on your computer (details provided in Section 1.5). Then extract the files from the installer package (.zip file) and run the executable ("*setup.exe*"). A new installer package can be relatively large (~1.2 gigabytes) because the database is embedded in the installer. However, once downloaded, the installation process is generally very fast.

Installation instructions are typically provided with each software release. Refer to these supplemental instructions for important additional information. Please note that your existing Setups will not be automatically transferred to newly installed versions of BenMAP. See instructions for exporting and importing databases in Chapter 9.

1.5   Uninstalling BenMAP-CE

To uninstall BenMAP-CE, go to **Control Panel, Programs and Features** and remove **BenMAP-CE.** Note that uninstalling BenMAP-CE does not also remove any results files that you have created with BenMAP-CE.

1.6   Contacts for Comments, Questions & Bug Reporting

For comments and questions, please contact Ali Kamal at the U.S. EPA.

Address: C539-07, U.S. EPA Mailroom, Research Triangle Park, NC 27711

Email: kamal.ali@epa.gov

Telephone: 919-541-4959

Alternatively, you can send a message at the BenMAP-CE website: https://www.epa.gov/benmap/forms/contact-us-about-benmap, or by simply emailing benmap@epa.gov.

To report programming bugs or suggest additions to the software in BenMAP-CE:

- Select the **Help** menu in the main window;

- Open the **Provide Feedback** form;

- Complete the form and submit the report.

1.7   Sources for More Information

For supplemental information on BenMAP-CE, such as articles and presentations, manuals, and training materials go to:

- U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards (OAQPS), BenMAP-CE website: https://www.epa.gov/benmap/

To interact with other BenMAP users, go to:

- BenMAP-CE Discussion Forum: https://forum.benmap.org/

To access the BenMAP-CE source code, go to:

- BenMAP-CE on Github: https://github.com/BenMAPCE/BenMAP-CE

For more information on conducting benefit analysis, see the following documents:

- U.S. EPA (various years). Costs and Benefits of the Clean Air Act. Available at: https://www.epa.gov/clean-air-act-overview/benefits-and-costs-clean-air-act

- U.S. EPA (2006). Final Regulatory Impact Analysis: 2006 National Ambient Air Quality Standards for Particulate Matter. Office of Air Quality Planning and Standards. *See*: Chapter 5. Available at: http://www.epa.gov/ttn/ecas/regdata/RIAs/Chapter%205--Benefits.pdf

- U.S. EPA (2008). Final Ozone NAAQS Regulatory Impact Analysis. Office of Air Quality Planning and Standards. March. *See:* Chapter 6. Available at: https://www3.epa.gov/ttn/ecas/regdata/RIAs/6-ozoneriachapter6.pdf

- U.S. EPA (2008). 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. EPA420-R-08-001a. May. *See*: Chapter 6. Available at: https://nepis.epa.gov/Exe/ZyPDF.cgi/P10024CN.PDF?Dockey=P10024CN.PDF

- U.S. EPA (2010). Final Regulatory Impact Analysis (RIA) for the SO2 National Ambient Air Quality Standards (NAAQS). Office of Air Quality Planning and Standards. June. *See:* Chapter 5. Available at: http://www.epa.gov/ttn/ecas/regdata/RIAs/fso2ria100602ch5.pdf

- U.S. EPA (2010). Final Regulatory Impact Analysis (RIA) for the NO2 National Ambient Air Quality Standards (NAAQS). Office of Air Quality Planning and Standards. June. *See:* Chapter 4. Available at: https://www3.epa.gov/ttn/ecas/docs/ria/naaqs-no2_ria_final_2010-01.pdf

- U.S. EPA (2010). Guidelines for Preparing Economic Analyses. Office of the Administrator, National Center for Environmental Economics. EPA 240-R-10-001. December. Available at: https://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0568-50.pdf/$file/EE-0568-50.pdf

- U.S. EPA (2011). Regulatory Impact Analysis (RIA) for the Final Transport Rule. Office of Air and Radiation. June. See: Chapter 5. Available at: http://www.epa.gov/airtransport/pdfs/FinalRIA.pdf

- U.S. EPA (2011). Regulatory Impact Analysis for the Final Mercury and Air Toxics Standards. Office of Air Quality Planning and Standards. December. See: Chapter 5. Available at: http://www.epa.gov/ttnecas1/regdata/RIAs/matsriafinal.pdf

- U.S. EPA (2012). Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. *See*: Chapter 5. Available at: http://www.epa.gov/ttnecas1/regdata/RIAs/finalria.pdf

1.8   Frequently Asked Questions (General)

Is BenMAP-CE free? Is there a Terms of Use agreement? Are there any restrictions on using BenMAP-CE?

BenMAP-CE is free. There is no Terms of Use agreement and there are no restrictions on using BenMAP-CE. Feel free to share it with others.

How do I know which version of BenMAP-CE I am using? How do I know if I have the most current version of BenMAP-CE? How do I get the most current version?

You can identify the version of BenMAP-CE you are using by going to the Help menu and choosing **About**. Here you will see the version number and contact information. To determine whether you have the most recent version of BenMAP-CE, you can check the BenMAP-CE website (http://www.epa.gov/benmap/), which will have the latest version that is publicly available. Alternatively, you can use the contact information to inquire about any upcoming versions of the model.
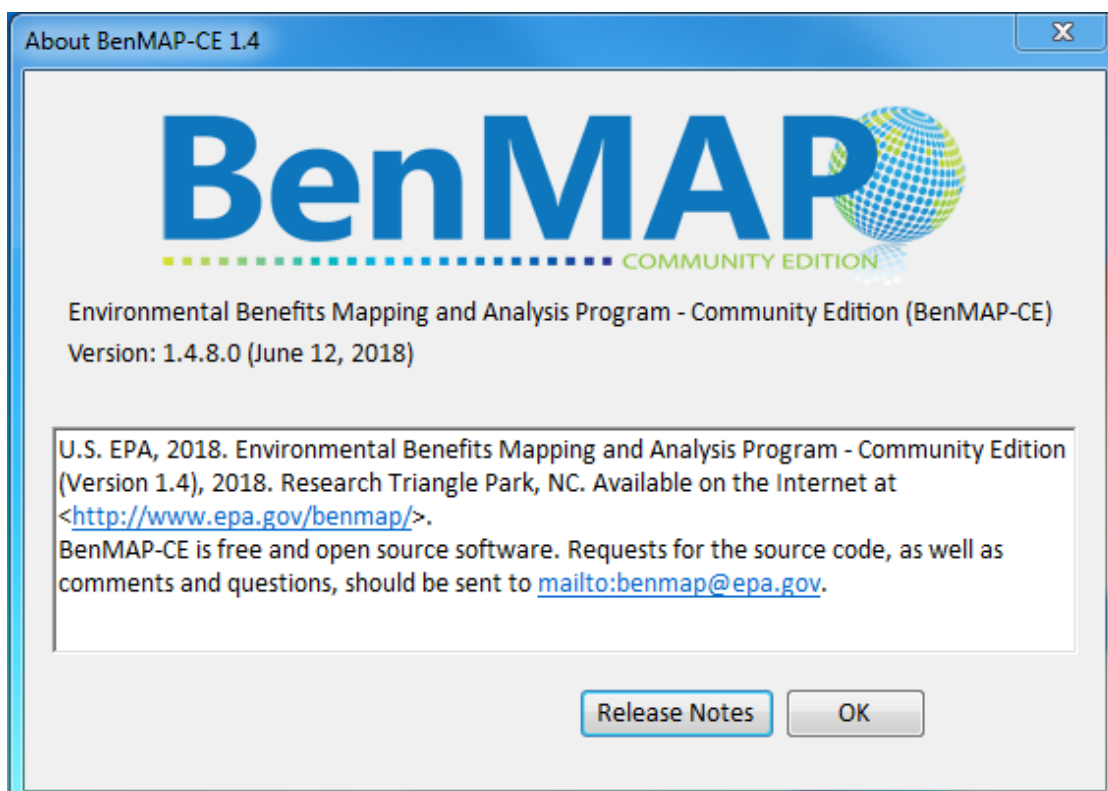
Why don't my files created with an older version of BenMAP work with BenMAP-CE?

Files created with an older version of BenMAP will not, in most cases, work with BenMAP-CE because of changes to the program. For example, later versions of BenMAP-CE have the capability to handle population data differentiated by ethnicity.

For this reason, after completing an analysis with BenMAP-CE, it is always good to archive the BenMAP-CE installer along with the files used in your analysis, so that you will always be able to reproduce your work in the future.

Why are my pop-up windows too small? Why are buttons missing?

If the BenMAP-CE pop-up windows do not show the entire content (display seems cut off or buttons are missing), please check the display properties for your computer. Locate your **Control Panel**, then your **Display** settings, and choose the screen resolution associated with **96 DPI** (display pixels per inch). Here is what the screen looks like for Windows 7 Professional operating system.
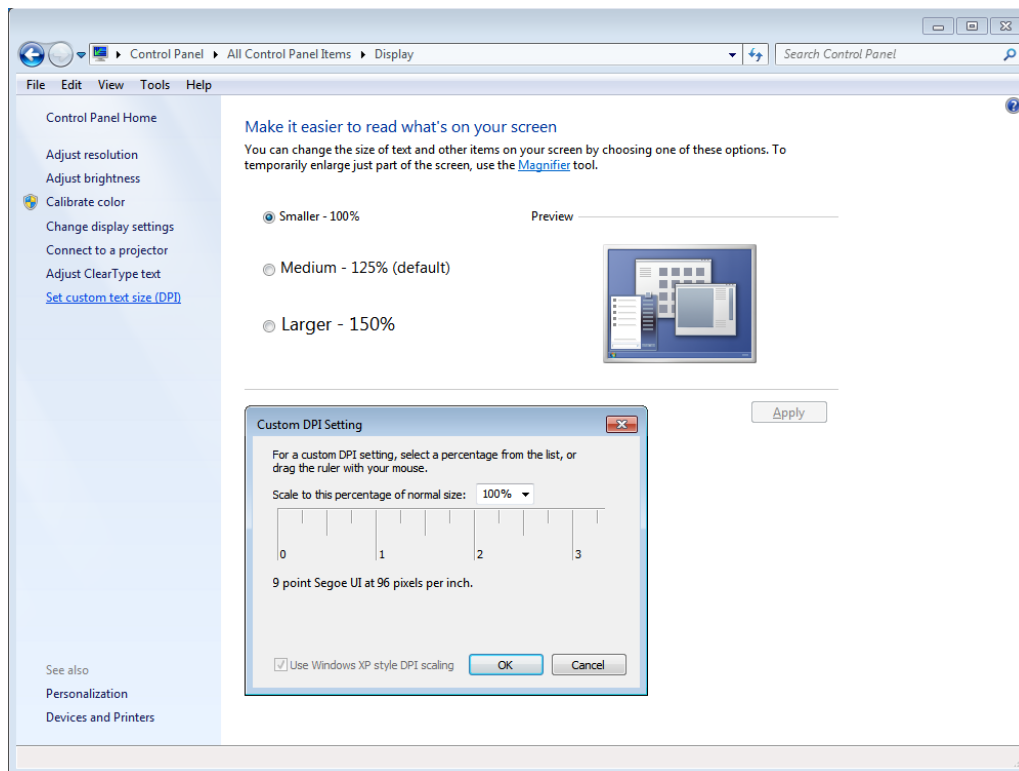
Why do I get different results than someone else?

There are many possible reasons why your results might differ from someone else's results. One good place to start is the Audit Trail Reporting option. With the Audit Trail you can examine the assumptions and selections that you have made to generate your results and compare your selections with those made in another analysis.

What do I need to be aware of if I use BenMAP-CE for a local scale analysis?

Perhaps the most important issue is to make sure that you have identified the resolution of your analysis and have the appropriate grid definitions loaded into BenMAP-CE. See Chapter 4 (Section 4.1.1) to read about grid definitions. The next key

step, which is closely connected to the grid definitions, is to determine the data that you want to use. Data such as air quality modeling, incidence data, and population data need to match the grid definitions that you are using. You also need to be careful about the formatting of your data when loading it into BenMAP-CE. Chapter 4 also provides information on loading data into BenMAP-CE.

Does BenMAP-CE estimate effects of air pollution that are not related to human health (i.e., ecological effects)?

No. BenMAP-CE does not currently have impact functions to estimate other than human health effects. In principle, it would be possible to estimate ecological effects, as BenMAP-CE is designed to combine different types of geographically variable data. To do so, you would need to develop and load data and impact functions appropriate to estimating ecological effects of interest.

How can I get training for BenMAP-CE?

A variety of training resources are available, including self-paced exercises, online interactive modules and instructor-led training are available at: https://www.epa.gov/benmap/benmap-ce-training-materials

Where can I find the source code for BenMAP-CE?

BenMAP-CE is an open source program and the development team welcomes contributions and scrutiny from the user community. If you are interested in receiving a current copy of the source code, see https://github.com/BenMAPCE/BenMAP-CE.

CHAPTER 2
Terminology

In this chapter…

- **Find definitions for common terms used in the BenMAP-CE tool and in this manual.**

Active Layer. In the GIS window, the active layer is the top-most data layer. All queries or statistical analysis of the map will act upon this top-most layer.

Aggregation. The summing of grid cell level results to a larger spatial scale, such as county, state, or national levels.

Aggregation, Pooling, and Valuation (APV) Configuration. APV configurations store your preferences regarding how to aggregate your results, whether and how to pool your results, and any economic valuation functions you have applied. For example, an APV file might aggregate your estimated change in incidence to the U.S. county level, it might pool across multiple hospital admission health impact functions and it could include an economic valuation function. APV configurations are stored in files with an *.apvx* file extension. The results derived from an APV configuration have an *.apv**r**x* file extension. APV files are by default stored in the *APV* folder, and APV results files are by default stored in the *CFGR* folder.

Air Quality Surface. An air quality surface contains modeled or monitored air pollution data in a series of cells; these cells may be a regular shape (like a 12km by 12km grid) or an irregular shape (like a county or census tract). These surfaces are also referred to as air quality grids. BenMAP-CE uses one air quality grid to represent the baseline scenario and a second grid to represent the control scenario. These baseline and control grids must share the same geographic structure. The program calculates the difference between baseline and control grids as an input to the health impact function. Air Quality Grids are stored in files with an *.aqgx* file extension.

Air Quality Metric. The metric expresses the time period over which air quality values are modeled or observed and whether that modeled or observed air quality value is an average, maximum or minimum. For example, the metric DailyMean represents the average concentration for the sampling day. This could be taken directly from a single 24-hour observation or from an average of hourly (or more frequent) observations. In addition to the time period, some metrics also specify the method used for averaging or aggregation. For example, a typical ozone metric D8HourMax represents the highest of the 8-hour moving averages during the day.

Air Quality Model. Air quality models are valuable air quality management tools. Models are mathematical descriptions of pollution transport, dispersion, and related physical and chemical processes in the atmosphere. Air quality models (like CMAQ[1] and CAMx[2]) are used to estimate the air pollutant concentration at specific locations, which are referred to as receptors, or over a spatial area that has been divided into uniform grid squares. The number of receptors or grid-cells in a model far exceeds the number of monitors one could typically afford to deploy in a monitoring study. Therefore, models provide a cost-effective way to analyze pollutant impacts over a wide spatial

---

[1] Community Multi-scale Air Quality (CMAQ) Model is available at:
http://www.epa.gov/amad/Research/RIA/cmaq.html or https://www.cmascenter.org/cmaq/.
[2] Comprehensive Air Quality Model with Extensions (CAMx) is available at: http://www.camx.com/.

area where factors such as meteorology, topography, and emissions from both local and remote sources could be important. BenMAP-CE does not contain an air quality model.

Attainment. The state of meeting the National Ambient Air Quality Standard (NAAQS) for a pollutant. A geographical area that meets the NAAQS is called an "attainment area."

Audit Trail. This is a report that contains a record of all the choices involved in creating a particular file. Audit trails can be created for any file that BenMAP-CE creates.

Background Concentration. The concentration of a pollutant, generally in the absence of human sources.

Background Incidence. The incidence of a given adverse effect due to all causes including air pollution. Also called baseline incidence rates.

Baseline Scenario. The air quality levels prior to whatever policy change you are evaluating. The baseline is sometimes referred to as "Business as Usual." The baseline scenario is usually considered to be the reference scenario against which to compare a potential "control scenario", in which air quality levels are changed from the baseline levels.

Beta. The coefficient for the health impact function. The value of beta (ß) typically represents the percent change in a given adverse health impact per unit of pollution.

Closest Monitor. The procedure by which data from the closest monitor is used to represent air pollutant levels in a population grid cell. BenMAP-CE can also scale the data from the closest monitor with air pollution modeling data. BenMAP-CE includes two types of scaling - "temporal" and "spatial". See "Scaling" for additional information.

Community Multi-scale Air Quality (CMAQ) Model. An open-source photochemical grid air quality model that the U.S. EPA and others rely upon to predict levels and changes in pollutant concentrations.

Concentration-Response (C-R) Function. A C-R function estimates the relationship between adverse health effects and ambient air pollution, and is used to derive health impact functions (defined below). You will often see that the term C-R function and health impact function are used interchangeably.

Configuration. A Configuration stores the health impact functions and model options used to estimate adverse health effects. Configurations are stored in files with a *.cfgx* file extension. CFGX files are by default stored in the *CFG* folder. The results derived from a Configuration have a *.cfgrx* file extension. CFGR files are by default stored in the *CFGR* folder.

Contingent Valuation. A survey-based economic technique for the valuation of non-market resources, such as environmental preservation or avoidance of air pollution health risk.

Control Scenario. In a modeling study, this is a sensitivity scenario in which emissions from one or more source sectors are changed (increased or decreased) from a given "baseline scenario". The control scenario generally represents air quality levels after a new policy has been implemented.

Core BenMAP-CE. The fully-featured benefits analysis program that accepts user-defined air quality data, quantifies health impacts, aggregates, values and pools results (details available in Section 3.1). Non-Core BenMAP-CE features include Command Line and PopSim, described in the following Chapters and Appendices.

Cost of Illness (COI). The cost of illness includes the direct medical costs and lost earnings associated with illness. These estimates generally understate the true economic value of reductions in risk of a health effect, as they include just the direct expenditures related to treatment and lost earnings but not the value of avoided pain and suffering.

Currency Year. The value of the currency based on the year specified. Valuation estimates should use a consistent currency year to account for inflation. For example, you might want to report the valuation estimate in 2000 dollars to make it easier to compare with your cost analysis, which uses that same currency year.

Deltas. The difference between two data points. As used in BenMAP-CE, mapping the air quality deltas shows the change in air pollution between the baseline air quality grid and the control air quality grid.

Discount Rate. In a cost-benefit analysis, the discount rate is a quantitative method to account for the fact that people generally value future benefits and costs less than current costs and benefits. Typically, if a benefit occurs over multiple years, the economic benefit would be discounted.

Endpoint. An endpoint is a subset of an endpoint group, and represents a more specific class of adverse health effects. For example, within the endpoint group *Mortality*, there might be the endpoints *Mortality, Long Term, All Cause* and *Mortality, Long Term, Cardiopulmonary*.

Endpoint Group. An endpoint group represents a broad class of adverse health effects, such as premature mortality or hospital admissions. BenMAP-CE only allows pooling of adverse health effects to occur within a given endpoint group, as it generally does not make sense to sum the number of cases of disparate health effects, such as premature mortality and hospital admissions.

Epidemiology. The study of factors affecting the health and illness of populations. Epidemiological studies cannot prove that a specific risk factor actually causes the disease being studied but can only show that a risk factor is associated (correlated) with a higher incidence of disease in the population exposed to that risk factor.

FIPS Code. Federal Information Processing Standard codes. Each state in the United States is assigned a 2-digit code. For example, "01" refers to Alaska, "37" refers to North Carolina, and "56" refers to Wyoming.

Fixed Effect Pooling. Fixed effect pooling is one method to combine two or more distributions of health impact or economic value estimates into a single new distribution. Fixed effect pooling assumes that there is a single true underlying relationship between these component distributions, and that differences among estimated parameters are the result of sampling error. Weights for the pooling are generated via inverse variance weighting, thus giving more weight to the studies that exhibit lower variance and less weight to the input distributions with higher variance. See Random Effects Pooling below for additional information regarding pooling techniques.

Fixed Radius. An option to interpolate air quality data points that uses all monitors within a fixed radius (or distance) of a given point of interest. All monitors are used and weighted by their relative distance.

Geographic Area. Designation of a grid definition for linkage to a specific health impact function.

GIS. Geographic Information System. A GIS is a system of hardware and software used for storage, retrieval, mapping, and analysis of geographic data.

Global Burden of Disease. The World Health Organization global burden of disease (GBD) study measures burden of disease using the disability-adjusted-life-year (DALY). This time-based measure combines years of life lost due to premature mortality and years of life lost due to time lived in states of less than full health. The DALY metric was developed in the original GBD 1990 study to assess the burden of disease consistently across diseases, risk factors and regions.[3]

Grid Cell. One of the many geographic, or spatial, components within a Grid Definition. These can be regularly or irregularly shaped.

Grid Definition. A BenMAP-CE Grid Definition provides a method of breaking a geographic region into areas of interest (Grid Cells) in conducting an analysis. This can be done in two ways - by loading a Shapefile (a particular type of GIS file) or by specifying a regularly shaped grid pattern. These are referred to as Shapefile Grid Definitions and Regular Grid Definitions, respectively. Typically a Shapefile Grid Definition is used when the areas of interest are political boundaries with irregularly shaped borders, while a Regular Grid Definition is used when the areas of interest are uniformly shaped grids (e.g., rectangles). All grid definitions must have a unique (i.e., non-repeating) column and row index.

---

[3] For more information on the GBD, see: http://www.who.int/topics/global_burden_of_disease.

Health Impact Function. A health impact function calculates the change in adverse health effects associated with a change in exposure to air pollution. Based on a C-R function, a typical health impact function has inputs specifying the air quality metric and pollutant; the age, race and ethnicity of the population affected; and the incidence rate of the adverse health effect.

Incidence. The total number of adverse health effects in a geographic region in a given unit time. In BenMAP-CE, this is the total number of adverse health effects avoided due to a change in air pollution levels.

Incidence Rate. The background rate of a health effect per person in a given geographic region. The average number of adverse health effects per person per unit of time, typically a day or a year. The incidence rate must be expressed at the same time scale as the remainder of the health impact function. For example, a health impact function quantifying day-to-day changes in premature death must specify a daily death rate.

Income Growth Adjustment. Adjusting certain valuation functions to reflect increases in real income over time. Generally, an increase in real income implies an increase in the willingness to pay (WTP).

Interpolation. The process of estimating the air quality level in an unmonitored area by using one or more nearby air quality monitors. BenMAP-CE uses two types of interpolation procedures: one is to simply choose the closest monitor, the other is to use a technique called Voronoi Neighbor Averaging. These interpolation methods are discussed in more detail in Appendix B.

Lat/Long. Latitude and longitude information to specify the geographic coordinates of a spatial location. The CMAQ model data are usually provided for each grid cell identified by the latitude and longitude of the grid cell's center point. Latitude identifies the north-to-south location of a point on the Earth. Longitude identifies the east to west location of a point on the Earth.

Layer. In GIS, a layer represents a logical separation of mapped data usually representing a theme, such as political boundaries, roads, ozone data, number of mortalities avoided, etc.

Layer Statistics. The summary statistics that correspond to the active layer in BenMAP-CE. For example, "mean", "standard deviation" or "max" of $PM_{2.5}$ air quality grid.

Metadata. Data that serves to provide context or additional information about other data. BenMAP-CE stores a minimum set of standardized metadata fields for imported data files (e.g., file name, file date, reference, import date, and description). For certain data types, additional metadata are recorded. For example, GIS metadata will include information about datum, geographic coordinate system, resolution, and units.

Micrograms per Cubic Meter (µg/m³). The unit of measure for particulate matter in the NAAQS. This unit represents the mass of PM and other particle pollutants found in a cubic meter of air.

Model Data. Pollutant concentration data that are generated by running an air quality model such as CMAQ. This is different from "monitor data," which are based upon observed concentrations.

Monetize. In the context of human health benefits assessment, this is the practice of expressing society's preferences for avoiding certain health effects as an economic value (e.g., in U.S. dollars). In BenMAP-CE we estimate monetized benefits by using either Willingness to Pay or Cost of Illness valuation functions (see above and below).

Monitor Data. Pollutant concentration data that are based upon measurements from an air quality monitor. "Raw" monitor data usually refers to data that are taken directly from measurement networks, with no additional processing of the data. Monitor data are different from "model data," which are based upon numerical predictions from an air quality model.

Monitoring. Actual measurements of air pollution concentrations. The U.S. EPA has monitoring data, as well as other information related to monitoring, available through its Air Quality System (AQS): https://www.epa.gov/aqs.

Monte Carlo Simulation. A technique used in BenMAP-CE to quantify the confidence intervals around mean incidence and economic value estimate by randomly sampling an uncertainty distribution around the effect coefficients or willingness to pay estimates.

Morbidity. A measure of being diseased or afflicted by an illness (generally non-fatal).

Mortality. A measure of the number of deaths in a given population.

National Ambient Air Quality Standards (NAAQS). The U.S. EPA establishes levels for pollutants that are considered harmful to public health and the environment. The Clean Air Act established two types of national air quality standards. Primary standards set limits to protect public health, including the health of "sensitive" populations such as asthmatics, children, and the elderly. Secondary standards set limits to protect public welfare, including protection against decreased visibility and against damage to animals, crops, vegetation, and buildings. The U.S. EPA has set NAAQS for six principal pollutants, which are called "criteria" pollutants: carbon monoxide, lead, nitrogen dioxide, ozone, particulate matter ($PM_{2.5}$, PM_{10}), and sulfur dioxide.

Odds Ratio. A quantitative measure reported in epidemiology studies of the relationship between exposure to air pollution and a health outcome. Odds Ratios must be converted to beta coefficients to be used in BenMAP-CE

Ordinality. In relation to air quality monitors, 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 four, then a monitor can have as many as three daily averages (assuming that we are using the daily average metric to define our standard) greater than your standard without violating the standard. In terms of rollback, if it has more than three daily averages in exceedance of the standard, then the rollback technique will be applied to that monitor.

Ozone (O_3). BenMAP-CE focuses on ground level or "bad" ozone, which is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen (NOx) and volatile organic compounds (VOCs) in the presence of sunlight. Emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors, and chemical solvents are some of the major sources of NOx and VOC. Breathing ozone can trigger a variety of health problems, particularly for children, the elderly, and people of all ages who have lung diseases such as asthma. Ground level ozone can also have harmful effects on sensitive vegetation and ecosystems.

Particulate Matter. Particulate matter, also known as particle pollution or PM, is a complex mixture of extremely small particles and liquid droplets. Particle pollution is made up of a number of components, including acids (such as nitrates and sulfates), organic chemicals, metals, and soil or dust particles. Once inhaled, these particles can affect the heart and lungs and cause serious health effects. Includes $PM_{2.5}$ (particles less than 2.5 microns in aerodynamic diameter), PM_{10} (particles less than 10 microns in aerodynamic diameter), and $PM_{10-2.5}$ (particles between 2.5 and 10 microns in aerodynamic diameter).

Parts per Million (ppm). This unit represents the concentration of the pollutant in a million parts of air. Carbon monoxide is often measured in units of ppm.

Parts per Billion (ppb). This unit represents the concentration of the pollutant in a billion parts of air. Ozone concentrations in BenMAP-CE are reported in units of ppb.

POC (Parameter Occurrence Code). An identifier used by U.S. EPA to distinguish between multiple monitors at the same site that are measuring the same parameter. For criteria pollutants, multiple monitors may be collocated to check precision. For combining data at the site level, the POC identifies the primary monitor (most frequent sampling). (POC appears in BenMAP-CE's advanced filtering options for monitor data.)

Pooling. The combining of different sets of data. BenMAP-CE has several pooling methods, including fixed effects, fixed/random effects, and subjective weighting. Appendix K discusses the pooling approaches available in BenMAP-CE.

Point Mode. When defining the configuration, you may choose to either estimate adverse health effects in point mode or using percentiles. The point mode simply means that BenMAP-CE will use the mean value of the coefficient in the health impact function.

Population Exposure versus Personal Exposure. Population (or ambient) exposure refers to the average air pollution level measured in a grid cell. In contrast, personal exposure keeps track over the course of a day the exposure individuals encounter in different micro-environments, such as the freeway, outdoors and indoors. BenMAP-CE only represents population exposure.

Population-weighted Air Quality. Modeled or monitored ambient concentrations that have been weighted according to the number of people exposed.

Prevalence Rate. The percentage of individuals in a given population who already have a given adverse health condition. Used to calculate changes in health conditions among those who already have a health condition, such as asthmatics.

Random Effect Pooling. Random effect pooling is one method to combine two or more distributions of health impact or economic value estimates into a single new distribution. This approach allows the possibility that the estimated parameter from different studies may in fact be estimates of different parameters, rather than just different estimates of a single underlying parameter.

Regulatory Impact Analysis (RIA). A policy tool used to assess the likely effects of a proposed regulation or regulatory change. It usually involves detailed analyses to quantify the costs and benefits of the regulation.

Relative Risk. Relative risk typically is used as a measure of the change in risk of an adverse health effect associated with an increase in air pollution levels in an epidemiology study. More specifically, it is the ratio of the risk of illness with a higher pollution level to the risk of illness with a lower pollution level, where the "risk" is defined as the probability that an individual will become ill.

Rollback. The process by which monitor data are reduced to a different level. BenMAP-CE rolls back monitor data in three ways. Percentage rollback reduces all monitor observations by the same percentage. Incremental rollback reduces all observations by the same increment. Rollback to a standard reduces monitor observations so that they just meet a specified standard.

Setup. A BenMAP-CE setup encapsulates all of the data needed to run analyses for a particular geographic area—a city, an entire country, etc. These data consist of grid definitions, pollutants, monitor data, incidence and prevalence rates, population data, health impact functions, variables, inflation rates, and valuation functions.

Shapefile. A shapefile is a particular type of GIS file, and has a *.shp* extension. These files are accompanied by companion files with *.shx* and *.dbf* extensions, and can be used to create Shapefile Grid Definitions.
See http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf for more information.

Sum Dependent Pooling. Summing two or more incidence or valuation results, assuming the underlying functions are correlated. For example, summing the incidence of respiratory hospital admissions for two different age groups quantified using C-R functions from the same study.

Sum Independent Pooling. Summing two or more incidence or valuation results, assuming the underlying functions are independent (uncorrelated). For example, summing the incidence of respiratory hospital admissions quantified using C-R functions from different studies using different methods.

Threshold. BenMAP-CE's advanced settings for health impact functions allows you to specify an air quality threshold; this is an air quality level below which benefits are not calculated. For example, if the threshold is 5 μg/m^3, then only areas with PM$_{2.5}$ concentrations equal to or greater than 5 μg/m3 will be included in estimating health incidence results. Specifying a threshold does not affect the shape of the C-R function used to quantify impacts.

Unit Value. A unit value is the estimated mean economic value of avoiding a single case of a particular health effect.

User-defined Weights Pooling. User-defined weights let you specify the weights that you want to use when combining two or more distributions of results. The weights should sum to one. If not, BenMAP-CE normalizes the weights so that they do.

Valuation Function. Valuation functions are used by BenMAP-CE to estimate the economic values of changes in the incidence of health effects. These are selected within an Aggregation, Pooling, and Valuation Configuration (APV Configuration).

Variable Datasets. Health Impact functions and valuation functions may sometimes refer to variables other than those for which BenMAP-CE automatically calculates values. For example, some valuation functions reference the median income within each area of analysis. To facilitate this, BenMAP-CE allows you to load datasets of variables for use in functions, which may be used globally or may vary geographically (meaning they are associated with a particular Grid Definition).

VNA (Voronoi Neighbor Averaging). An algorithm used by BenMAP-CE to interpolate air quality monitoring data to an unmonitored location. BenMAP-CE first identifies the set of monitors that best "surround" the center of the population grid cell, and then takes an inverse-distance weighted average of the monitoring values. This is discussed in detail in Appendix B.

WTP (Willingness to Pay). The willingness of individuals to pay for a good or service, such as a reduction in the risk of illness. In general, economists tend to view an individual's WTP for an improvement in environmental quality as the appropriate measure of the value of a risk reduction. An individual's willingness to accept (WTA) compensation for not receiving an improvement is also a valid measure. However, WTP

is generally considered to be a more readily available and conservative measure of benefits.

CHAPTER 3

Overview of BenMAP-CE Features

In this chapter...

- **Get an overview of the features available with the Core Program.**
- **Learn about additional BenMAP-CE modules.**
- **Learn about the Tools and Help menu options.**
- **Find descriptions of the various outputs including types of files, results, maps and reports available from BenMAP-CE.**

Chapter 3 Table of Contents

Upon starting BenMAP-CE for the first time, you will see the following **Welcome** screen.
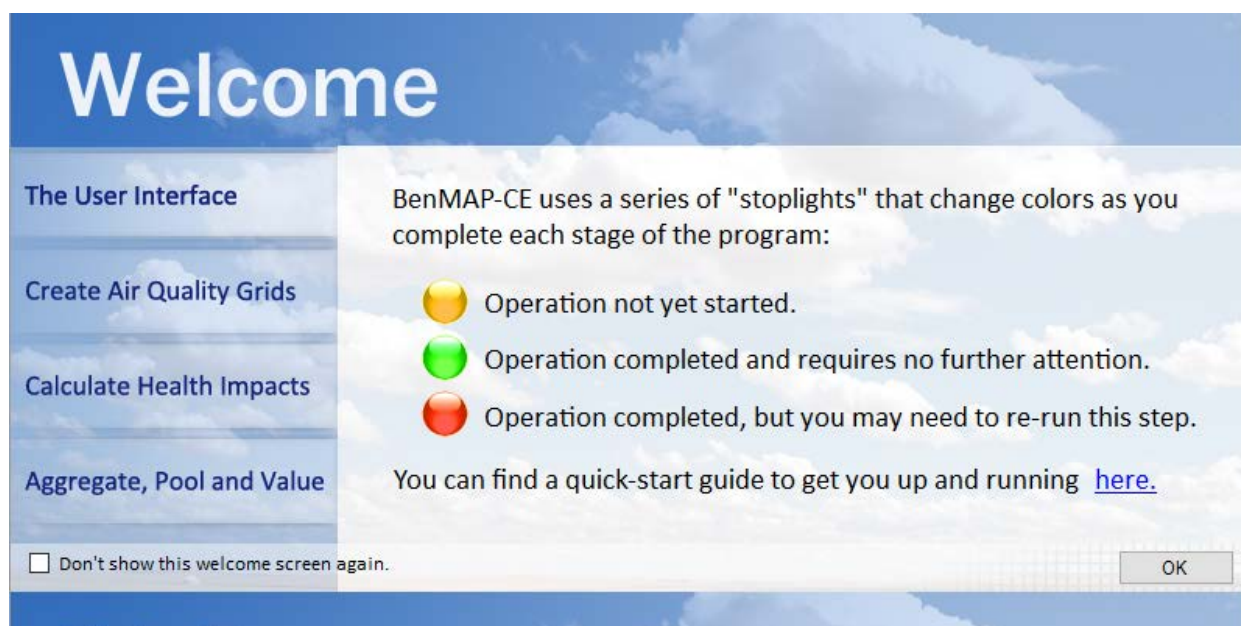
The **Welcome** screen gives a brief description of the user interface and highlights the "stoplight" metaphor used in BenMAP-CE to indicate the status of analytical steps performed using the tree menu on the left side of the main window. Clicking the links on the left side of the **Welcome** screen (e.g., **Create Air Quality Grids**) will provide information about each feature. You will also find a link to the BenMAP-CE website for downloading the most current BenMAP-CE software and other reference information. If you do not wish to see the **Welcome** screen at program start-up, check the option '*Don't show this window screen again'* in the lower left-hand corner of the screen. To re-enable the **Welcome** screen, go to **Tools** menu on the main BenMAP-CE window, select **Options,** and check the option for *Show Start Window*. Press **OK** on the **Welcome** screen to close this window and display the BenMAP-CE main window.
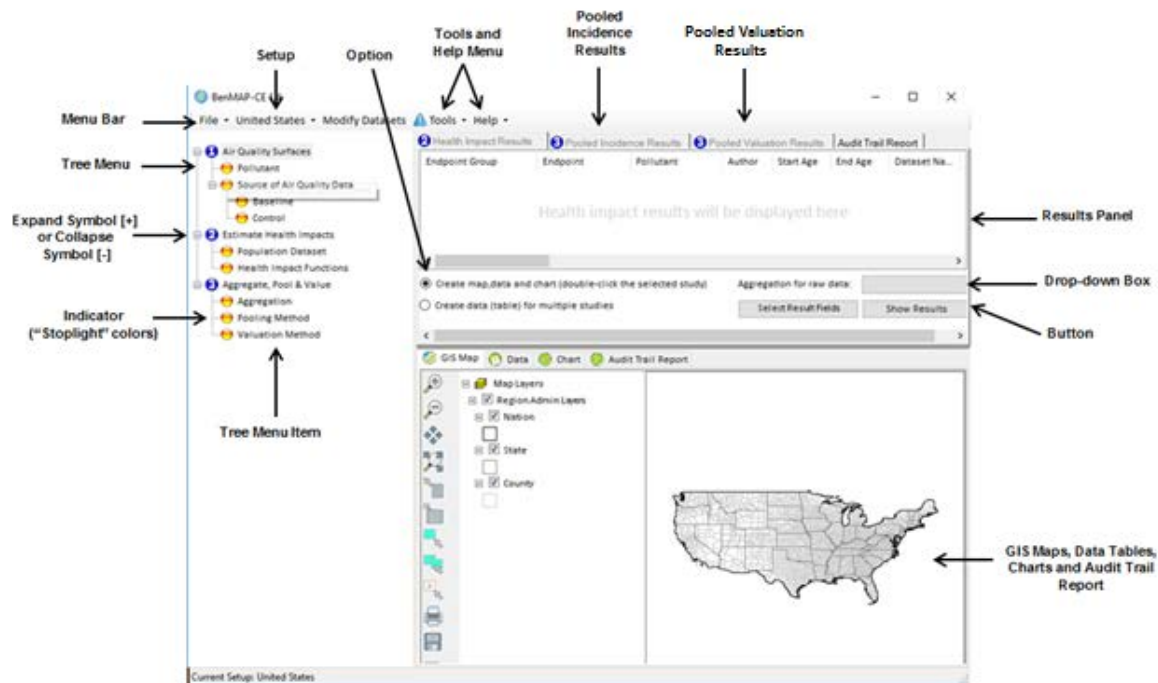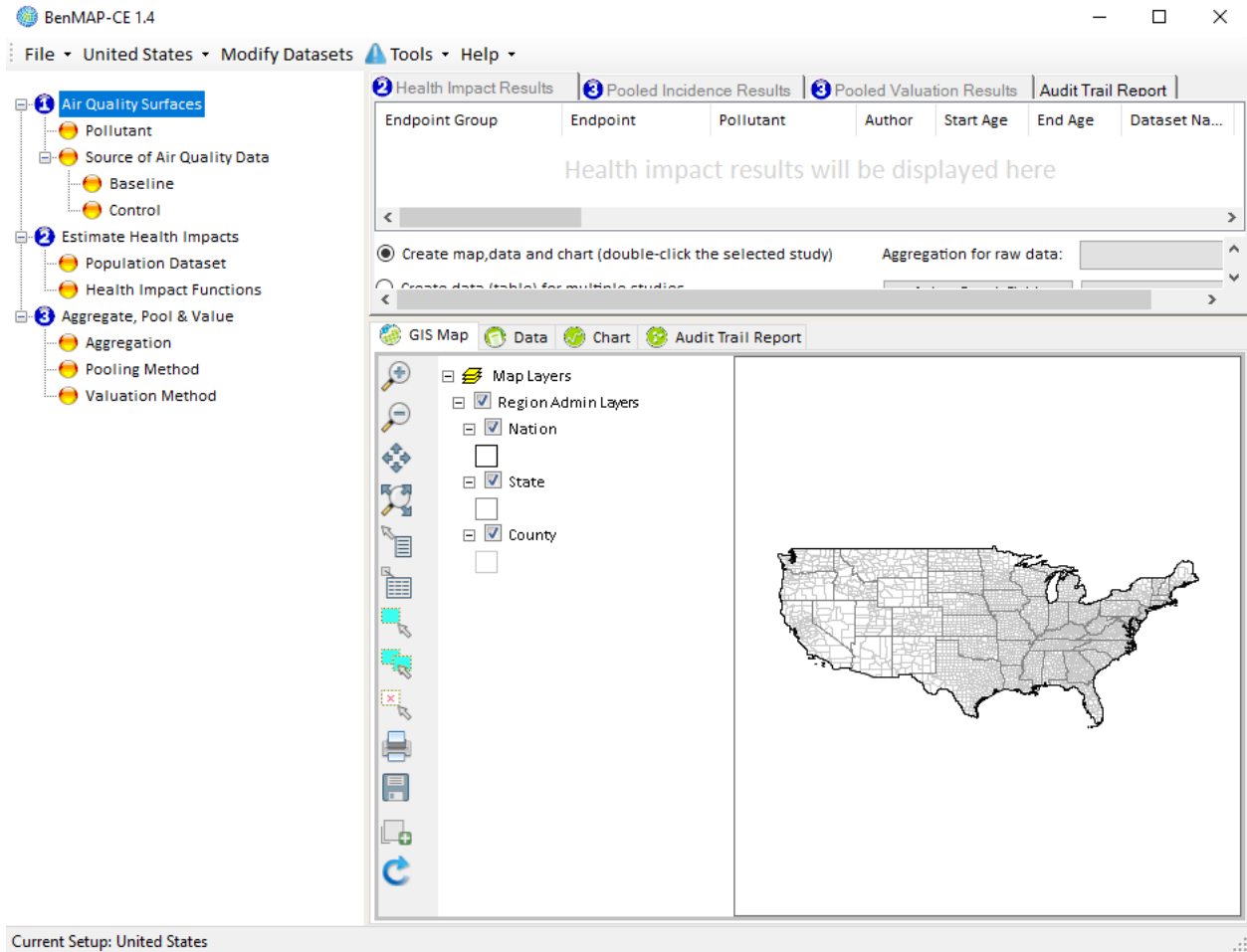
The tree menu on the left-hand window pane lists the analytical steps used in the Core Program. The tree menu items allow you to perform a highly customized health impact analysis. The **Tools** menu at the top of the screen is for less frequently used functions, such as importing and exporting data and special modules described in later sections.

The first section in this Chapter describes the Core Program features. The second section describes the additional functions found in the **Tools** and **Help** menus. The next section covers BenMAP-CE output options. Lastly, we answer some frequently asked questions. Note that this chapter provides an overview of functionality, not keystroke by keystroke instructions. Those may be found in Chapters 4 through 9.

3.1   Core Program

Beginning at the main BenMAP-CE window, you can choose the **Setup** you want to use for an analysis by selecting it from the **Setup** drop-down list. You are then ready to begin using the features available through the tree menu.
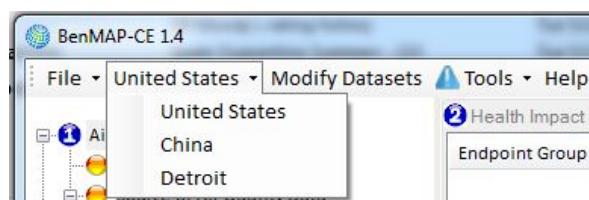
The tree menu takes you through the steps of an analysis. The first step, **Air Quality Surfaces**, allows you to select the pollutant of interest, and then specify the baseline and control air quality surfaces. The second step, **Estimate Health Impacts**, lets you choose the population dataset for a particular analysis, and then specify the health impact functions to estimate the incidence of adverse health effects. The last step, **Aggregate, Pool & Value**, gives you options for combining the health effects estimates and choosing economic valuation functions.

3.1.1 Create Air Quality Surfaces

BenMAP-CE is not an air quality model. Instead, it relies on externally-created air quality modeling data inputs. To estimate population exposure to air pollution, BenMAP-CE combines population data with air quality surfaces, which it generates using some combination of air quality modeling and/or monitoring data. In BenMAP-CE, air quality surfaces can be described as air quality grids (the structure) that have been populated with air pollution values (the data). The following is a brief description of each step. For detailed instructions, see Chapter 4: Loading Data and Chapter 5: Creating Air Quality Surfaces.

Pollutant

The **Pollutants** section of a setup specifies the pollutants that BenMAP-CE will analyze and defines the air quality metrics to be used by BenMAP-CE. You are not importing air pollution data, but rather naming your pollutants and defining the measures or metrics BenMAP-CE will use when assessing the health impacts associated with each pollutant.

Grid Definition

Air quality surfaces contain air pollution exposure estimates for a particular **Grid Definition**, as defined in the **Modify Datasets** window. **Grid Definitions** are typically comprised of either regularly shaped rectangles covering the region of analysis, or irregularly shaped polygons corresponding to political boundaries.

Modeling and Monitoring Data

To generate air quality grids, you can use air quality modeling data and air quality monitoring data in three different ways, as discussed below. However, once generated, all air quality grids have the same structure, and have the same *.aqgx extension that BenMAP-CE uses to designate these file types.

▪ **Model Data.** *Model Data* grid creation simply takes raw model data and converts it into a file that BenMAP-CE recognizes as an air quality grid. This type of grid definition allows you to directly specify the air pollution values for each grid cell in a Grid Definition.

▪ **Monitor Data.** *Monitor Data* grid creation uses air pollution monitoring data to estimate air pollution levels for each grid cell in the selected Grid Definition. This may be done using one of three interpolation procedures - *Closest Monitor*, *Voronoi Neighborhood Averaging* (VNA), or *Fixed Radius*. With closest monitor, BenMAP-CE simply uses the data of the monitor closest to each grid cell's centroid. With VNA, BenMAP-CE first identifies the set of monitors that most closely "surround" each grid cell, and then calculates an inverse-distance weighted average of the data from these neighboring monitors. With fixed radius, BenMAP-CE constrains the VNA interpolation to a user-specified distance around each monitor.

▪ **Monitor Rollback.** *Monitor Rollback* grid creation allows you to reduce, or roll back, monitor data using three methods: *Percentage Rollback*, *Incremental Rollback*, or *Rollback to a Standard*. Percentage rollback reduces all monitor observations by the same percentage. Incremental rollback reduces all observations by the same increment. Rollback to a standard reduces monitor observations so that they just meet a specified standard. After the monitor data is rolled back, it may be directly interpolated (as in *Monitor Data* grid creation) or combined with modeling data. This approach is described in more detail in Chapter 5: Creating Air Quality Grids, as well as in the Appendix A: Monitor Rollback Algorithms.

Once an air quality grid is created, it can be saved as a .aqgx file and reopened without any additional inputs needed.
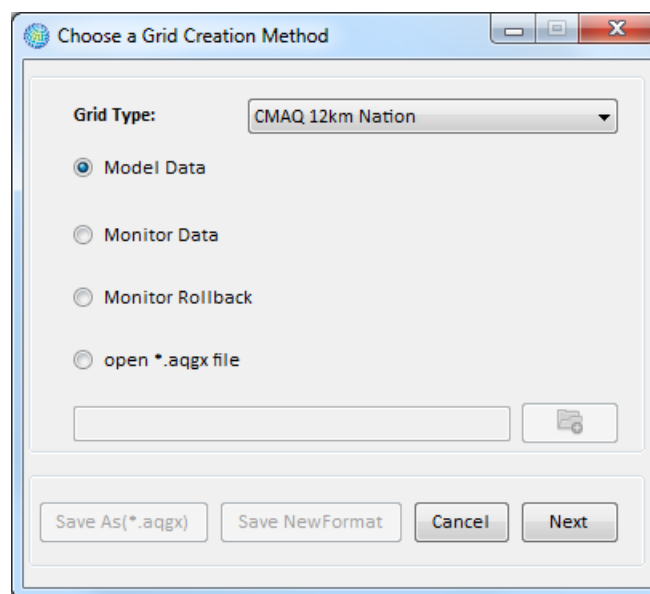
3.1.2 Estimate Health Impacts

The **Estimate Health Impacts** section allows you to calculate the change in the incidence of adverse health effects associated with changes in air quality. There are three steps in the process. The following is a brief description of each step. For detailed instructions, see Chapter 6: Estimating Incidence.

- **Step 1.** Specify the *Population Dataset* and *Population Year*.
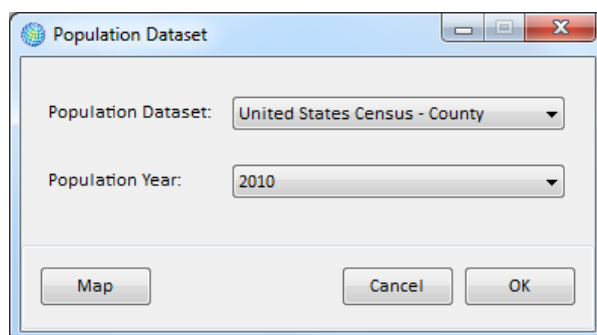
- **Step 2.** Choose the *Health Impact Functions* that will be used to estimate the incidence of adverse health effects.
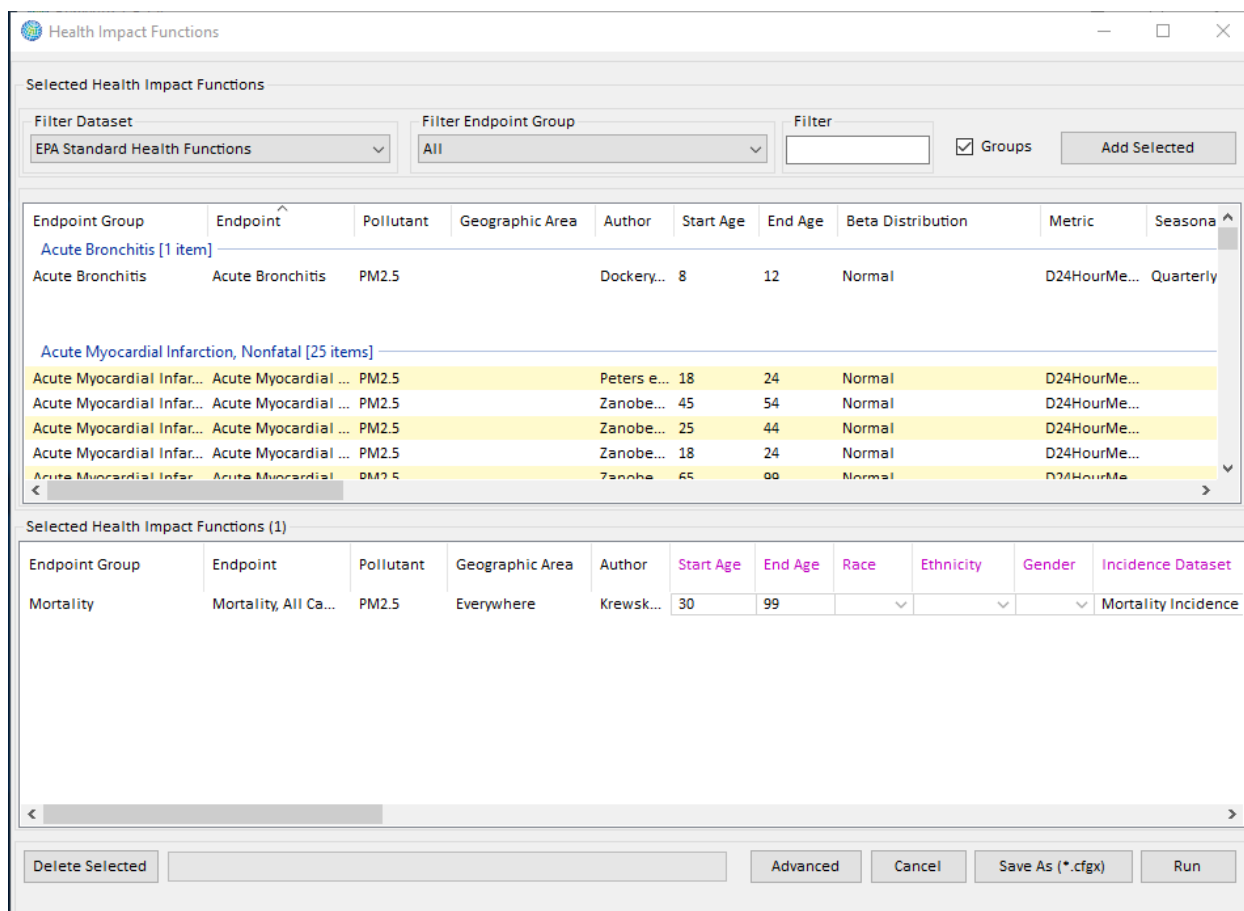
Step 3. BenMAP-CE performs a full Monte-Carlo analysis to quantify the confidence intervals around mean incidence and economic value estimates by randomly sampling an uncertainty distribution around the effect coefficients or willingness to pay estimates. This step typically occurs automatically using default parameters for the Monte-Carlo analysis. However, you may change these defaults in Advanced Settings, as shown in the screenshot below. In general, the computation time increases as you specify additional percentiles to report from the Monte-Carlo generated distribution. If you want to replicate the Monte-Carlo distribution from another analysis, then you may also specify the *Random Seed*.) Specify the *Air Quality Threshold*, or a lowest value for air quality data. Any observations which fall below this threshold will be replaced with the threshold value in all calculations.
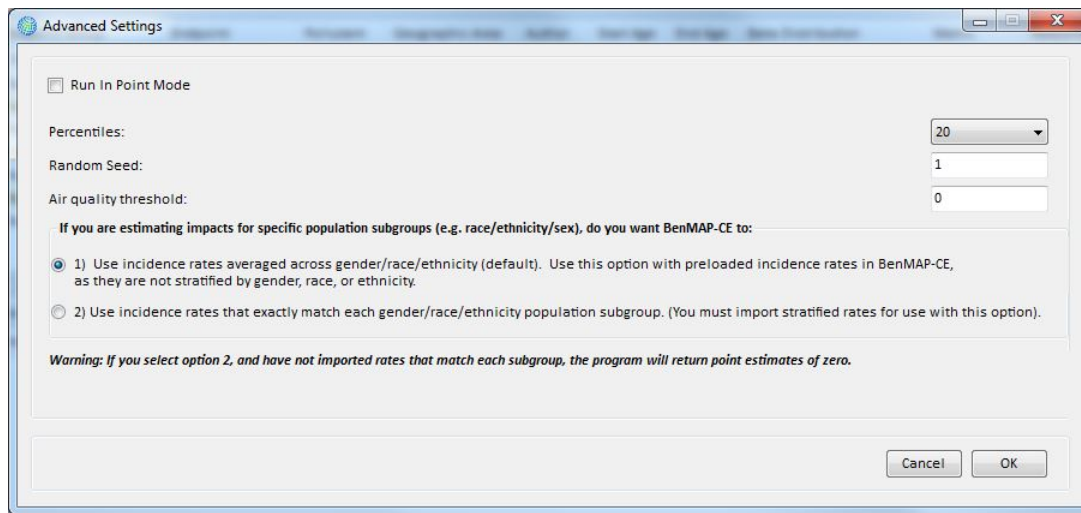
BenMAP-CE can store configuration choices in a user-named file with a *.cfgx* extension, and can store incidence change estimates in a user-named file with a *.cfgrx* extension.

3.1.3   Aggregate, Pool, and Value

The **Aggregate, Pool, & Value (APV)** feature on the BenMAP-CE tree menu allows you to aggregate and pool previously calculated incidence estimates and place an economic value on these pooled and aggregated incidence estimates. You can also aggregate the economic values, and finally pool the aggregated economic values. There are several steps in this process. The following is a brief description of each step. For detailed instructions, see Chapter 7: Aggregating, Pooling, and Valuing.

- ▪ **Step 1.** Choose the desired aggregation levels.
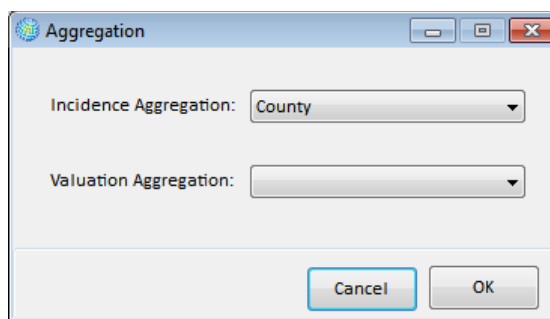
- ▪ **Step 2.** Choose the desired pooling and aggregation options for the incidence results.
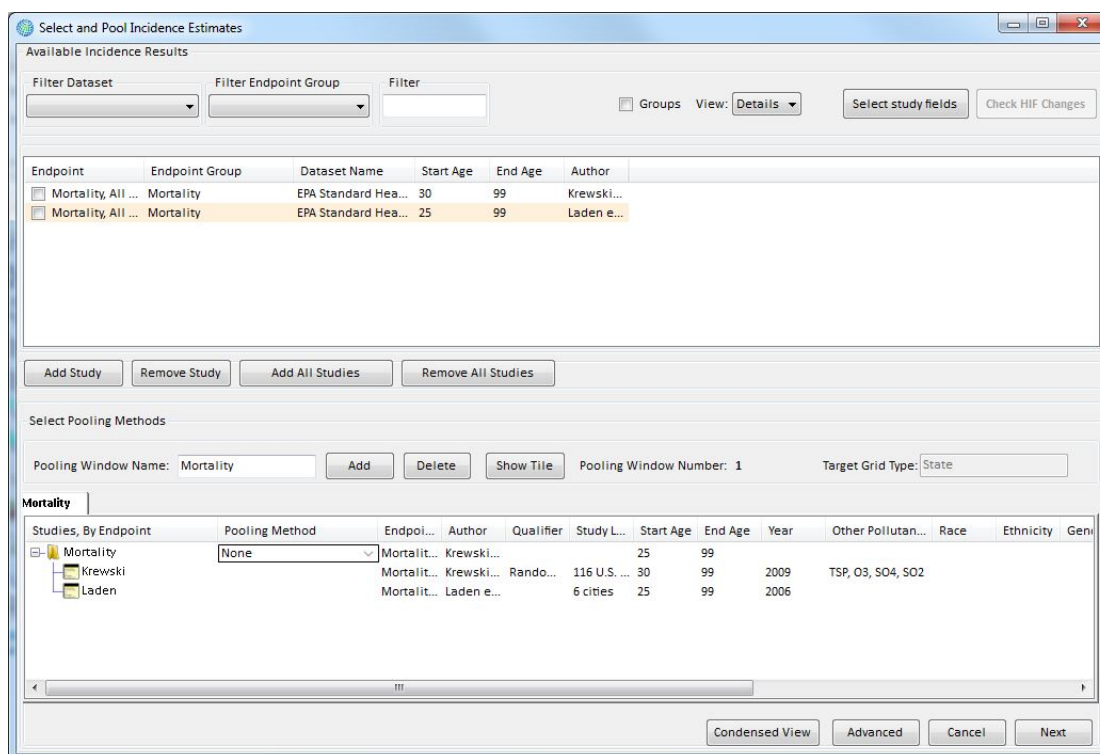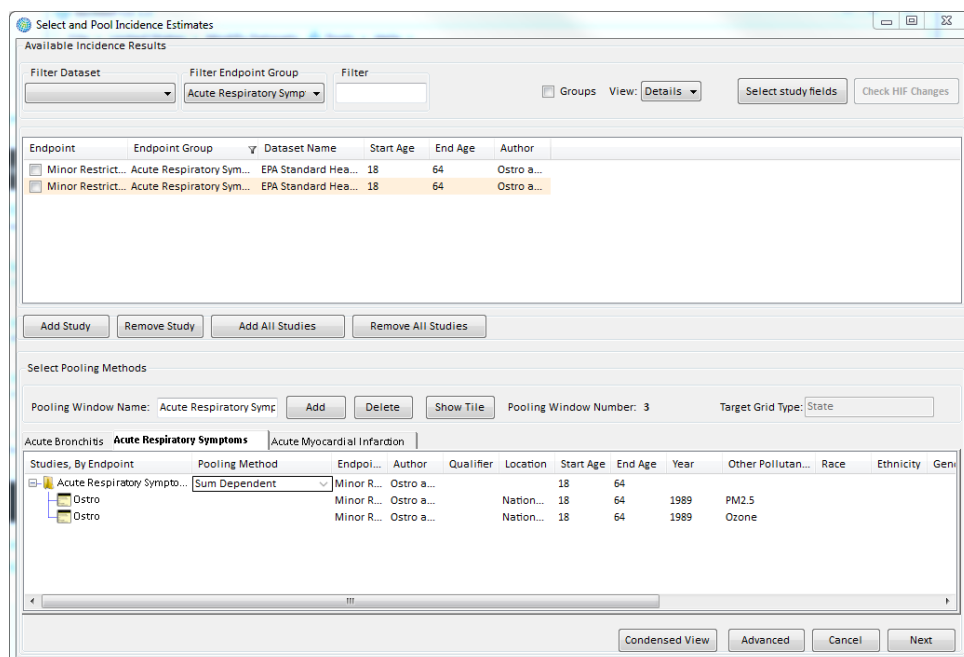
- **Step 3.** Choose the economic valuation functions to apply to the pooled and aggregated incidence results.
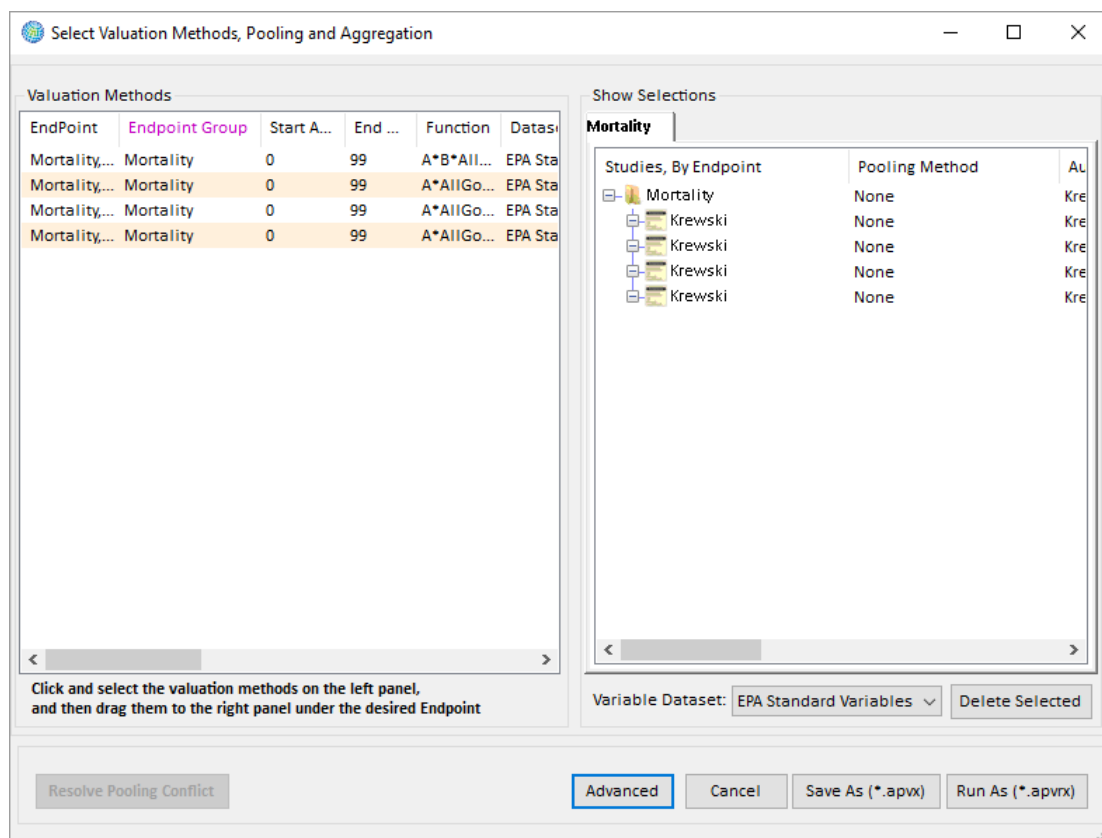
BenMAP-CE can store APV configuration choices in a user-named file with an *.apvx* extension, and can store APV configuration results in a user-named file with an *.apvrx* extension. As needed, you can access both files for later use.

3.2   Menus

There are five menu options found at the top of the main window: **File**, **Setup**, **Modify Datasets**, **Tools** and **Help**.

- **File.** This menu provides options for selecting saved project files (*.projx*) to open, starting a new project file, saving your work in a project file, and exiting the program.

- **Setup.** BenMAP-CE comes pre-loaded with datasets for the *United States, China*, and *Detroit* setups. The selected setup will be displayed in the menu bar, next to the File menu. You may click on the selected setup to see a menu of available setups. Each setup includes the information needed to run analyses for a particular geographic area. To learn more about modifying setups, see Chapter 4: Loading Data.

- **Modify Datasets.** BenMAP-CE stores the information needed to run analyses for a particular geographic area, such as a city, region, or nation in a single dataset called a setup. Many users will never need to modify the preloaded setups. However, the **Modify Datasets** menu provides tools to add, modify (load additional datasets), or delete these setups if needed. This is discussed in detail in Chapter 4: Loading Data.

- **Tools.** This menu provides access to data import and export functionality in addition to a number of other features. An overview is provided below in Section 3.2.1.

- **Help.** This menu provides access to a *Quick Start Guide* (available on EPA's website), information *About* BenMAP-CE, and a form to *Provide Feedback* about software errors or requested features. An overview is provided below.

3.2.1   Tools Menu

The **Tools** menu has several options: *Aggregate Air Quality Surface, Database Export, Database Import, Online Database Export, Online Database Import, Export Air Quality Surface, GBD Rollback, Monitor Data Conversion, Neighbor File Creator, PopSim Options, and Compute Grid Crosswalks.* A brief description is given below, but further information can be found in Chapter 9: Tools Menu.

- *Aggregate Air Quality Surface.* Create a new air quality surface for a specified grid definition (e.g., County) from an existing air quality surface created with a different (and generally finer) grid definition (e.g., 12km CMAQ).

- *Database Export.* Export all or part of BenMAP-CE's internal database to a database (.bdbx) file or multiple csv or shapefiles which can later be used on another computer or by another user. Manually loading data into BenMAP-CE can be time and labor intensive, so this tool can be quite useful in sharing data with other users or computers.   This tool can be used to share all or part of an existing setup, or to backup Setups prior to upgrading BenMAP.

- *Database Import.* Import data created using the *Database Export* tool into a specified setup.

- *Online Database Export* (Feature currently disabled). Allows users to post/share data for BenMAP-CE via a cloud-based system.

- *Online Database Import* (Feature currently disabled)*.* Allows users to download data shared by others for BenMAP-CE via a cloud-based system.

- *Export Air Quality Surface*. Generate a text file (*.csv*) with all of the data in air quality surface, including summary statistics such as mean, median, minimum, and maximum.

- *GBD Rollback Tool*. The *GBD Rollback* tool allows you to select a country, region, or group of countries and see the impact of lowering $PM_{2.5}$ emissions based on the data from the 2010 GBD study[1]. The outputs include the baseline and policy case $PM_{2.5}$ concentrations as well as the population-weighted air quality change.

- *Monitor Data Conversion.* Create a monitor file in BenMAP-CE format using an external data file in a different format.

- *Neighbor File Creator*. Create a text file (*.txt*) identifying "neighbor" monitors and associated interpolation weights for each grid cell in an air quality grid.

- *PopSim.* A dynamic population simulation that incorporates the cumulative effects of air pollution on different age groups over time.

- *Options.* Select options for start-up and exit screens, validation logs and default set-ups.

- *Compute Grid Crosswalks.* Remove all crosswalks for selected setups and re-create them. This tool is for creating or repairing crosswalks which were broken due to database error or force quitting of the application.

---

[1] Exposure Assessment for Estimation of the Global Burden of Disease Attributable to Outdoor Air Pollution. Michael Brauer, Markus Amann, Rick T. Burnett, Aaron Cohen, Frank Dentener, Majid Ezzati, Sarah B. Henderson, Michal Krzyzanowski, Randall V. Martin, Rita Van Dingenen, Aaron van Donkelaar, and George D. Thurston. *Environmental Science & Technology* **2012** *46* (2), 652-660

3.2.2 Help Menu

The **Help** menu has a few options to choose from: *Quick Start Guide link, User Documentation, About,* and a form to *Provide Feedback.*

- *Quick Start Guide.* A link will open a webpage to the U.S. EPA site for BenMAP-CE training materials. On the webpage, users can download a series of self-paced exercises for seven different regions of the world, including two exercises in Spanish and one in French. Each self-paced exercise will take you through the basic operations of BenMAP-CE data imports and decision-making.

- *User Documentation.* Opens a link to the U.S. EPA site with the BenMAP-CE user manual and appendices.

- *About.* Opens a window that displays information about the program (e.g., software version, contact information, and a suggested citation). You can click on **Release Notes** to read about software modifications and any known issues.

- *Provide Feedback.* This feature allows you to submit any problems that you may encounter while running BenMAP-CE or requested features to the BenMAP-CE development team. There are fields to provide your contact information (optional) and information about the error or requested feature. Your feedback will be logged into an issue tracking system for U.S. EPA to evaluate.

3.3   Outputs

3.3.1   Results

If you are interested in viewing or exporting the results of an analysis, these reports can be accessed by clicking on the appropriate **Results** tabs from the upper portion of the main BenMAP-CE window. The number next to the tab description indicates the associated "step" of the BenMAP-CE analysis (from the tree menu).

Select the appropriate tab for the type of results you wish to create:

- **Health Impact Results** uses a *Configuration Results* file (with the *.cfgrx* extension) to create a map, data table and bar chart for incidence results based on the selected health impact studies. "Raw" incidence estimates are those that have not been aggregated, pooled or valued. See Chapter 6: Estimating Incidence, for detailed instructions for creating health impact results.

- **Pooled Incidence Results** uses an *Aggregation, Pooling, and Valuation Results* file (with the *.apvrx* extension) to create maps, data tables and bar charts for incidence, aggregated incidence and pooled incidence results. See Chapter 7: Aggregation, Pooling, and Valuation, for detailed instructions for creating pooled incidence results.

- **Pooled Valuation Results** also uses an *Aggregation, Pooling, and Valuation Results* file (with the *.apvrx* extension) to create maps, data tables and bar charts for valuation, aggregated valuation and pooled valuation. See Chapter 7: Aggregate, Pool, and Value, for detailed instructions for creating pooled incidence results.

The results data are viewable on the **Data** tab in the lower right frame of the BenMAP-CE main window (see below). You can also view the results in a **GIS Map** (described in the following section), or simple **Chart** format. All results can be exported as comma-separated value files (*.csv*), which can be read into spreadsheet and database programs.

GIS Map	Data	Chart	Audit Trail Report

Column	Row	D24HourMean	QuarterlyMean
1	29	11.95	11.95
1	31	10.69	10.69
1	37	11.66	11.66
1	39	10.22	10.22
1	41	10.91	10.91
1	45	10.90	10.90
1	49	12.01	12.01
1	53	10.12	10.12
1	57	11.12	11.12

3.3.2 Maps

GIS Maps can be viewed in the lower right frame of the BenMAP-CE main window. Once an air quality surface is displayed, you can choose which layers to view by selecting or deselecting items in the GIS table of contents. Spatial layers (except regional administrative layers) are semi-transparent so that overlapping layers are viewable. You can export a formatted graphics file (e.g., .png format) with the map legend and title. For detailed instructions on GIS maps, see Chapter 8: GIS/Mapping.

3.3.3 Audit Trail Report

The **Audit Trail Report** provides a summary of the options selected in the various parts of the analysis. You may generate an audit trail with any of the file types used in BenMAP-CE: Air Quality Grids (with the *.aqgx* extension), Configurations (with the *.cfgx* extension), Configuration Results (with the *.cfgrx* extension), Aggregation, Pooling, and Valuation Configurations (with the *.apvx* extension), and Aggregation, Pooling, and Valuation Configuration Results (with the *.apvrx* extension). The report itself has a tree structure,. Below is an example of an Audit Trail Report.
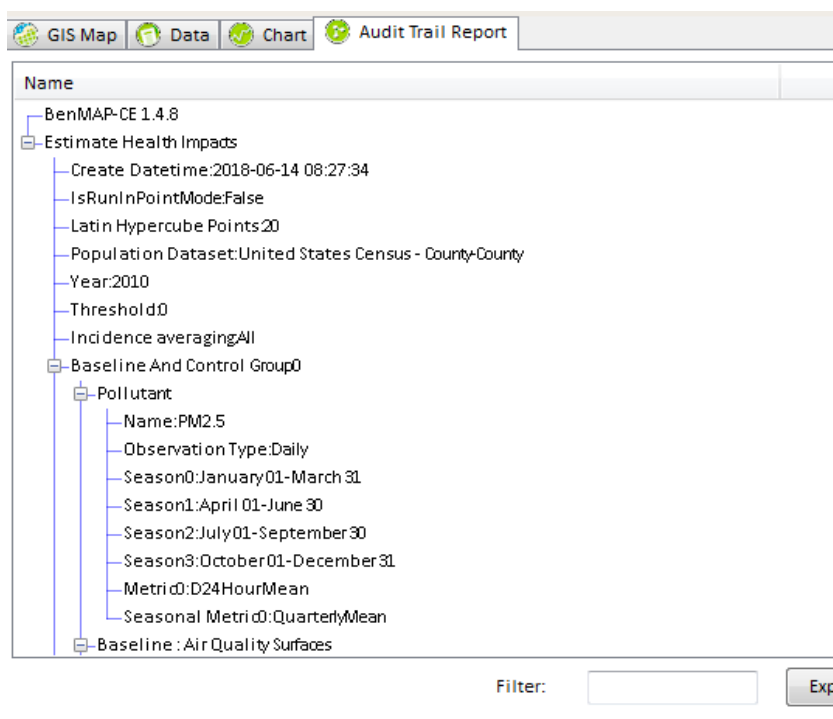
Note that each successive step in an analysis contains a summary of its inputs and attributes, and those of each previous step in the analysis. For example, in the above report the attributes of the *Health Impact Function* file used to generate the APV Results are present in the *Estimate Health Impacts* node. Similarly, the metadata for both the baseline and control air quality grids are present under the *Estimate Health Impacts* node. For more information on audit trails, see Chapter 8: GIS/Mapping.

3.3.4   Dataset Validation Reports

You may load data to BenMAP-CE to tailor the analysis to your specific needs (click **Modify Datasets** from the main menu). Loading data requires specific formatting. BenMAP-CE offers a validation option to confirm that the proper headings and data types are present in the selected file. The validation routines also check that values are within reasonable ranges for certain types of data. If the file does not meet the validation requirements, error and/or warning messages will be reported. For more information about loading and validating data, see Chapter 4: Loading Data.

3.3.5   File Types

BenMAP-CE has a number of file types that you can use to store the settings used in a BenMAP-CE analysis, the results of an analysis, as well as maps and reports. Table 3-1 presents the names of the different file types, their functions, and their default folder locations.

Table 3-1. File Types Generated by BenMAP-CE

File Extension	Description	Default Folder Location[1]
*.aqgx	Air quality grid.	Result\AQG
*.apvx	Aggregation, Pooling, and Valuation configuration specifying the aggregation levels, pooling options, and valuation methods used to generate aggregated incidence estimates, pooled incidence estimates, valuation estimates, aggregated valuation estimates, and pooled valuation estimates.	Result\APV
*.apvrx	Aggregation, Pooling, and Valuation configuration results, containing incidence results at the grid cell level, aggregated incidence results, valuation results, aggregated valuation results, and pooled valuation results.	Result\APVRX
*.bdbx	BenMAP-CE Database Export tool creates files which can contain individual datasets or entire setups. These are saved in a specific format for importing to BenMAP-CE.	User-specified
*.cfgx	Configuration specifying the health impact functions and other options used to generate incidence estimates.	Result\CFG
*.cfgrx	Configuration results, containing incidence results at the grid cell level.	Result\CFGR
*.csv	Reports (Results Tables) and PopSim tool results are exported as *.csv files, which may be viewed in a text editor, or in programs such as Excel. Users may also export portions of BenMAP setups (e.g., baseline incidence rates) as .csv files.	Result\CFGRX or Result\APVRX, PopSim
*.rtf	Validation results from data imports	ValidationResults
*.shp	Shape files generated by BenMAP-CE's geographic information maps system. These files can be viewed within BenMAP-CE or within shape file viewers, such as ArcView.	AppData\... \Shapefiles[2]
*.xlsx	GBD Rollback Tool results are exported as .xlsx files, which may be viewed in a spreadsheet tool such as Excel (there is also a .csv option).	GBD

[1] Most files generated by BenMAP-CE are stored within the User's directory under C:\Users\<user name>\Documents\My BenMAP-CE Files\.

[2] Shape files (*.shp) are stored at

C:\Users\<user name>\AppData\Local\BenMAP-CE\Data\Shapefiles\<setup name>\.

3.4   Frequently Asked Questions

When creating reports from *.cfgrx and *.apvrx files, why do some of the variables that I have checked appear as blanks?

When results are pooled, some of the identifying information for individual health impact functions gets lost. For example, when pooling endpoints within the same endpoint group, such as "HA, Pneumonia" and "HA, Chronic Lung Disease" (both within

"Hospital Admissions, Respiratory"), there is no longer a unique endpoint name for the pooled result. So, BenMAP-CE would leave the endpoint name blank.

How do I export my results?

Identify the type of report that you want to create, then refer to the Section 3.3.1 in this chapter on exporting reports.

How do I determine what the Column and Row refer to?

The Column and Row are variables designed to uniquely identify each grid cell in the grid definition. In the case of the County grid definition, the Column refers to the state FIPS code and the row refers to the county FIPS code. One way to get a good sense of the Column and Row variables is to create a map and then view where particular Column and Row variables occur in the map.

Chapter 4
Loading Data

In this chapter…

- **Learn how to create a new setup for your project.**
- **Learn more about the file structure for data inputs.**
- **Learn how to export and import a setup.**

Chapter 4 Table of Contents

BenMAP-CE can store the information needed to run analyses for a particular geographic area, such as a city, region, or nation, in a single dataset. This dataset is called a "Setup" and consists of 10 categories of data:

- Grid definitions
- Pollutants
- Monitor data
- Incidence and prevalence rates
- Population data
- Health impact functions
- Variable data (socioeconomic variables)
- Inflation rates
- Valuation functions
- Income growth data.

Grouping the data in this way has a number of advantages. It makes it easy to organize and view the data, export the data (either a whole setup or a portion of a setup), and import setups generated by others.

In this chapter we discuss how to add, modify, and delete a setup. It is important to keep in mind that if you delete one part of a setup, you may be affecting other parts of the setup. We discuss this further below.

M**any users will never need to modify the setup**. If you are performing an analysis with the pre-loaded *United States*, *China, or Detroit* setups, you may find that BenMAP-CE contains all of the data you need to perform your analysis, and no additional modifications are necessary.

We will also discuss how to load (import) data into an existing setup. There are a number of steps involved in formatting and loading these data, so it is important to carefully review the steps in this chapter. For most dataset types, BenMAP-CE provides a validation tool to help you check your data file format (column names, required columns, and data types) before import. Validation reports are provided which describe any errors or warnings with the associated row number and column name.[1]

4.1   Add, Modify, and Delete a Setup

To add a new setup, modify an existing setup, or delete a setup, choose **Modify Datasets** from the menu bar. This will bring up the **Modify Datasets** window. The United States setup, which comes preinstalled with BenMAP-CE, includes a variety of datasets and looks like this:

---

[1] Validation reports are saved to C:\Users\<user name>\Documents\My BenMAP-CE Files\ValidationResults.
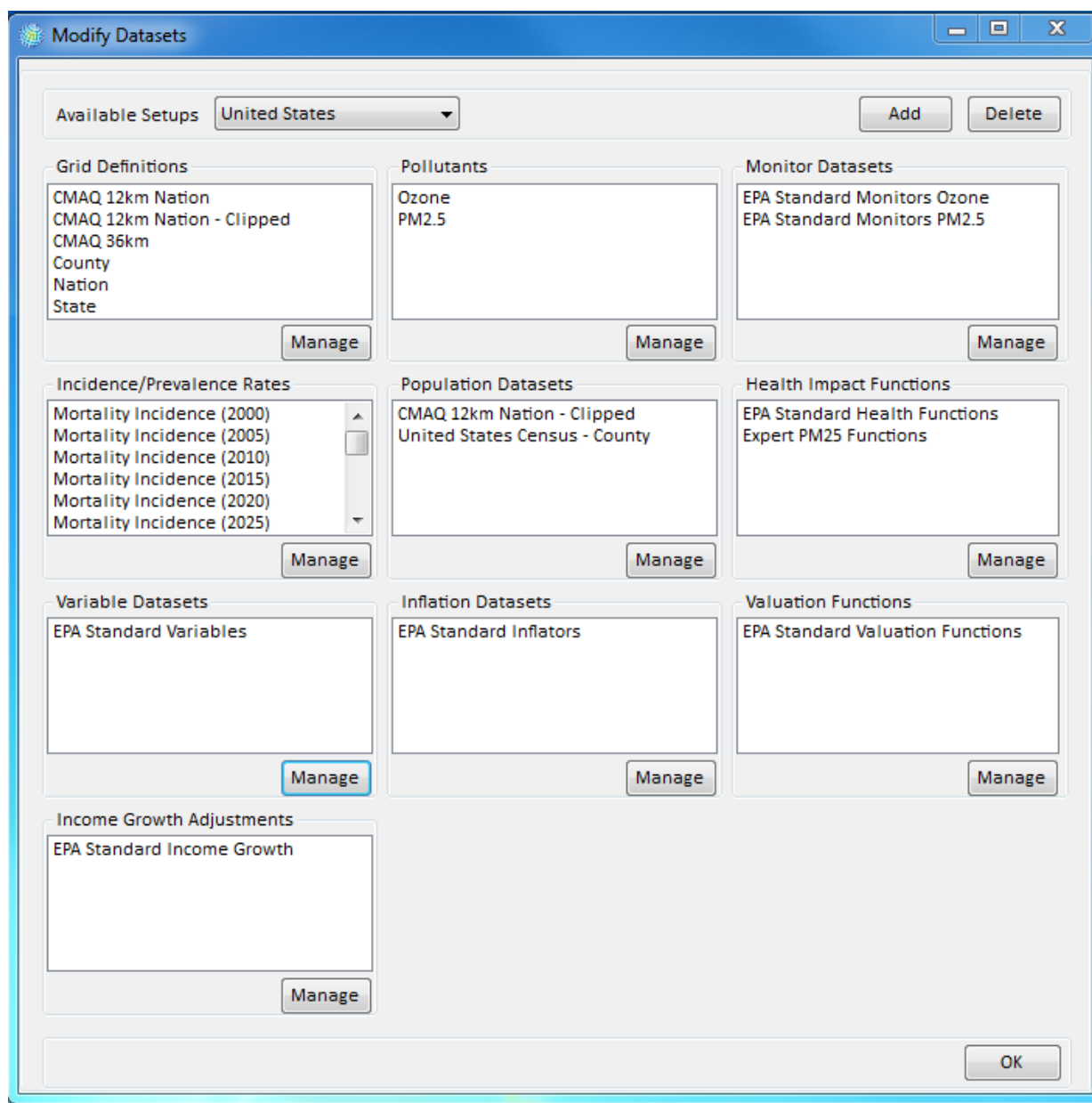
Add a Setup. To add a setup (e.g., to add a new country that was not pre-loaded with BenMAP-CE), click the **Add** button. The **New Setup** window will appear where you can type a name for the new setup.
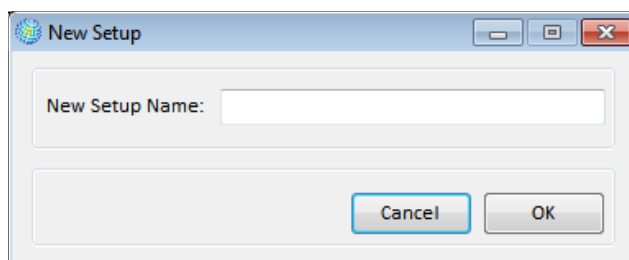
After naming the new setup, you can define the elements that comprise a setup. Table 4-1 lists the 10 dataset types within BenMAP-CE and indicates which types of data are needed to perform certain analyses.

Table 4-1. BenMAP-CE Data Elements

Dataset Type	Required to Estimate Health Impacts	Required to Quantify Economic Values
Grid Definitions	✓	✓
Pollutants	✓	✓
Monitor Datasets (or Modeled Data)	✓	✓
Incidence/Prevalence Rates	✓	✓
Population Datasets	✓	✓
Health Impact Functions	✓	✓
Variable Datasets		
Inflation Datasets		✓
Valuation Functions		✓
Income Growth Adjustments		✓

Some of the elements of a setup are fundamental and should be entered before the others, namely, **Grid Definitions** and **Pollutants**. The **Incidence/Prevalence Rates, Population**, and **Variable Datasets** depend on the **Grid Definitions**, and the **Monitor** and **Health Impact Functions** datasets depend on the **Pollutants** that you have defined. Therefore, it is best to start by defining your **Grid Definitions** and **Pollutants**, and then define the other elements of the setup.

<u>**Modify a Setup**</u>. To modify a setup, click **Modify Datasets** on the menu bar. Choose the setup for modification from the drop-down list of **Available Setups** (the default value is *United States*). Then, click on the **Manage** button under one of the ten components comprising a setup. The sections below provide more information for each of these components.

<u>**Delete a Setup**</u>. To delete a setup, click **Modify Datasets** on the menu bar. Choose the setup for deletion from the drop-down list of **Available Setups**. Click the **Delete** button. You will then be asked to confirm your decision.
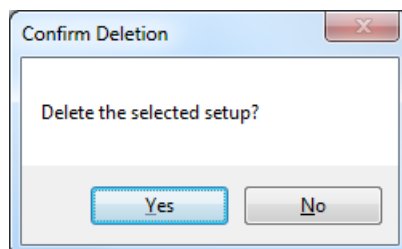
4.1.1 Grid Definitions

A BenMAP-CE **Grid Definition** specifies geographic units (i.e., grid cells) that serve two purposes: (1) the program assigns air quality data, population data, baseline incidence rates, and health impact functions to these grids in order to calculate impacts; and (2) you can use these grids to report results in the GIS tool. You can define a grid in one of two ways: by loading a Shapefile (a particular type of GIS file) or by specifying a regularly shaped grid pattern. These are referred to as **Shapefile Grid Definitions** and **Regular Grid Definitions**, respectively. A **Regular Grid Definition** is used when you want to specify a grid that is regularly shaped (e.g., 12 x 12 km squares). A **Shapefile Grid Definition** can be used to create either grids that are regularly shaped or grids that match an irregular shape, like a political boundary. All **Shapefile Grid Definitions** must contain an attribute table with a unique (i.e., non-repeating) column and row index.

At least one **Grid Definition** should be created to outline the area of interest for the BenMAP-CE analysis (a city boundary, for example). Additional grid definitions can also be created for subdivisions of that area for which (a) data are available (see the Air Monitoring, Population, Incidence and Prevalence, and Variables sections below), or (b) reports or maps are desired.

For example, an analysis for the United States might use one or more of the following grid definitions:

- **Nation** – this Shapefile Grid Definition contains an outline of the United States (just the lower forty-eight states), defining an overall area of interest.

- **State** – this Shapefile Grid Definition contains state borders, for use in generating reports and maps with results aggregated to the state level.

- **County** – this Shapefile Grid Definition contains county borders, for use with county-based population and incidence rate data.

- **CMAQ 12km Nation** – this Shapefile Grid Definition contains grid cells that are roughly 12 kilometers on each side, for use with air quality modeling data.

To start adding or modifying grid definitions, click **Modify Datasets** on the menu bar. Click on the **Manage** button below the **Grid Definitions** box. The **Manage Grid Definitions** window will appear.
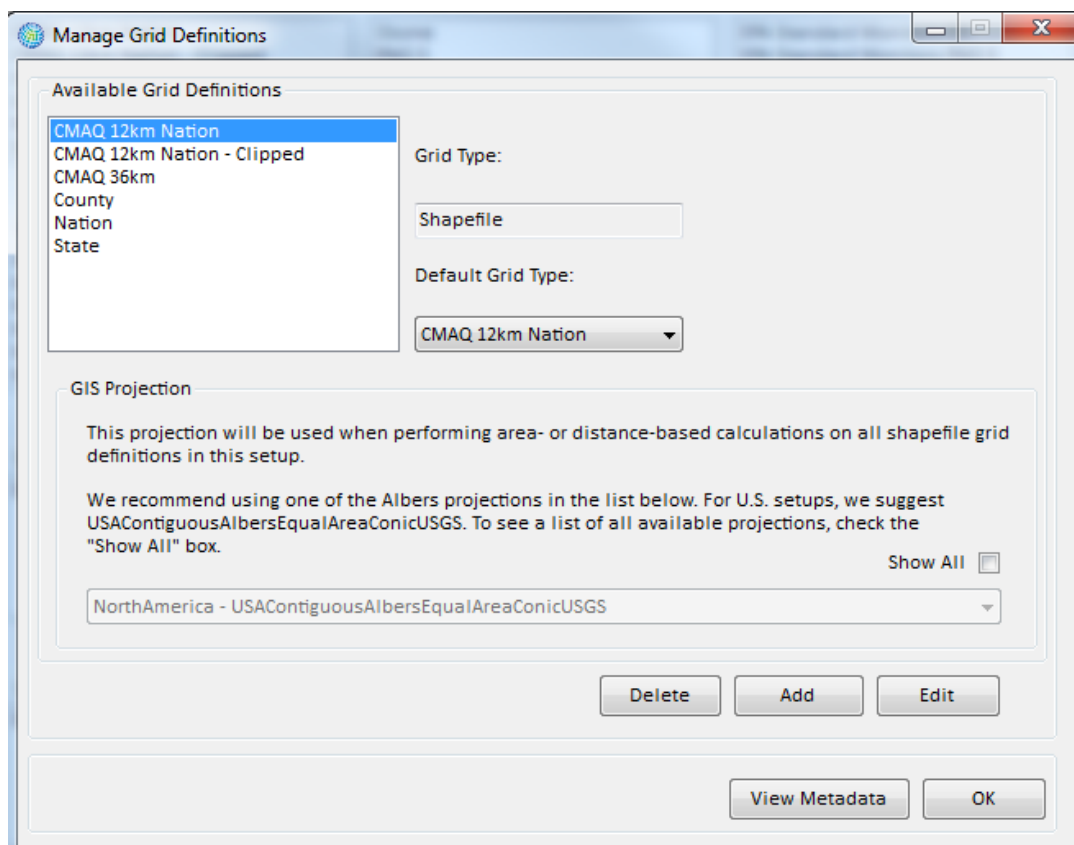
Click on the **Add** button to display the **Grid Definition** window. Provide a name for the Grid Definition in the **Grid ID** field and specify the **Grid Type**: *Shapefile Grid* or *Regular Grid*.

When you add a new Setup, you will be asked to select a GIS projection. The projection you specify will be used when performing area- or distance-based calculations on all shapefile grid definitions in this setup. We recommend selecting a regional setting that best represents the geographic area you are evaluating from one of the Albers projections provided in the list box. (For U.S. setups, we suggest USAContiguousAlbersEqualAreaConicUSGS.)
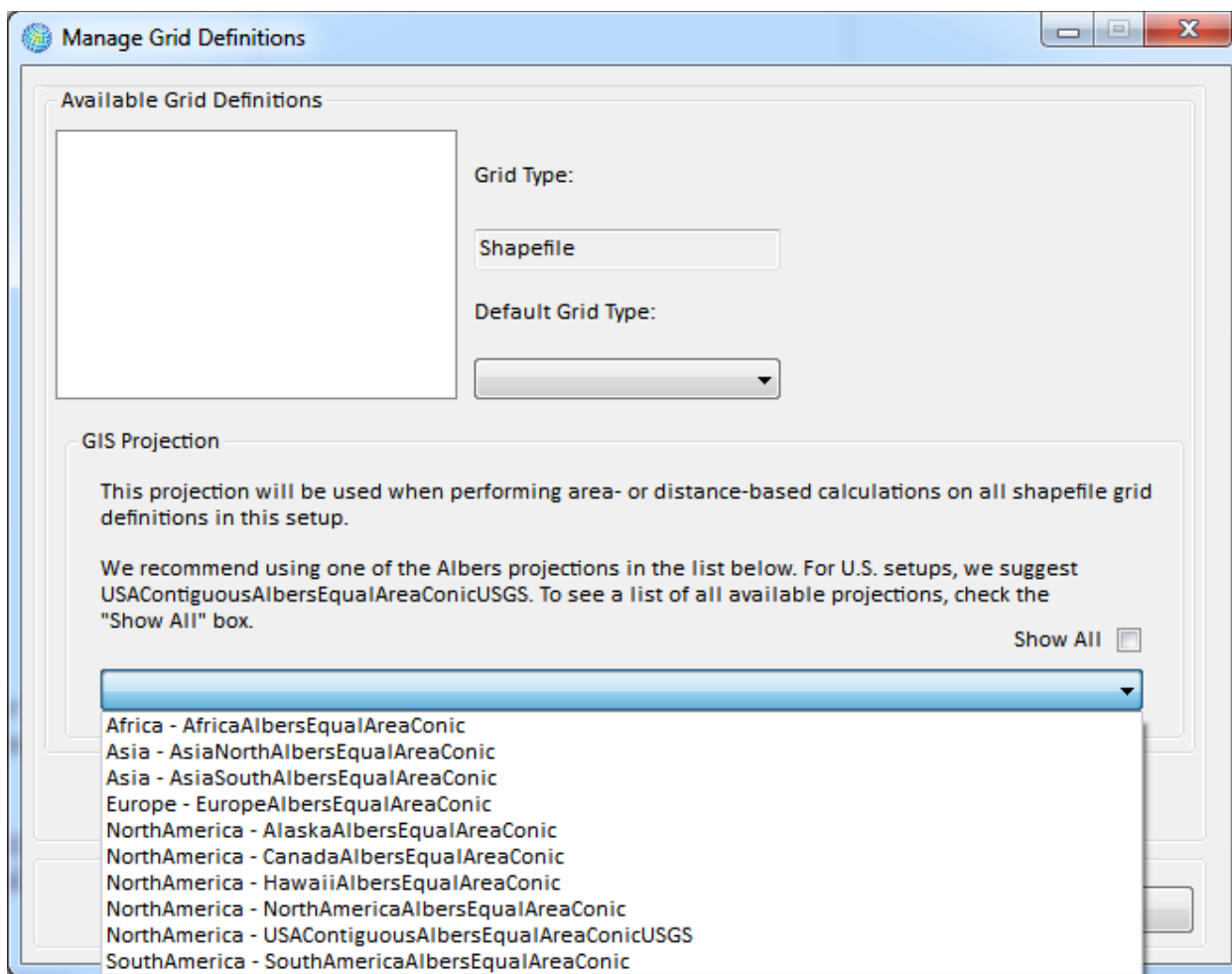
If you have reason to choose another type of projection (e.g., in consultation with a geospatial analyst), you may check the "Show All" box and select from an expanded list of choices.
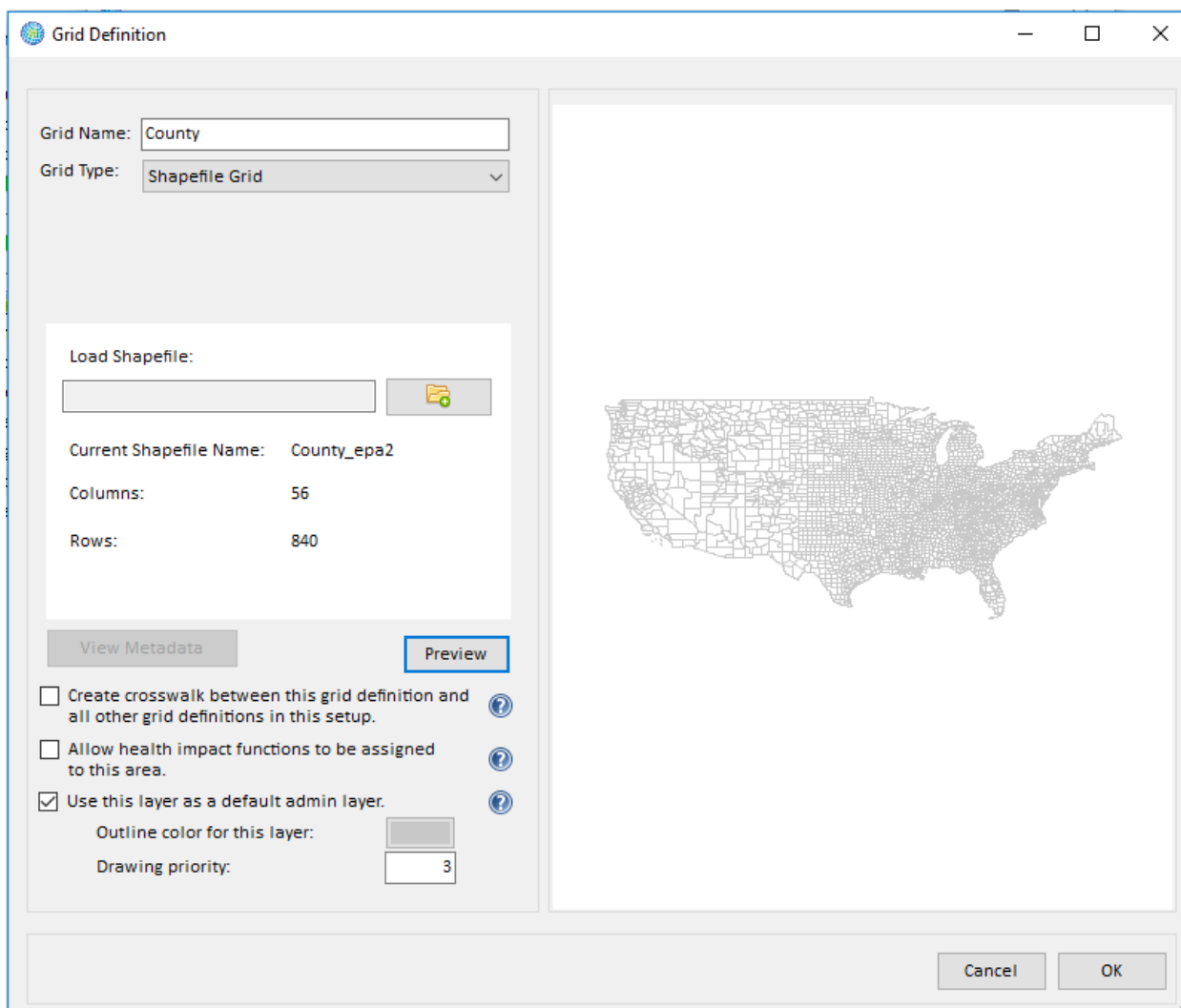
Starting with BenMAP-CE version 1.4, you can apply different health impact functions to different air quality grids within the same BenMAP run. This feature is designed to facilitate the application of region-specific air pollution effect coefficients (e.g., you can apply a Los Angeles-specific health impact function to that city and a national function to the remainder of the United States). If you plan to link specific health impact functions to a new grid definition, you must check the "Allow health impact functions to be assigned to this area" box in the lower left of the screen. Checking this box does not require you to assign a health impact function to this area, but enables this feature as an option.

Also new to version 1.4, you can now specify whether you want to use a grid definition as a default administrative (admin) layer; these layers are useful for showing political boundaries in GIS maps of BenMAP results. By checking the box "Use this layer as a default admin layer" you are telling BenMAP to draw this layer on the GIS results map by default. You may select multiple admin layers, such as a country boundary and state or province boundaries, and you can turn individual administrative layers on or off manually at any point using the GIS tool.

4.1.1.1   Regular Grid

Regular Grid Definitions are defined by a lower left corner (specified as decimal degree latitude and longitude, with West and South having negative values and East and North having positive values), a total number of columns and rows, a number of columns per degree longitude, and a number of rows per degree latitude. Individual cells within the resultant grid are numbered in sequential order (columns from left to right, rows from bottom to top) starting at (1, 1). These field values will be used to link the **Regular Grid Definition** with other sources of data, as discussed in more detail below.

To define a **Regular Grid**, start by selecting *Regular Grid* from the **Grid Type** drop-down menu. Type the name of the grid definition in the **Grid ID** box, and then define the number of **Columns** and **Rows** in the grid. To locate this grid geographically, provide the decimal degree coordinates for the lower left-hand corner of the grid in the **Minimum Longitude** and **Minimum Latitude** boxes.

To give the overall geographic size of the grid, provide the number of **Columns Per Longitude** and **Rows Per Latitude**. For example, if you specify 16 columns and 2 columns per degree longitude, then the grid will span 8 degrees of longitude. And if you specify 25 rows and 4 rows per degree latitude, then the grid will span 6.25 degrees of latitude.

Combining the numbers in this example, if the minimum longitude and latitude are -81 and 38 and the grid spans 8 degrees longitude and 6 degrees latitude, then the grid will run between -81 and -73 degrees longitude and between 38 and 44.25 degrees latitude.

After defining the grid, click the **Preview** button to see what the grid looks like. You may change the parameters and click the **Preview** again to see how the grid changes. When you are satisfied with the grid definition, click the **OK** button.

To calculate health impacts and economic benefits, BenMAP-CE uses air quality, population, and demographic data at different spatial scales. To do this, the program calculates a crosswalk (percentage file) that relates data at one spatial scale to another (e.g., 12km CMAQ grid to county). This step is performed only once per crosswalk and the results are saved to the database for subsequent calculations.

If you wish to pre-calculate the crosswalk (percentage overlap) between this grid definition and all other grid definitions in this setup, check the option box. This will take longer to load the data now, but can save you time later when you are ready to calculate health impacts. If you do not create the crosswalks at the grid definition stage, BenMAP-CE will create crosswalks as needed during the configuration or aggregation, pooling, and valuing stages.

The name of your newly defined grid will then appear in the **Manage Grid Definitions** window. You may click **Edit** to change the grid definition, **Delete** to permanently remove the grid that you just defined, or **Add** to define a new grid definition. **View**

Metadata is not applicable for **Regular Grid Definitions**. Click **OK** to return to the **Modify Datasets** window.

4.1.1.2   Shapefile Grid

Shapefiles used to create **Shapefile Grid Definitions** should be of the ESRI Shapefile format. Details on this format can be found at https://www.esri.com/library/whitepapers/pdfs/shapefile.pdf. When a shapefile is being used to create a new grid, BenMAP-CE will: (1) check to see if the file is projected to GCS NAD 83; and (2) if necessary, re-project the file (with notification) from the native projection to GCS NAD 83. Any shapefiles used must contain integer fields named Column (or Col) and Row, and each shape within the shapefile must contain a unique combination of values for these two fields. These column and row values are used, just as the Column and Row field values in Regular Grid Definitions, to link the Shapefile Grid Definition with other sources of data, as discussed in more detail below.

To add a **Shapefile Grid**, click on the **Add** button in the **Manage Grid Definitions** window, choose *Shapefile Grid* from the **Grid Type** dropdown menu, name the grid in the **Grid ID**, and browse for the correct shapefile by clicking on the small open-file icon just to the right of the **Load Shapefile** input box. After locating the file, click **Open**. This will choose the file, and bring you back to the **Grid Definition** window. To view the shapefile, click **Preview**. You can add metadata using the **View Metadata** button. Adding metadata allows you to supply a file **Reference** (e.g., person, organization, publication, or model that produced or supplied the values in the electronic file) and **Description** (or any relevant notes about the use or limitations of the data). Other minimal file attributes are pre-populated automatically during import.
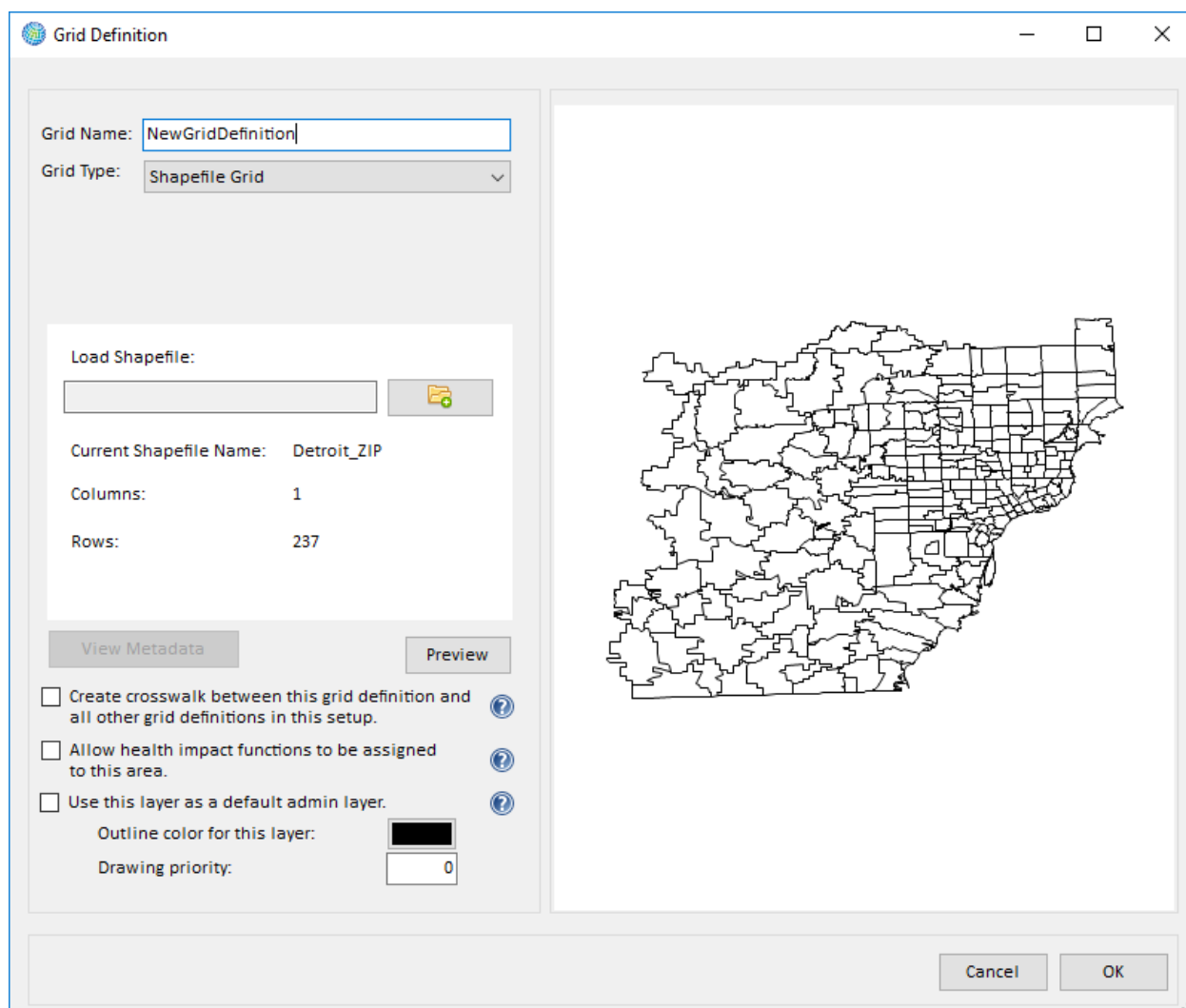
When you are satisfied that the shapefile looks correct and have checked the appropriate boxes on the lower left, click **OK**. This will save the shapefile and bring you back to **Manage Grid Definitions** window.[2] Note that the **Grid Type** box displays the type of grid for each of your grid definitions. The **View Metadata** button will allow you to see and edit any comments that were previously entered.

Click **OK** when you are finished loading grid definitions. The **Modify Datasets** screen will now list the **Grid Definitions** that you have just created. At any time, you may click the **Manage** button to add, modify, or delete grid definitions.

WARNING! If you delete a **Grid Definition,** you will permanently delete any gridded data that is dependent on it, such as any **Incidence/Prevalence, Population**, and

---

[2] Shapefiles are saved to C:\Users\<user name>\AppData\Local\BenMAP-CE\Data\Shapefiles\<setup name>\.

Variable Datasets that use this particular **Grid Definition**. As we discuss each of these other setup elements below, we will describe how this might happen.

4.1.2   Pollutants

The **Pollutants** section of a setup specifies the pollutants that BenMAP-CE will analyze and defines the air quality metrics to be used by BenMAP-CE. You are not importing air pollution data, but rather naming your pollutants and defining the measures or metrics BenMAP-CE will use when performing an analysis for each pollutant. You may include any pollutant, though typically air pollutants such as particulate matter, ozone, sulfur dioxide, and carbon monoxide are used in a BenMAP-CE analysis.

A key concept for pollutants is the **Metric**. Air quality metric describes the period of the day over which the pollutant observations are averaged. For example, a metric of *D24HourMean* is a daily average of hourly measurements. A metric of *D8HourMax* is the average of the 8-hour period during the day when pollutant levels are the highest (see Table 4-2 below). The air quality change must be expressed in a metric that matches the metric used by the health impact function; this concept is discussed further below.

In general, air pollution data in BenMAP-CE is hierarchical – a pollutant can have multiple **Metrics**, each of which has multiple **Statistics** (these are automatically calculated by BenMAP-CE) and which can have multiple **Seasonal Metrics**. Similarly, **Seasonal Metrics** have multiple **Statistics**. Furthermore, air pollution data can be provided to BenMAP-CE at any of these levels, in addition to the daily and hourly observation level, as described in more detail in Section 4.3.

Table 4-2. Example Calculation of D8HourMax

Hourly Period[1]	Hourly average O$_3$ concentration[2] (ppm)	Moving 8-hour average[3]	D8HourMax[4]
07:00	0.000	---	
08:00	0.005	---	
09:00	0.010	---	
10:00	0.015	---	
11:00	0.020	---	
12:00	0.025	---	
13:00	0.030	---	
14:00	0.035	---	
15:00	0.040	0.020	
16:00	0.045	0.025	
17:00	0.050	0.030	
18:00	0.055	0.035	
19:00	0.060	0.040	
20:00	0.055	0.044	
21:00	0.050	0.047	
22:00	0.045	0.048	
23:00	0.040	**0.049**	← **0.049**
00:00	0.035	0.048	
01:00	0.030	0.047	
02:00	0.025	0.044	
03:00	0.020	0.040	
04:00	0.015	0.035	
05:00	0.010	0.030	
06:00	0.000	0.024	

[1] Days for measuring ozone start and end at 7:00 AM local standard time.
[2] Hourly average is the average of individual measurements taken during the hour.
[3] Moving 8-hour average is the average of the hour and the proceeding 7 hours.
[4] *D8HourMax* is the maximum of the moving 8-hour averages.
Reference: Federal Register 79 FR 75233, Dec. 17, 2004.
https://www.federalregister.gov/articles/2014/12/17/2014-28674/national-ambient-air-quality-standards-for-ozone#h-193

Add a Pollutant

Air pollution data in BenMAP-CE is of two types: (1) point source monitoring data and (2) **Grid Definition**-based modeling data. For both types, the data must be associated with a particular pollutant. Table 4-3 describes these variables used to define a pollutant in BenMAP-CE.

Table 4-3. BenMAP-CE Pollutant Definitions

Pollutant Field Name	Notes
Pollutant ID	Unique name for the pollutant which will be referenced in health impact functions, associated with monitoring and modeling data, etc.
Observation Type	Pollutants may have hourly observations or daily observations. In the United States, Ozone has hourly observations, while PM_{10} and $PM_{2.5}$ have daily observations.
Metrics	Daily values calculated directly from daily observations, or through various mathematical manipulations of hourly observations. Typical ozone metrics include the highest hourly observations during the course of each day, the mean of all twenty four hourly observations, etc.
Seasonal Metrics	Seasonal values calculated from metric values. In the United States, for example, quarterly means are calculated for $PM_{2.5}$ from daily means.

To add pollutant definitions to BenMAP-CE, click **Modify Datasets** on the menu bar. Then click on the **Manage** button below the **Pollutants** group. The **Manage Pollutants** window will appear. Here you may click **Add** to add a new **Pollutant**, **Delete** to remove a previously defined pollutant, or **Edit** to modify an existing pollutant.
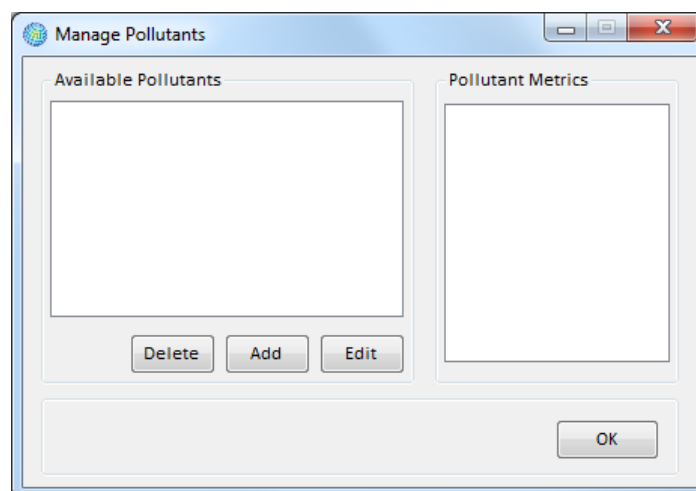
To start defining a pollutant, click the **Add** button and the **Pollutant Definition** window will appear. In the **Pollutant ID** box, you give a unique name for the pollutant (e.g., PM$_{2.5}$), and then define the characteristics of this pollutant – the **Observation Type** and **Metrics**.

The **Observation Type** identifies whether a pollutant is measured *Hourly* or *Daily*. In the United States, ozone, sulfur dioxide, carbon monoxide, and others have hourly observations, while particulate matter has daily observations.

Next you need to define a pollutant's **Metrics**. A pollutant has to have one or more metrics, which are daily values calculated directly from daily observations, or through various mathematical manipulations of hourly observations.

To add a **Metric**, click on the **Add** button below the **Metrics** box in the **Pollutant Definition** window. A default name 'Metric 0' will appear in the box. Since the default name is not very descriptive of a metric, it is best to change the name. Typical names used for metrics given in Table 4-4. These are provided just as an example, you may use any names that you like. However, keep in mind that the names that you use for your

metrics need to be consistent with the metric names that you include in your air pollution monitoring and modeling data, as well as your health impact functions. (We will discuss this further below.) Additionally, metric names are used to display pollutant concentrations in BenMAP-CE's mapping window. As such, they must be consistent with GIS naming conventions, meaning they must begin with a letter, and may only contain letters, numbers, and underscores.

Table 4-4. Examples of Metric Names

Name	Description
D1HourMax	Highest hourly value from 12:00 A.M. through 11:59 P.M.
D8HourMax	Highest eight-hour average calculated between 12:00 A.M. and 11:59 P.M.
D24HourMean	Average of hours from 12:00 A.M. through 11:59 P.M.

4.1.2.1   Hourly Metrics

Pollutants that are measured hourly (**Observation Type** = *Hourly*), such as ozone, sulfur dioxide, carbon monoxide, and others, must be characterized by a daily metric, which mathematically summarizes the hourly observations.

Table 4-4 lists some of the ways that metrics can be generated from hourly values. Note that these metrics are not arbitrarily chosen, and instead match the metrics used in epidemiological studies.

The **Detail** section of the **Pollutant Definition** window lets you define the metrics that you want to use. There are three options that you may choose using the **Hourly Metric Generation** drop-down list: *Fixed Window*, *Moving Window*, and *Custom*.

The **Fixed Window** option lets you define simple metrics which are calculated as statistics over a fixed window of hours (**Start Hour** and **End Hour**) within each day. The **Start Hour** should be less than or equal to the **End Hour**, and both can range from 0 to 23, where 0 stands for the period 12:00 am to 12:59 am, and 23 stands for 11:00 pm to 11:59 pm. The **Statistic** includes the *Mean*, *Median*, *Max*, *Min*, and *Sum.* Some examples follow:

- **D24HourMean**: The mean of the observations from 12:00 am through 11:59 pm. **Start Hour** = *0*. **End Hour** = *23*. **Statistic** = *Mean*.

- **D1HourMax**: The highest hourly value of the observations from 12:00 am through 11:59 pm. **Start Hour** = *0*. **End Hour** = *23*. **Statistic** = *Max*.

- **D12HourMean**: The mean of the daylight observations, defined as the period from 8:00 am through 7:59 pm. **Start Hour** = *8*. **End Hour** = *19*. **Statistic** = *Mean*.

The **Moving Window** option lets you consider metrics that are not based on the same set of hours each day. The **Window Size** defines the number of hours that will be considered together. The **Window Statistic** defines how the hours in the **Window Size** will be characterized. And the **Daily Statistic** defines how BenMAP-CE will use the statistics generated for each window.

For example, consider the highest eight-hour mean (*D8HourMax*) over the course of a day. You would have the following settings: **Window Size** = *8*. **Window Statistic** = *Mean*. **Daily Statistic** = *Max*. BenMAP-CE would calculate every possible eight-hour mean, starting with the eight-hour mean from 12:00 am through 7:59 am, and ending with the eight-hour mean from 4:00 pm through 11:59 pm. This would generate 17 possible eight-hour means. BenMAP-CE would then choose the eight-hour mean that has the highest value (see Table 4-2 above for an example).

The **Custom** tab lets you define **Metrics** using a mathematical function that you specify. These functions can include measures such as the sum of the number of hours of ozone exposure above 60 ppb. The possibilities are quite diverse, as evidenced by the range of functions and variables available for use as shown in Table 4-5. However, the syntax for using these functions is somewhat involved, so we have reserved discussion of this for the Appendix M: Function Editor.
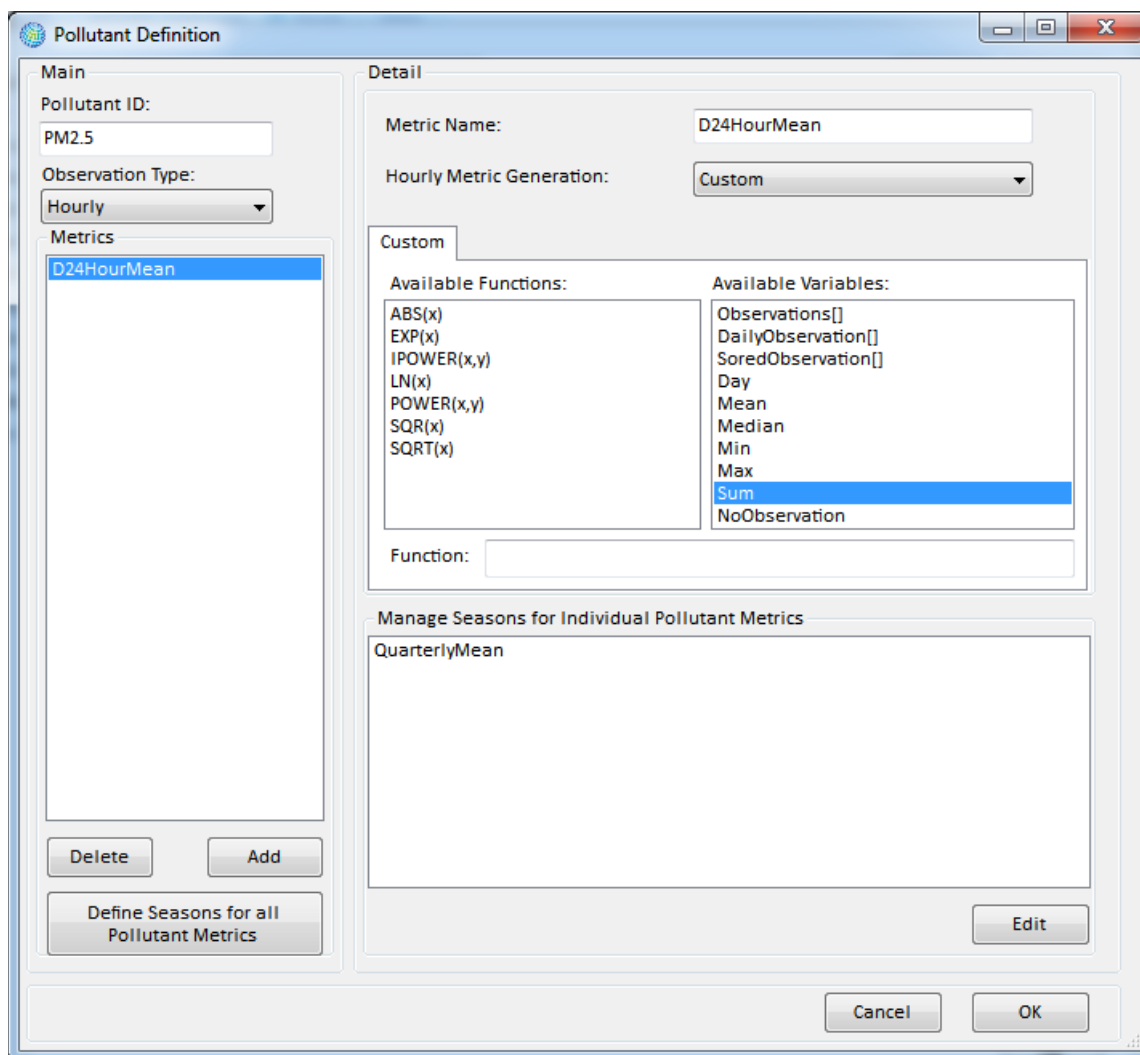
Table 4-5. Available Functions and Variables for Custom Metrics

Name	Description
Functions	
ABS(x)	Returns the absolute value of *x.*
EXP(x)	Returns *e* the power *x*, where *e* is the base of the natural logarithm.
IPOWER(x,y)	Returns *x* to the power *y* (*y* an integer value).
LN(x)	Returns the natural logarithm of *x.*
POWER(x,y)	Returns *x* to the power *y* (*y* a floating point value).
SQR(x)	Returns the square of *x.*
SQRT(x)	Returns the positive square root of *x.*
Variables	
Observations[i]	All hourly observations for the year (index begins at zero, typically ranging to 8,760).
DailyObservations[i]	All hourly observations for the day (indexed zero to twenty-three).
SortedObservations[i]	All hourly observations for the day, sorted from low to high (indexed zero to twenty-three).
Day	Index of the day whose metric value is being generated (index begins at zero).
Mean	Mean of the daily observations.
Median	Median of the daily observations
Min	Minimum of the daily observations.
Max	Maximum of the daily observations.
Sum	Sum of the daily observations.
NoObservation	Flag value indicating a missing observation (-345)

4.1.2.2 Manage Seasons for Individual Pollutant Metrics

Manage Seasons for Individual Pollutant Metrics allow you to aggregate daily **Metric** values over a portion of the year that you define. This has a number of uses. For example, if pollutant values vary greatly by season of the year, you can calculate separate pollutant measures for each season of interest. You might be interested in *Dry Season* versus *Wet Season*, or differences between *Winter*, *Spring*, *Summer*, and *Fall*.

To add seasonal metrics for individual pollutant metrics, first select the metric of interest (e.g., D24HourMean), then click on the **Edit** button below the **Manage Seasons for Individual Pollutant Metrics** box. The **Manage Seasons for Individual Pollutant Metrics** window will appear.

To add a **Seasonal Metric**, click on the **Add** button below the **Seasonal Metrics** box. A default name *Seasonal Metric 0* will appear in the box. Since the default name is not very

descriptive, it is best to change the name to something more informative such as the *QuarterlyMean*. As with the **Metric** names, keep in mind that the **Seasonal Metric** names need to be consistent with the metric names that you include in your air pollution monitoring and modeling data, as well as your health impact functions.

The next step is to define the seasons that you want associated with your **Seasonal Metric** name. For example, in the case of a *Quarterly Mean*, you would want to define four seasons. To start this process, click on the **Add** button below the **Seasonal Metric Seasons** box. Then, in the far right side of the window under **Selected Season Details**, give the **Start Date** and **End Date** for each season. To change the date, click on the month or day, and use the arrows to change the month or day accordingly.

Next, you need to choose the **Statistic** tab or the **Custom Function** tab to determine how the daily **Metrics** will be combined in each season. For example, you might choose the *Mean* from the drop-down list on the **Statistics** tab. This would calculate the mean of the daily metrics in each season. The **Custom** tab allows seasonal metric values to be calculated using customized functions, similar to those used to calculate daily metric values from hourly observations. See Appendix M: Function Editor for more detail on this topic.

Once you have finished defining the **Seasonal Metrics**, click **OK** to return to the **Pollutant Definition** window.

Click **OK** after defining each **Pollutant**. This will return you to the **Manage Pollutants** window.
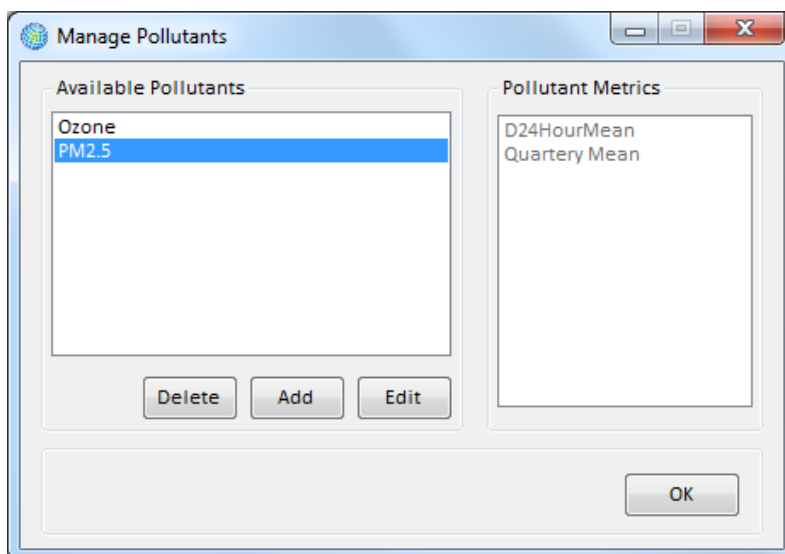
4.1.2.3 Define Seasons for all Pollutant Metrics

The **Define Seasons for all Pollutant Metrics** button on the **Pollutant Definition** window allows you to associate **Seasons** with a **Pollutant** and all of its associated metrics. These seasons differ somewhat from the **Seasonal Metrics** defined for individual pollutant metrics (discussed above). They are used to define:

- The portion of the year for which benefits are calculated for a **Pollutant**. You can think of **Seasons** as defining this period of the year "globally" for the pollutant, as it affects the portion of the year over which both **Metrics** and **Seasonal Metrics** are calculated. For example, in the United States ozone benefits are often only calculated for the ozone season, from May 1 through September 30.

- The portion(s) of the year for which missing pollutant concentrations are filled in by BenMAP-CE. That is, in order to calculate benefits, BenMAP-CE in certain cases needs to generate complete sets of metric values by estimating concentrations for those days that have missing observations. This can be important if certain seasons tend to have more missing values than others.

To define **Seasons** for a **Pollutant**, click the **Define Seasons for all Pollutant Metrics** button. This will bring up the **Define Seasons** window. For each season desired, click the **Add** button, select the appropriate **Start Date** and **End Date**, which define the days included in the season; and the appropriate **Start Hour** and **End Hour**, which define the hours included in monitoring period. The advanced options for PM$_{2.5}$ look like the following:
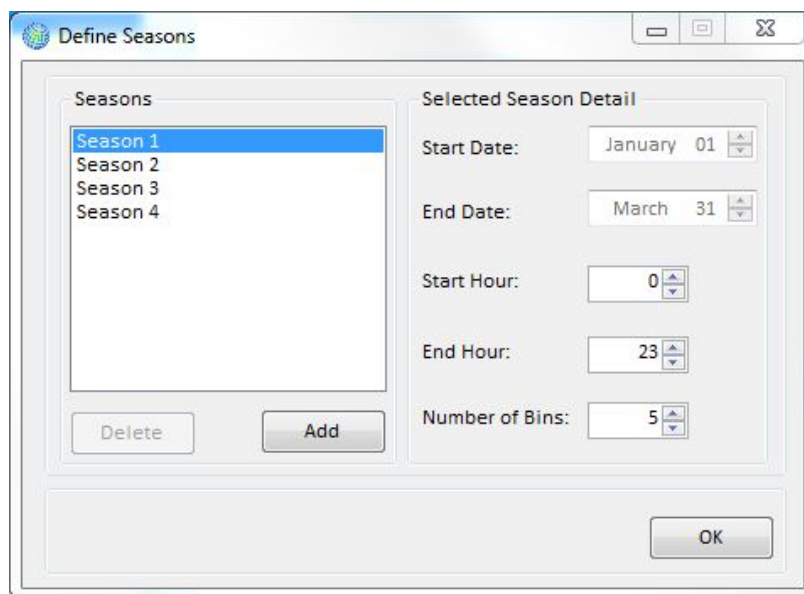
Once you have finished defining the **Seasons**, click **OK** to return to the **Pollutant Definition** window.

If you later wish to View or Edit a particular **Pollutant** definition, simply select the appropriate **Pollutant** within the **Available Pollutants** box and click the **Edit** button. When you are done, click **OK** to return to the **Modify Datasets** window.

After defining all of the pollutants that you want, click **OK**. This will return you to the **Modify Datasets** window.

4.1.3   Monitor Datasets

The **Monitor Datasets** section of the **Modify Datasets** window allows you to add air pollution monitoring data to your setup. Air pollution monitoring data may be used to estimate ambient pollution levels in each grid cell defined by a **Grid Definition**. BenMAP-CE uses a variety of procedures (such as *Voronoi Neighbor Averaging*, discussed later) to interpolate the monitor data points across the area of interest.

NOTE: Air pollution data in BenMAP-CE is of two types: (1) point source monitoring data and (2) **Grid Definition**-based modeling data. Both types of data must be associated with a particular pollutant that you have already defined. Only the point source monitoring data is stored in the setup database. The modeling data are loaded into BenMAP-CE as you need them for a particular analysis.

4.1.3.1   Add Monitor Datasets

To start, click on the **Manage** button below the **Monitor Datasets** box. The **Manage Monitor Datasets** window will appear. From this window you may **Add** monitoring data, view and **Edit** existing datasets, as well as **Delete** them. The section on the left

under **Available Datasets** lists the monitor datasets that are currently in the setup. The section on the right under the **Dataset Contents** identifies the number of monitors in each dataset by *Pollutant* and by *Year*. To view the metadata of a particular monitor dataset, select an **Available Dataset** and click on a row from the **Dataset Contents**, then click the **View Metadata** button. This allows you to view further information about references or descriptions of the file.
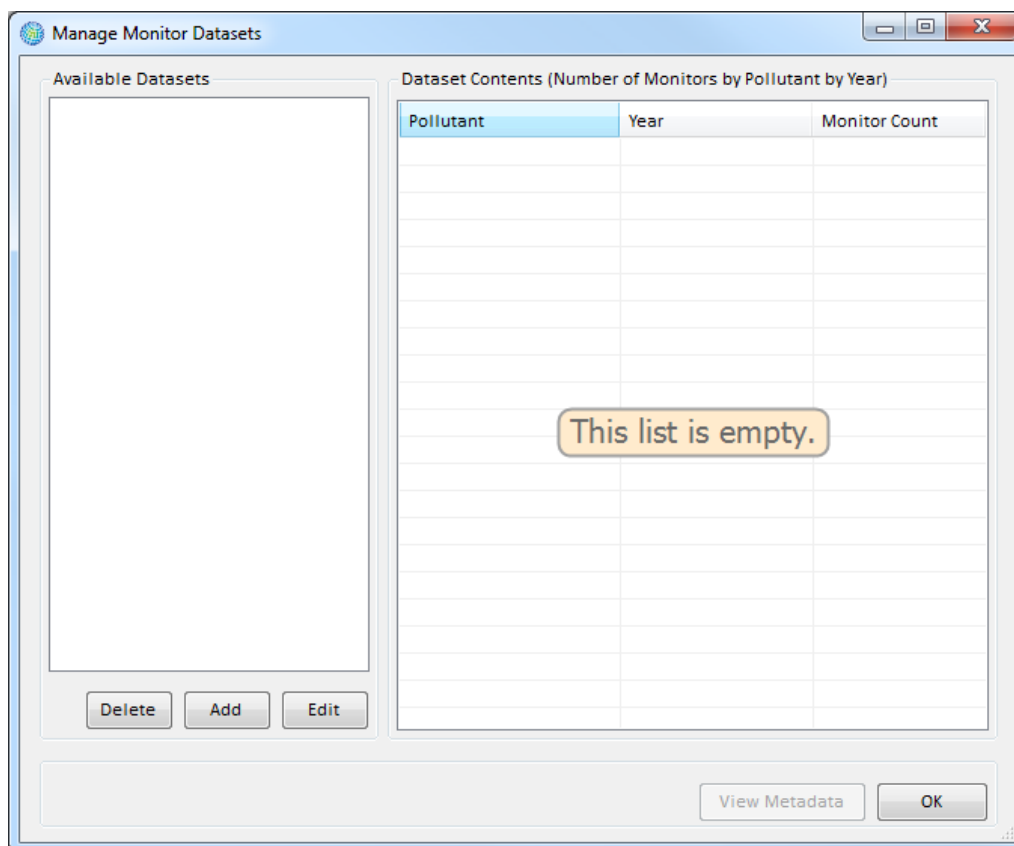
To start adding data, click the **Add** button. This will bring up the **Monitor Dataset Definition** window. Give the dataset a name in the **Dataset Name** box, choose the appropriate pollutant from the **Pollutant**[3] drop-down menu, and then type the 4-digit *Year* of the data in the **Year** box.

NOTE: The **Dataset** that you define can have one or more pollutants and multiple years of data (e.g., representing a particular monitoring network). However the data must be imported one pollutant and one year at a time.

---

[3] The pollutants in the **Pollutant** drop-down menu have been defined under the **Pollutants** box on the **Modify Datasets** window.

Monitor data must be formatted in a database file, with monitor definition information and monitor values in a single line.

After specifying the **Pollutant** and **Year**, click on the **Load Data From File** button to bring up a window from which you can **Browse** the BenMAP-CE Data directory to find the desired data file. Click **Open,** to choose the file.

Users are advised to click the **Validate** button before loading the monitor dataset. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings.[4]   The user also has the option of adding metadata to the file to save references and comments about that specific file. To add metadata, click the **View Metadata** button. After passing validation (and adding metadata if desired) click the **OK** button to bring you back to the **Monitor Dataset Definition** window.

---

[4] Validation reports are saved to C:\Users\<user name>\Documents\My BenMAP-CE Files\ValidationResults.

Repeat this procedure to load all of your monitoring data. If you try to load the data for the same pollutant/year combination, BenMAP-CE will warn you of the duplication.
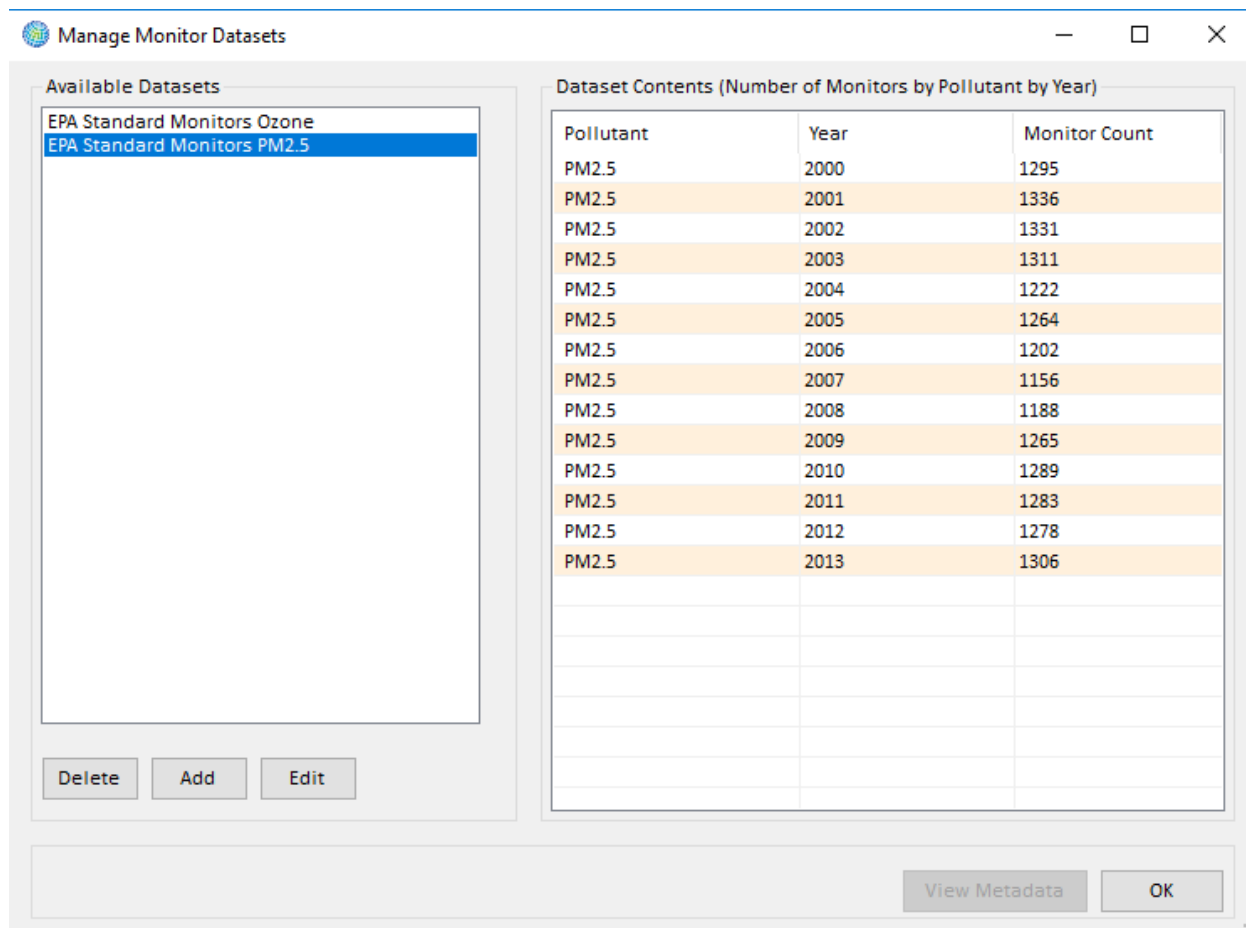
To see the years of data and the number of monitors each year, use the scrollbars on the bottom and on the right of the **Dataset Contents** box. To view the metadata for a particular entry in the dataset, choose an **Available Dataset** and row in the **Dataset Contents** and click the **View Metadata** button.

To delete existing datasets, select the dataset in the **Available Datasets** list and click the **Delete** button. To edit an existing dataset, select the dataset in the **Available Datasets** list and click the **Edit** button. (Note: Certain pre-loaded datasets cannot be edited. Instead if you select one of these, you will have the option to **Copy** the locked dataset and then you can edit the copied dataset.)

When you have finished loading your monitor data, click **OK** in the **Manage Monitor Datasets** window. This will take you back to the **Modify Datasets** window, which will show the name of the **Dataset**(s) that you just entered.

4.1.3.2   Format for Monitor Data

Monitor data is required to be formatted in a single database file, with monitor definition information and monitor values in a single line. Tables 4-6a and 4-6b list the variables in the monitor dataset and provide a sample of what a data file might look like.

NOTE: The monitor data files do not specify the pollutant with which the data is associated—this is specified by the user when loading the monitor data into BenMAP-CE.

Table 4-6a. Required Format, Air Monitoring Data File Variables

Variable	Type	Required	Notes
Monitor Name	Text	Yes	Unique name for each monitor in a particular location.
Description	Text	No	Description of the Monitor.
Longitude	Numeric (double)	Yes	Values should be in decimal degree format. Values in the eastern hemisphere are positive, and those in the western hemisphere are negative.
Latitude	Numeric (double)	Yes	Values should be in decimal degree format. Values in the northern hemisphere are positive, and those in the southern hemisphere are negative.
Metric	Text	No	This variable is either blank (signifying that the Values are Observations, rather than Metric values), or must reference an already defined Metric (e.g., 1-hour daily maximum) for the appropriate Pollutant.
Seasonal Metric	Text	No	This variable is either blank (signifying that the Values are not Seasonal Metric values) or must reference an already defined Seasonal Metric for the Metric (e.g., mean of the 1-hour maximum values for the months of June through August).
Statistic	Text	No	This is an annual metric, which is either blank (signifying that the values are not annual statistics) or must be one of: None, Mean, Median, Max, Min, Sum. (e.g., mean of the 1-hour maximum for the year)
Values	Text	Yes	If Metric is blank, values are supplied as a comma-delimited string of values for the year [e.g., 365 or 366 (leap year) values for daily data, 8760 or 8784 (leap year) values for hourly data]. If Metric is defined, but Seasonal Metric and Statistic are blank, 365 or 366 metric values. If Seasonal Metric is defined, but Statistic is blank, n seasonal metric values. If Statistic is defined, one annual statistic value for either the Metric (if Seasonal Metric is blank) or the Seasonal Metric. Missing values are signified with a period ('.').

Table 4-6b. Required Format, Sample Air Monitoring Data File

Monitor Name	Monitor Description	Latitude	Longitude	Metric	Seasonal Metric	Statistic	Values
260050003881011	'RESIDENTIAL, SUBURBAN'	42.7678	-86.1486	D24HourMean	QuarterlyMean	Mean	11.67
260170014881011	'RESIDENTIAL, URBAN AND CENTER CITY'	43.5714	-83.8907	D24HourMean	QuarterlyMean	Mean	11.62
260210014881011	'COMMERCIAL, RURAL'	42.1978	-86.3097	D24HourMean	QuarterlyMean	Mean	11.58
260490021881011	'RESIDENTIAL, URBAN AND CENTER CITY'	43.0472	-83.6702	D24HourMean	QuarterlyMean	Mean	11.56
260650012881011	'RESIDENTIAL, URBAN AND CENTER CITY'	42.7386	-84.5346	D24HourMean	QuarterlyMean	Mean	11.49
260770008881011	'COMMERCIAL, URBAN AND CENTER CITY'	42.2781	-85.5419	D24HourMean	QuarterlyMean	Mean	11.44
260810007881011	'INDUSTRIAL, URBAN AND CENTER CITY'	42.9561	-85.6791	D24HourMean	QuarterlyMean	Mean	11.4
260810020881011	'INDUSTRIAL, URBAN AND CENTER CITY'	42.9842	-85.6713	D24HourMean	QuarterlyMean	Mean	11.36
260990009881011	'COMMERCIAL, SUBURBAN'	42.7314	-82.7935	D24HourMean	QuarterlyMean	Mean	11.43
261010922881011	'RESIDENTIAL, URBAN AND CENTER CITY'	44.3070	-86.2426	D24HourMean	QuarterlyMean		11.43,11.07,10.98,10.90
261130001881011	'FOREST, RURAL'	44.3106	-84.8919	D24HourMean	QuarterlyMean		11.44,14.23,11.20,9.30
261150005881011	'AGRICULTURAL, RURAL'	41.7639	-83.4719	D24HourMean	QuarterlyMean		11.39,13.71,10.04,11.24
261210040881011	'COMMERCIAL, URBAN AND CENTER CITY'	43.2331	-86.2386	D24HourMean	QuarterlyMean		11.28,9.23,15.03,11.10
261250001881011	'RESIDENTIAL, SUBURBAN'	42.4631	-83.1832	D24HourMean	QuarterlyMean		11.27,11.12,13.15,12.03
261390005881011	'RESIDENTIAL, SUBURBAN'	42.8945	-85.8527	D24HourMean	QuarterlyMean		11.27,10.05,10.9,12.13
261470005881011	'RESIDENTIAL, SUBURBAN'	42.9533	-82.4562	D24HourMean	QuarterlyMean		11.26,12.13,15.00,10.01
261610008881011	'COMMERCIAL, URBAN AND CENTER CITY'	42.2406	-83.5996	D24HourMean	QuarterlyMean		11.24,14.05,11.03,10.10
261630001881011	'COMMERCIAL, SUBURBAN'	42.2286	-83.2082	D24HourMean	QuarterlyMean		11.23,9.23,13.04,10.09
261630015881011	'COMMERCIAL, URBAN AND CENTER CITY'	42.3028	-83.1065	D24HourMean	QuarterlyMean		11.23,11.24,10.45,10.04
261630016881011	'RESIDENTIAL, URBAN AND CENTER CITY'	42.3578	-83.0960	D24HourMean	QuarterlyMean		11.2,10.73,11.23,10.45
261630019881011	'RESIDENTIAL, SUBURBAN'	42.4308	-83.0001	D24HourMean			.,.,13.01,.,.,17.19,.,.,3.89,.,.,13.86,.,.,5.9
261630025881011	'COMMERCIAL, SUBURBAN'	42.4231	-83.4263	D24HourMean			.,.,13.48,.,.,8.73,.,.,6.079,.,.,14.534,.,.,3.
261630033881011	'INDUSTRIAL, SUBURBAN'	42.3067	-83.1488	D24HourMean			.,.,15.675,.,.,12.72,.,.,6.46,.,.,16.24,.,.,6.
261630036881011	'COMMERCIAL, SUBURBAN'	42.1873	-83.1539	D24HourMean			.,.,16.52,.,.,15.01,.,.,5.41,.,.,15.39,.,.,5.0
261630038881011	'RESIDENTIAL, URBAN AND CENTER CITY'	42.3350	-83.1096	D24HourMean			.,.,20.52,.,.,21.85,.,.,5.88,.,.,19.28,.,.,5.1
261630039881011	'RESIDENTIAL, URBAN AND CENTER CITY'	42.3233	-83.0685	D24HourMean			.,.,20.14,.,.,7.59,.,.,8.17,.,.,13.11,.,.,7.59

4.1.4 Incidence and Prevalence Rates Data

Most health impact functions, such as those developed from log-linear or logistic health impact functions, estimate the percent change in a health effect associated with a pollutant change. In order to estimate the absolute change in incidence using these functions, the baseline incidence rates (and in some cases the prevalence rate) of the adverse health effect are needed.

The incidence rate is the number of health effects per person in the population per unit of time, and the prevalence rate is the percentage of people that suffer from a particular chronic illness. For example, the incidence rate for asthma attacks may be 25 cases per asthmatic individual per year, and the prevalence rate (measuring the percentage of the population that is asthmatic) might be six percent of the total population.

NOTE: For both incidence and prevalence rates, BenMAP-CE allows the user to have rates that vary by race, ethnicity, gender, and age group. BenMAP-CE can support multiple sets of incidence and prevalence rates, if the rates differ by year or by grid definition.
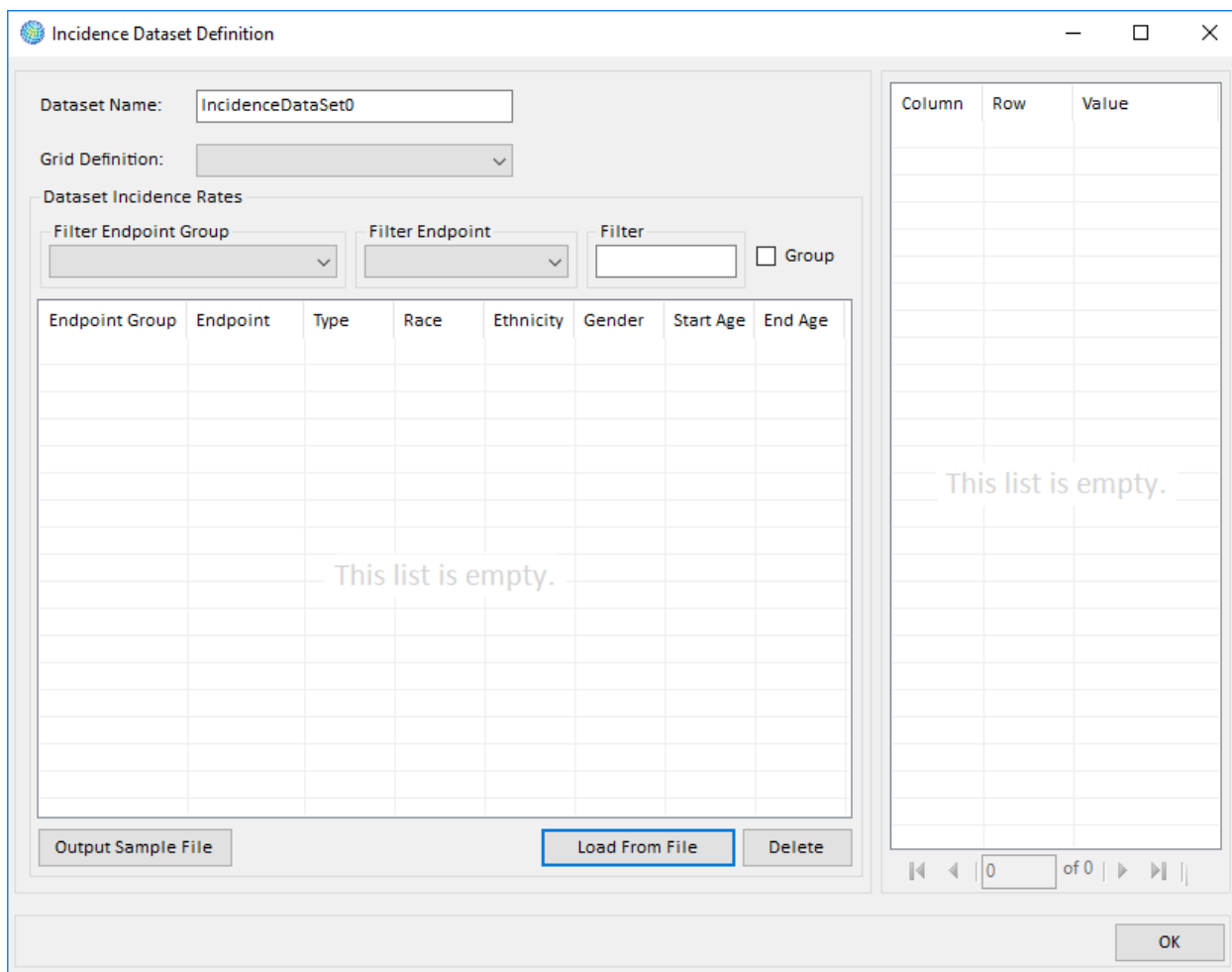
4.1.4.1 Add Incidence/Prevalence Rates

To start adding incidence and prevalence data files, click on the **Manage** button below the **Incidence/Prevalence Rates** box. The **Manage Incidence Datasets** window will appear.

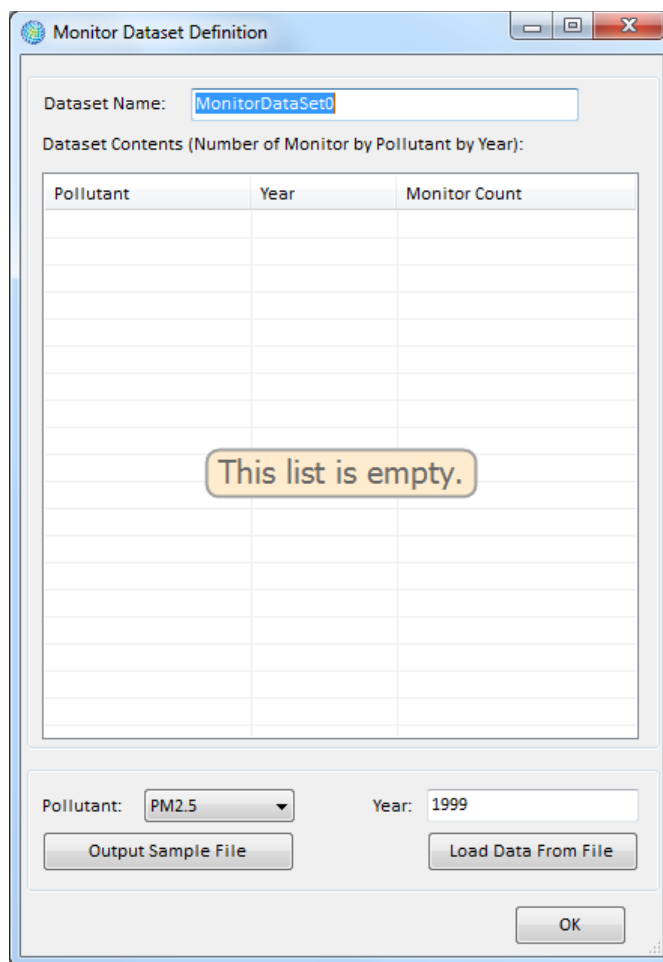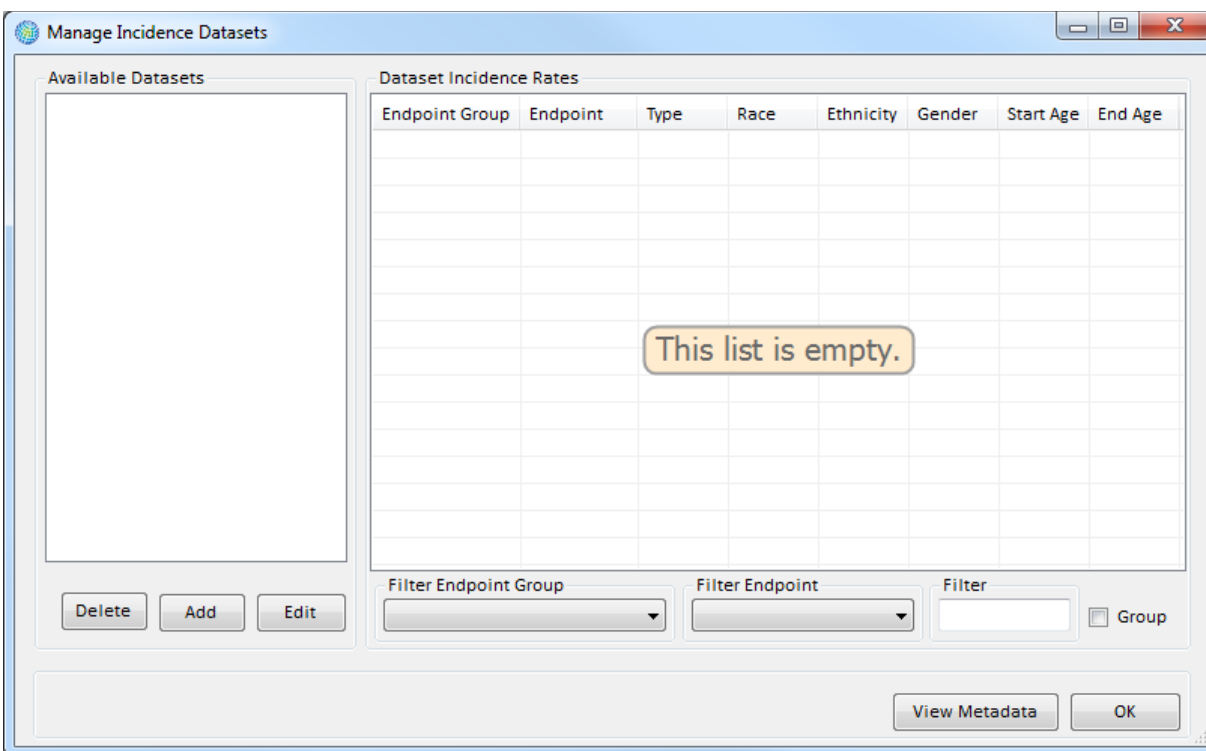In this window you may **Add**, **Edit**, and **Delete** datasets. The section on the left under **Available Datasets** lists the incidence/prevalence datasets that are currently in the setup. The section on the right under the **Dataset Incidence Rates** identifies the rates in the selected dataset.

To add a dataset, click the **Add** button. This will bring up the **Incidence Dataset Definition** window. Give a name to the dataset that you are creating by typing a name in **Dataset Name** box.

NOTE: If you have multiple incidence or prevalence datasets that vary, for example, by year and grid definition, then use the name to provide a reference to the year and grid definition (e.g., "Mortality Incidence (2000)").

In the **Grid Definition** drop-down list choose the item that matches the grid definition used to develop the incidence/prevalence dataset. Note: The incidence and prevalence rate data must use the same column/row information as the matching grid definition. Click the **Load From File** button. Then click on the **Browse** button, to browse for the dataset file. (The format for the dataset is detailed in the next sub-section.)
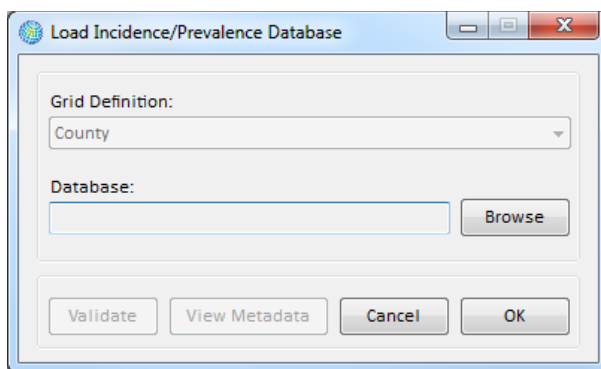
After locating the file, click **Open**. Click the **Validate** button before loading the data. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings. You also have the option of adding metadata to the dataset. This is done by clicking the **View Metadata** button and adding any references or descriptions that you see fit. Click **OK** on the **Load Incidence/Prevalence Database** window to load the selected file. The **Incidence Dataset Definition** window will appear, displaying the rates in the data file that you just loaded.
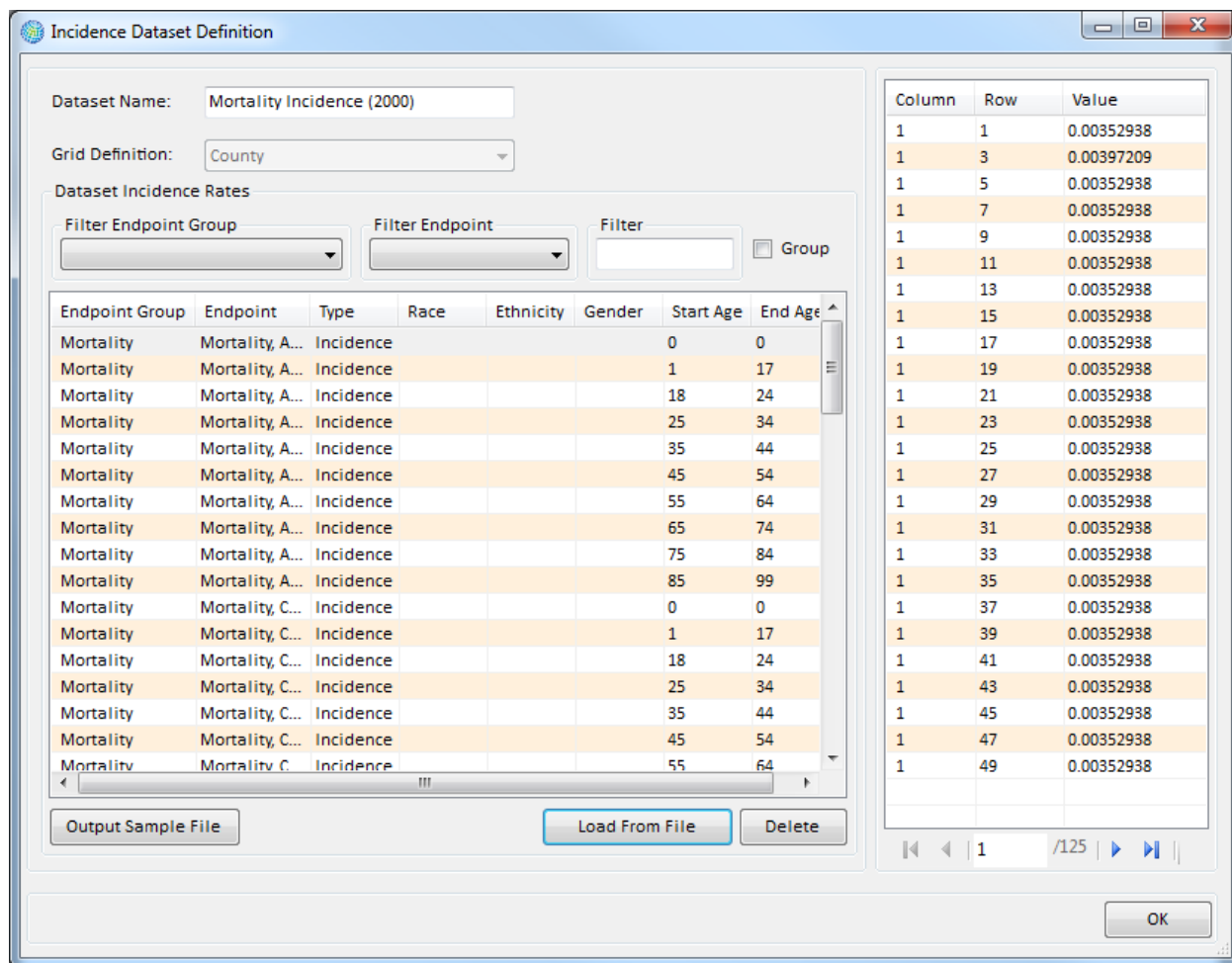
If the data look correct, click **OK**. This will return you to the **Manage Incidence Datasets** window. To view any metadata that was added, select an **Available Dataset** and a **Rate**, and click the **View Metadata** button. The user can view or edit the metadata of the imported files.
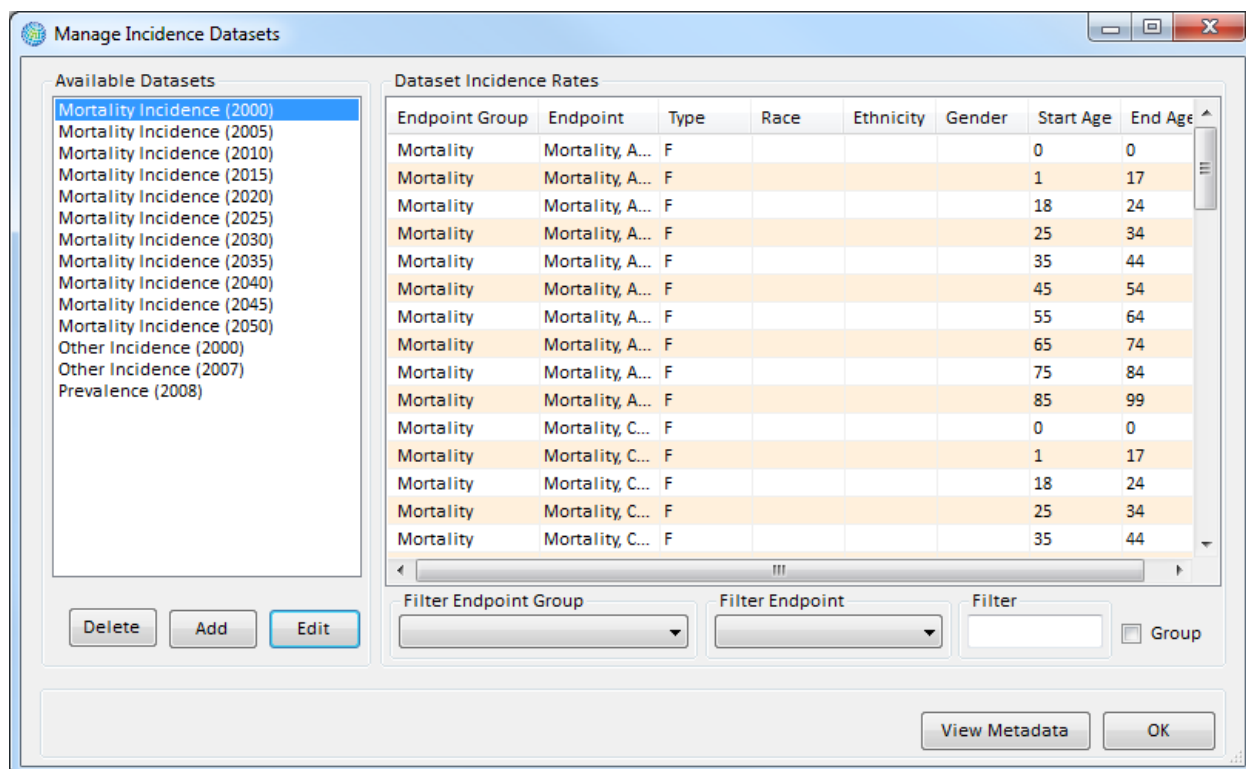
Follow the same procedure for any additional incidence/prevalence datasets that you want to add to the setup database. When you have finished adding data, click **OK** in the **Manage Incidence Datasets** window. The **Incidence/Prevalence Rates** box in the **Modify Datasets** window will show the datasets that you have entered.

4.1.4.2   Format for Incidence/Prevalence Data

Table 4-7a presents the variables that can be used in incidence and prevalence datasets, and Table 4-7b presents a sample dataset that follows this format.

Table 4-7a. Health Incidence and Prevalence Dataset Variables

Field Name	Type	Required	Notes
Endpoint Group	Text	Yes	If this does not reference an already defined Endpoint Group, one will be added.
Endpoint	Text	Yes	If this does not reference an already defined Endpoint for the Endpoint Group, one will be added.
Race	Text	No	Should either be blank (signifying All Races)

Field Name	Type	Required	Notes
			or reference a defined Race, such as "Black" (from one or more Population Configurations).
Ethnicity	Text	No	Should either be blank (signifying All Ethnicities) or reference a defined Ethnicity, such as "Hispanic" (from one or more Population Configurations).
Gender	Text	No	Should either be blank (signifying All Genders) or reference a defined Gender (from one or more Population Configurations).
Start Age	Integer	Yes	Specifies the low and high ages, inclusive. For example, Start Age of "0" and End Age of "1" include infants through the first 12 months of life and all one-year old infants.
End Age	Integer	Yes	
Column	Integer	Yes	The Column and the Row link the incidence/prevalence data with cells from a Grid Definition.
Row	Integer	Yes	
Value	Numeric (double)	Yes	The incidence/prevalence rate for the specified demographic group for this location.
Type	Text	No	If value is a prevalence rate, then "Prevalence" should be specified. Otherwise BenMAP-CE assumes that the value is an incidence rate.

Table 4-7b. Sample Health Incidence Dataset

Endpoint Group	Endpoint	Year	Race	Gender	Ethnicity	Start Age	End Age	Column	Row	Type	Value
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	2	Incidence	2.62E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	3	Incidence	7.73E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	4	Incidence	1.68E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	5	Incidence	5.45E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	6	Incidence	1.02E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	7	Incidence	2.72E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	8	Incidence	5.05E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	9	Incidence	3.03E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	10	Incidence	5.89E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	11	Incidence	3.11E-06
Hospital Admissions, Respiratory	HA, Asthma	2002				0	99	1	12	Incidence	3.30E-06

4.1.5 Population Data

The population data is used to estimate population exposure and in turn any adverse health effects associated with a change in air pollution. BenMAP-CE allows you to

specify race, ethnicity, gender, and age of the population, as well as the year of the population estimate.

Population data loaded into BenMAP-CE must be associated with a **Population Configuration**, which defines the races, ethnicities, genders, and age ranges present in the data. Race, ethnicity, and gender are unique text values representing population subgroups. Age ranges are defined by integer values for starting age and ending age (inclusive), and a unique text value representing the name of the age range. For example, '0TO1' might be used as a name for the age range defined by a start age of zero and an end age of one, thus consisting of infants through the first twelve months of life and all one-year old infants. The population data provided to BenMAP-CE should then contain population values for all combinations of race, ethnicity, gender, and age range. The population values may be non-integer values.

Population data must also be associated with a **Grid Definition** which specifies the geographic areas for which the data is available (see for more details the section on Grid Definitions). If population data is available for multiple grid definitions (cities and neighborhoods, for example), you can have the option of using different sets of population data for different analyses.

BenMAP-CE can also estimate populations for Grid Definitions for which no population data is available by calculating spatial overlap percentages with Grid Definitions for which data is available.

4.1.5.1 Add Population Data

To add population data to BenMAP-CE, click on the **Manage** button below the **Population Datasets** box in the **Modify Datasets** window. The **Manage Population Datasets** window will appear.
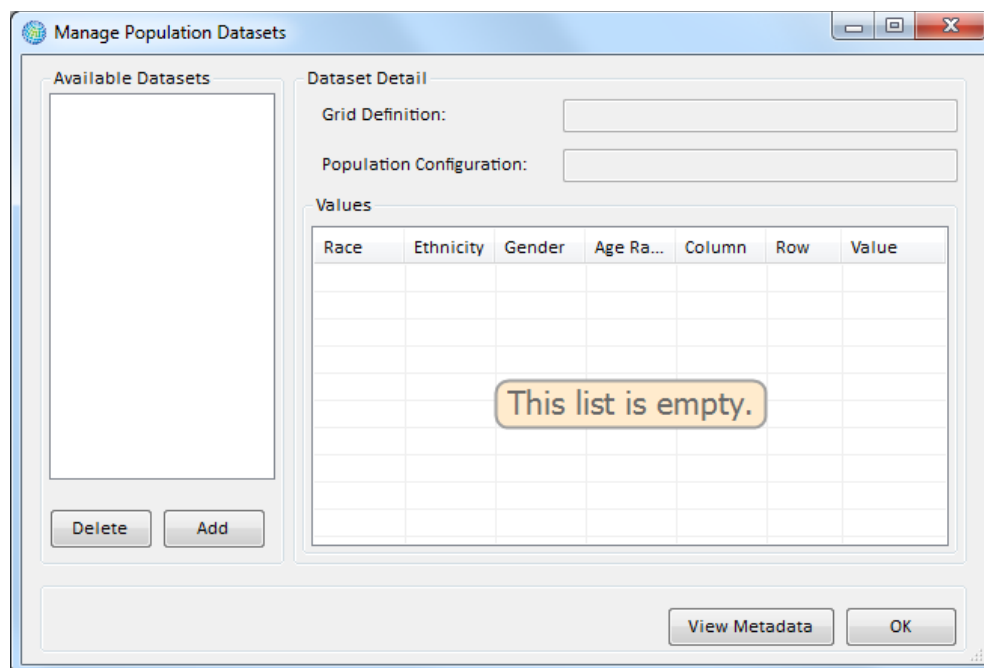
Click on the **Add** button to display the **Load Population Dataset** window. Name the dataset using the **Population Dataset Name** box.

The **Grid Definition** drop-down list provides the list of existing grid definitions. Choose a grid definition that matches your population dataset.

The **Population Configuration** section allows you to define the variables that are in the population data file to be loaded into BenMAP-CE. Use the drop-down list to choose an existing population configuration and then view it by clicking the **View** button, or you may click the **Add** button and define a new population configuration. Clicking the **Add** button will open a **Population Configuration Definition** window where you can enter the fields that appear in the file that will be later uploaded (discussed in more detail below).

The **Browse** button to the right of the Database box allows you to find the data file that you want to load into BenMAP-CE. Once you click **Open** and load the file, the **Validate** and **View Metadata** buttons become active. You can click the **Validate** button before the file is loaded into the **Manage Population Datasets** form. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings. You can also add metadata, which includes references and descriptions, by clicking the **View Metadata** button.

If you wish to run an analysis based on an air quality grid not already defined in BenMAP-CE, you may need to import a new population file matched to that grid definition. The PopGrid program allocates the 2010 block-level U.S. Census population to a user-defined grid, creating a population file ready for importation to BenMAP-CE.[5]

The *Use Population Growth Weights* checkbox should be checked when using population data generated by the PopGrid software application. The population weights file assists in forecasting population levels. See Appendix J for a more detailed discussion of population growth weights in the *United States* setup.

Defining a Population Configuration

If you are performing an analysis outside of the U.S. and are loading your own population data, you will need to first create a new **Population Configuration**. The Population Configuration defines the age range (**Start Age** and **End Age**), **Race**, **Ethnicity**, and **Gender** variables in your population database.

---

[5] The PopGrid program may be downloaded from EPA's website. See: http://www2.epa.gov/benmap/benmap-community-edition.

Check that your population variables align with the population configuration already defined for your setup (the spellings must match exactly). If your data fails to load correctly, you will need to go back and either develop a new population configuration to match your data, or you need to revise your population database so that it matches the population configuration.

> It is <u>critical</u> that the age, race, ethnicity, and gender variables defined in the population configuration match your population input data exactly, otherwise BenMAP-CE will fail to load the population data.

To add a population configuration, click the **Add** button in the **Load Population Dataset** window.

In the **Population Configuration Name** box, replace '*PopulationConfiguration0'* with a name of your choosing. Under the **Races** box click on **New** and type in the name for any races present in your population data. The names appear in both the **Races** list box and the **Available Races** list box. (If you later create alternative population configurations, you can simply drag the relevant names from the **Available Races** list box into the **Races** list box.) Similarly, under the **Available Genders** and **Available Ethnicity** list boxes, click on **New** and type in the name for any ethnicity and gender identifiers present in your population data.

If you want to remove a selected category from the **Races**, **Genders**, or **Ethnicity** list boxes, then highlight the category that you want to deselect and click the **Remove** button. (It is not possible to delete categories from the **Available Races**, **Available Genders**, or **Available Ethnicity** list boxes.)

The next step is to create the age ranges that match the age ranges in your population file. To start click on the **Add** button below the **Age Ranges** list box. The **AgeRange Definition** window will appear. Type in the name of the age variable in the **Age Range ID** box and the upper bound of the age range in the **High Age** box. (BenMAP-CE automatically fills in the value for the **Low Age** box.) For example, the age range names (with corresponding low and high ages) might include the following: *0to0*, *1to4*, *5to9*, *10to14*, *15to19*, *20to24*, *25to29*, *30to34*, *35to39*, *40to44*, *45to49*, *50to54*, *55to59*, *60to64*, *65to69*, *70to74*, *75to79*, *80to84*, and *85up*. The choice of the names is up to you. However, you must be sure that the names exactly match those in your population input file.
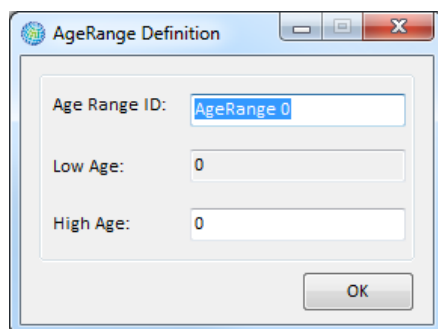
Click **OK** when you have defined the age range. If you make a mistake and want to delete an age definition after you have entered it, click on the **Delete** button. This will remove the last age range that you have entered. (Click on it twice if you want to remove the last two age groups that you entered.) The population configurations can be quite detailed, as in the case of the *United States Census* population configuration that comes loaded with BenMAP-CE.

Click **OK** on the **Population Configuration Definition** window to return to the **Load Population Dataset** window.

Click **OK** on the **Load Population Dataset** window to return to the **Manage Population Datasets** window. To view or edit any metadata that was previously added, click the **View Metadata** button. Click **OK** on the **Manage Population Datasets** window. In the **Population Datasets** box of the **Modify Datasets** window you should see an entry for the population dataset that you just loaded.
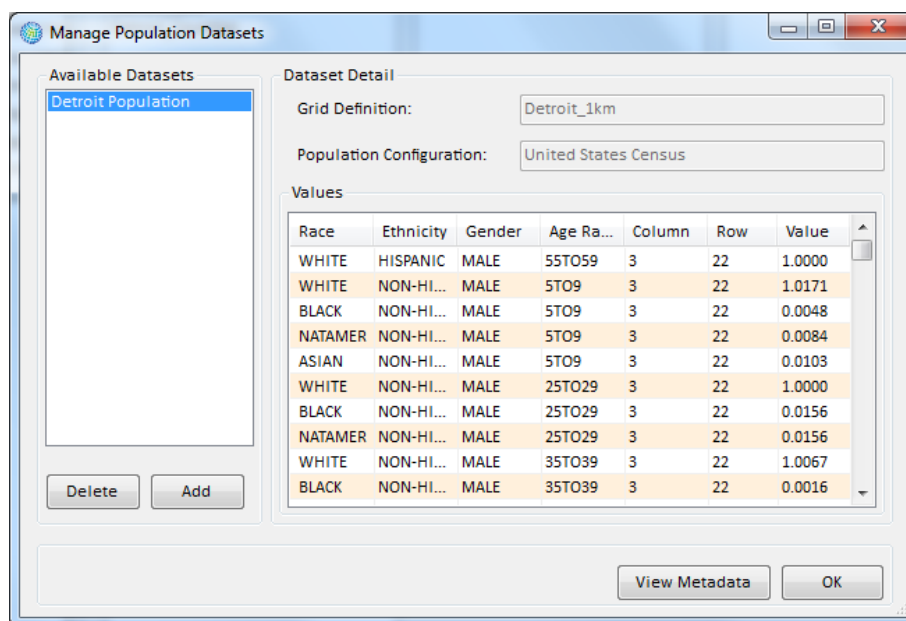
4.1.5.2   Format for Population Data

Table 4-8 presents the variables that can be used in population datasets. Note that the names you define for age ranges do not need to follow the same pattern used in this manual; the age ranges should be based on what seems most appropriate for you. However, it is critical that the age, race, ethnicity, and gender variables in your population input data exactly match those defined for the population configuration, otherwise BenMAP-CE will fail to load the population data

Table 4-8. Population Dataset Variables

Variable	Type	Required	Notes
AgeRange	Text	Yes	References a defined age range in the associated Population Configuration.
Column	Integer	Yes	The column and the row link the population data with cells in a Grid Definition.
Row	Integer	Yes	
Year	Integer	Yes	The year of the data. Note that this may include historical population estimates (such as from a census), as well as population forecasts.
Population	Numeric (double)	Yes	Population estimate. Note that the estimate is not restricted to integers.
Race	Text	Yes	References a defined race in the associated Population Configuration.
Ethnicity	Text	Yes	References a defined ethnicity in the associated Population Configuration. If no ethnicity is specified in the data, "ALL" should be listed throughout the entire column.

Variable	Type	Required	Notes
Gender	Text	Yes	References a defined gender in the associated Population Configuration.

4.1.6 Health Impact Functions

Health impact functions calculate the change in the number of adverse health effects among a certain population associated with a change in exposure to air pollution. A typical health impact function has inputs specifying the pollutant; the metric (daily, seasonal, and/or annual); the age, race, ethnicity, and gender of the population affected; and the incidence rate of the adverse health effect.

Health impact functions are subdivided by user-specified types of adverse health effects. The broadest category is the **Endpoint Group**, which represents a broad class of adverse health effects, such as premature mortality, cardiovascular-related hospital admissions, and respiratory- related hospital admissions, among other categories. (BenMAP-CE only allows pooling of adverse health effects to occur within a given endpoint group, as it generally does not make sense to sum or average together the number of cases of disparate health effects, such as premature mortality and chronic bronchitis.) The **Endpoint Group** may then be subdivided by user-specified **Endpoints**. For example, the respiratory-related hospital admission **Endpoint Group**, may have separate **Endpoints** for asthma-related hospital admissions and chronic bronchitis-related hospital admissions.

There are a wide range of variables that can be included in a health impact function, to specify the parameters of the function and to identify its source, such as the **Author**, **Year**, and **Location** of the study, as well as other **Pollutants** used in the study. The bibliographic **Reference** for the study may be included, as well as any additional information needed to identify a particular impact function. (The **Reference** variables are useful for this.) A number of health impact functions have been developed based on epidemiological studies in the United States and Europe. However, researchers have conducted an increasing number of epidemiological studies in Asia and Latin America that can be used to develop more location-specific impact functions. There are a number of issues that arise when deriving and choosing between health impact functions that go well beyond this user manual. *Hence, it is important to have a trained health researcher assist in developing the impact function data file.* [6]

4.1.6.1 Add Health Impact Functions

To add health impact functions to BenMAP-CE, click on the **Manage** button below the **Health Impact Functions** box in the **Modify Datasets** window. The **Manage Health Impact Functions Datasets** window will appear.

---

[6] U.S. EPA-default configuration and pooling setup files for ozone and $PM_{2.5}$ health impact assessments are available on the BenMAP-CE website. See: http://www.epa.gov/benmap/benmap-community-edition.
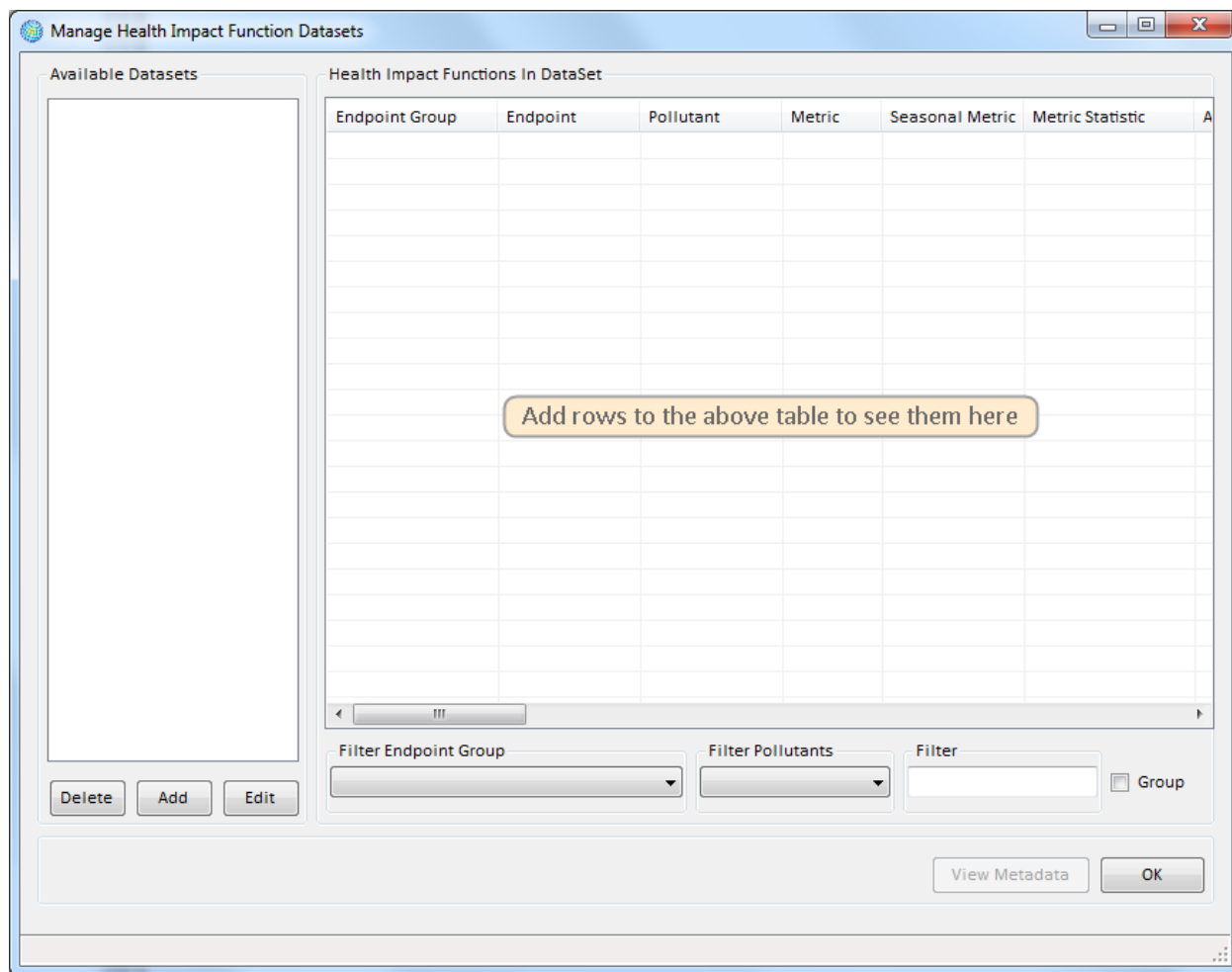
In this window you may **Add**, **Edit**, and **Delete** datasets. The section on the left under **Available Datasets** lists the health impact function datasets that are currently in the setup database. (See Appendices E, F, and G for more information about the pre-loaded health impact functions.) The section on the right under the **Health Impact Functions in Dataset** lets you view the functions in a selected dataset.

To add a new dataset, click the **Add** button. The **Health Impact Function Dataset Definition** window will appear. Type the name that you want to use for the dataset in the **Health Impact Function Dataset Name** box.

You may then enter functions into this dataset through an externally created database by clicking the **Load From File** button. Alternatively, you may **Add**, **Delete**, and **Edit** individual functions within BenMAP-CE.

To add a database, click the **Load From File** button. In the **Load Health Impact Dataset** window, click the **Browse** button and then find and select the health impact function database that you want to load into your setup. Click **Open**. If validation is required, then you will have to click the **Validation** button before the file can be imported. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings.

You can also add metadata (references and descriptions) to the file that is about to be imported by clicking the **View Metadata** button. Click the **OK** button on the **Load Health Impact Dataset** window to load the dataset.

The **Health Impact Function Dataset Definition** window will reappear, and you can then view the health impact functions that you have loaded into your dataset.
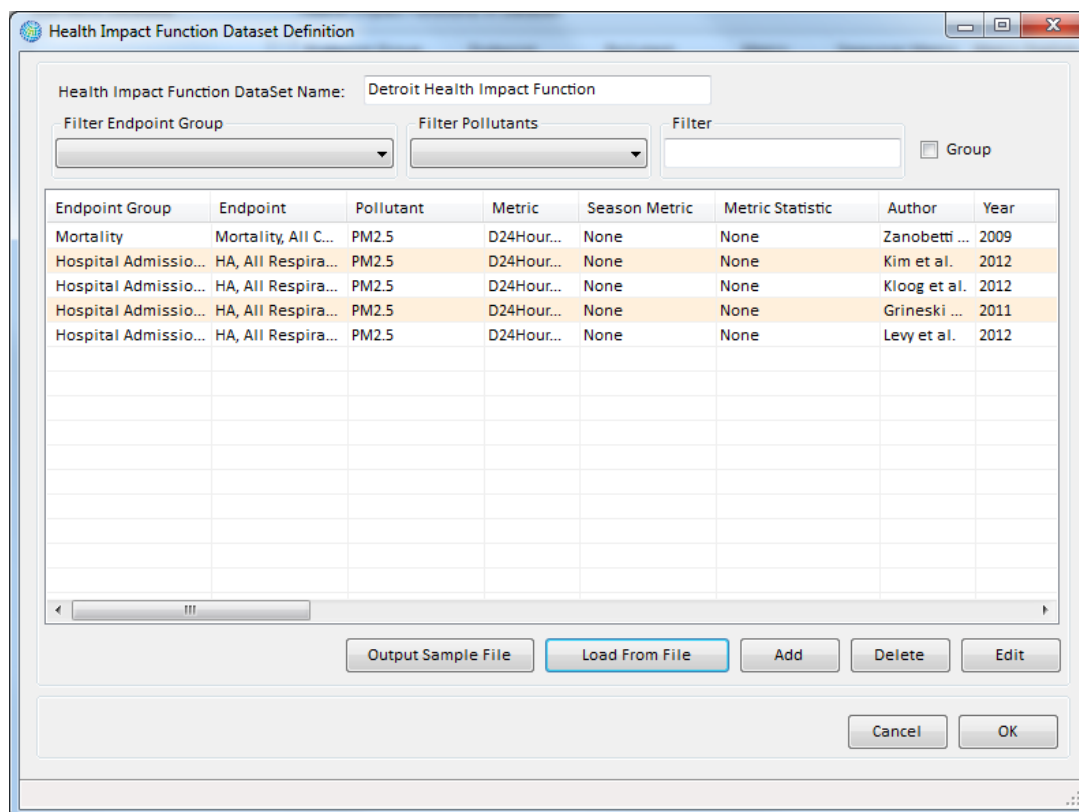
By clicking and holding the cursor on a column header, you may move it to provide the most useful display. For example, by clicking and holding on the **Pollutant** column header and then dragging it to the far left of the window, you can sort all of the health impact functions by **Pollutant**. (Note rearranging the columns is only for display and has no effect on the underlying health impact functions in the database.)

Clicking **OK** brings you back to the **Manage Health Impact Function Datasets** window. The new dataset you just loaded will be displayed in the list of **Available Datasets** and the associated functions will be displayed in the **Health Impact Functions In Dataset** grid to the right. If you have more than one dataset, you can select the dataset by clicking on it.
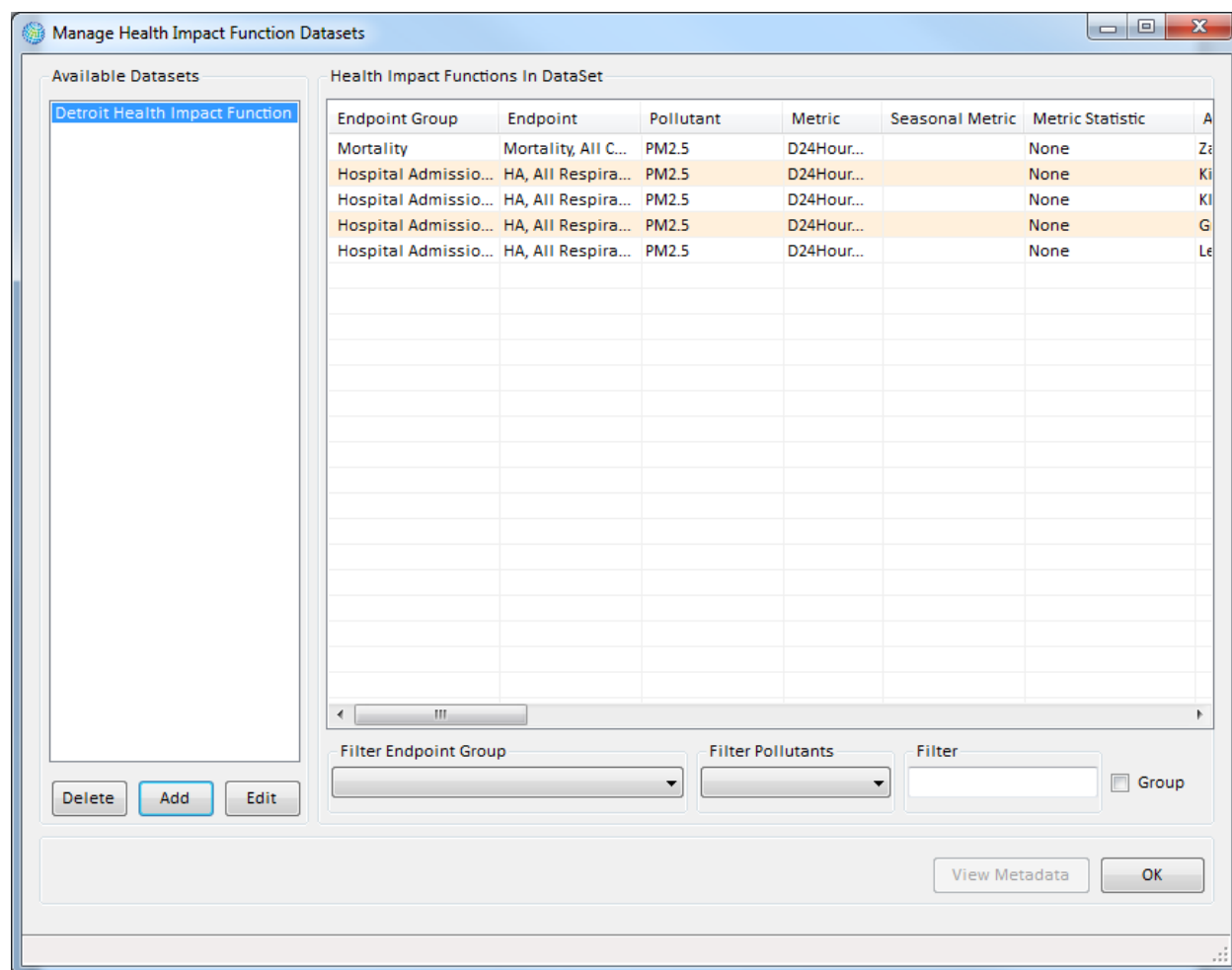
To edit an existing function, first click to select the dataset in the list of **Available Datasets**. Next, select a particular function in the data grid under **Health Impact Functions in Dataset**. Then, click the **Edit** button. The **Health Impact Function Dataset Definition** window appears, and you may change any of the values in the boxes and the drop-down lists. When you are finished, click **OK**.

From the **Health Impact Function Dataset Definition** window you can also add health impact functions to the ones that are already in your dataset. Click the **Add** button and a blank **Health Impact Function Definition** window will appear and you can then create new health impact functions. (See Appendix M: Function Editor for additional information about the syntax for developing functions with this editor.)

Starting in Version 1.4, you can link a health impact function to a specific geographic area for which you have uploaded a grid definition by selecting that area in the **Apply Function To** dropdown menu. This prevents you from applying the function to other grid definitions but may be appropriate for functions derived using sub-national data. By default this menu is set to "Entire Area", which means the application of the health impact function is unrestricted.

After defining the new health impact function, click **OK**. This will take you back to the **Health Impact Function Dataset Definition** window. When you are finished with any editing or adding of health impact functions, click **OK**. From the **Manage Health Impact Function Datasets** window, you can also select an **Available Dataset** and **Data**

row and view the Metadata. To view the Metadata associated with the data file, click the **View Metadata** button to view and edit existing references and descriptions. Click **OK** on the **Manage Health Impact Function Datasets** window when you are satisfied with all your inputs. The **Modify Datasets** window will appear. Here in the **Health Impact Functions** box you should see an entry for any health impact function datasets that you have loaded.

4.1.6.2   Format for Health Impact Functions

Table 4-9 presents the variables that can be used in health impact function datasets.

Table 4-9. Health Impact Function Dataset Variables

Variable	Type	Required	Notes
Endpoint Group	Text	Yes	If this does not reference an already defined Endpoint Group, one will be added.
Endpoint	Text	Yes	If this does not reference an already defined Endpoint for the Endpoint Group, one will be added.
Pollutant	Text	Yes	Should reference an already defined Pollutant.
Metric	Text	Yes	Should reference an already defined Metric for the Pollutant.
Annual Statistic	Text	No	Should either be blank (signifying no annual metric value) or be one of: None, Mean, Median, Min, Max, Sum.
Seasonal Metric	Text	No	Should either be blank (signifying no Seasonal Metric value) or reference an already defined Seasonal Metric for the Metric.
Race		No	Should either be blank (signifying All Races) or reference a defined Race.
Ethnicity		No	Should either be blank (signifying All Ethnicities) or reference a defined Ethnicity.
Gender		No	Should either be blank (signifying All Genders) or reference a defined Gender.
Start Age	Integer	Yes	Specifies the low and high ages, inclusive. For example, Start Age of '*0*' and End Age of '*1*' includes infants through the first 12 months of life and all one-year old infants.
End Age	Integer	Yes	
Author	Text	No	The author(s) of the study from which the function is derived.
Apply Function To		No	The specific geographic area to which you would like to apply the health impact function; unrestricted by default ("Entire Area")
Year of Publication	Integer	Yes	The year of publication of the study.

Variable	Type	Required	Notes
Qualifier	Text	No	Provides additional information to identify a particular health impact function, such as when a particular study has multiple functions.
Location Name		No	Type of study area. For the 'United States' setup, choose between 'State', 'County', and 'MSA (metropolitan area)'.
Location	Text	No	The specific location of the study.
Co-Pollutants Specified in Regression Model	Text	No	Identifies other pollutants that were included simultaneously in the estimation equation for the pollutant of interest.
Reference	Text	No	Bibliographic reference, included to identify the source in the health literature.
Function	Text	Yes	The functional form, interpreted (executed) by BenMAP-CE when running an analysis to estimate air pollution-related health impacts. For example, the log-linear form is as follows: '*(1-(1/ EXP(Beta*DELTAQ)))*Incidence*POP*'.
Baseline Incidence Function	Text	Yes	The functional form, interpreted (executed) by BenMAP-CE to estimate health impacts due to all causes. This typically has the form: '*Incidence*POP*'.
Beta Distribution	Text	No	If the Beta has no distribution, any value is acceptable. Otherwise, should be one of: *Normal, Triangular, Poisson, Binomial, LogNormal, Uniform, Exponential, Geometric, Weibull, Gamma, Logistic, Beta, Pareto, Cauchy, Custom*.
Beta	Numeric (double)	No	Mean value of the Beta distribution.
Beta Parameter 1	Numeric (double)	No	Parameter 1 of the Beta distribution (meaning depends on the distribution - for Normal distributions this represents the standard deviation).
Beta Parameter 2	Numeric (double	No	Parameter 2 of the Beta distribution (meaning depends on the distribution - for Normal distributions this is not required).
Name A	Text	No	Description of variable A.
A	Numeric (double)	No	A constant value which can be referenced by the Function.
Name B	Text	No	Description of variable B.
B	Numeric (double)	No	A constant value which can be referenced by the Function.
Name C	Text	No	Description of variable C.

Variable	Type	Required	Notes
C	Numeric (double)	No	A constant value which can be referenced by the Function.
Incidence Dataset	Text	No	Specifies the dataset from which incidence data will be derived. The user may choose from multiple datasets. (Initially this field may be left blank.) See Section 4.1.4.
Prevalence Dataset	Text	No	Specifies the dataset from which prevalence data will be derived. The user may choose from multiple datasets. (Initially this field may be left blank.) See Section 4.1.4.
Variable Dataset	Text	No	Specifies the dataset from which "variable" data (e.g., income data) will be derived. The user may choose from multiple datasets. (Initially this field may be left blank.) See Section 4.1.7.

While the mean value of the Beta distribution is always available in BenMAP-CE, the values associated with Beta Parameter 1 and Beta Parameter 2 may change depending on the type of distribution. Table 4-10 provides a list of the variables associated with each type of beta distribution in BenMAP-CE.

Table 4-10. Beta Distribution Types and Variables

Distribution	Formula	Beta Parameter 1	Beta Parameter 2	Notes
Normal	$\dfrac{1}{\sigma\sqrt{2\pi}} \times e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	Standard deviation (sigma)	N/A	The Normal distribution has two parameters - the mean, mu, and the standard deviation, sigma.
Triangular	$\dfrac{2(x-a)}{(b-a)(c-a)}$ for $a \leq x \leq c$ $\dfrac{2(b-x)}{(b-a)(b-c)}$ for $c \leq x \leq b$	Minimum value (a)	Maximum value (b)	The Triangular distribution has three parameters - the minimum value (a), the maximum value (b), and the most likely value (c). BenMAP-CE uses the mean value, the minimum, and the maximum to calculation the most likely value.
Poisson	$\lambda^x e^{-\lambda}/x!$	Lambda	N/A	The Poisson distribution has a single parameter, lambda.
Binomial	$\dfrac{n!}{x!\,(n-x)!}\, p^x(1-p)^{n-x}$	n	p	The Binomial distribution has two parameters, n and p.
LogNormal	$1/\sigma\sqrt{2\pi} \cdot e^{-(x-\mu)^2/2\sigma^2}$	Standard deviation (sigma) of the corresponding Normal	N/A	The LogNormal distribution has two parameters - the mean of the corresponding Normal distribution, mu, and the standard deviation of the

Distribution	Formula	Beta Parameter 1	Beta Parameter 2	Notes
		distribution		corresponding Normal distribution, sigma.
Uniform	$1/(B-A)\ for\ A < x < B$	A	B	The Uniform distribution has two parameters, A and B, which define the interval on which the distribution is defined.
Exponential	$\dfrac{1}{\mu}e^{-x/\mu}$	Mu	N/A	The Exponential distribution has one parameter, mu.
Geometric	$p(1-p)^x\ for\ 0 < p < 1$	p	N/A	The Geometric distribution has one parameter, p.
Weibull	$\left(\dfrac{\beta}{\alpha}\right)\left(\dfrac{x}{\alpha}\right)^{\beta-1}e^{-(x/\alpha)^{\beta}}$	alpha	beta	The Weibull distribution has two parameters, alpha and beta.
Gamma	$\dfrac{1}{b^{a}\Gamma(a)}x^{a-1}e^{-x/b}$	a	b	The Gamma distribution has two parameters, a and b.
Logistic	$\dfrac{e^{(x-m)/b}}{\left|b\right|\left[1+e^{(x-m)/b}\right]^2}$	m	b	The Logistic distribution has two parameters, m and b.
Beta	$\dfrac{1}{\beta(a,b)}x^{a-1}(1-x)^{b-1}$	a	b	The Beta distribution has two parameters, a and b.
Pareto	ab^{a}/x^{a+1}	a	b	The Pareto distribution has two parameters, a and b.
Cauchy	$\dfrac{1}{\pi}\cdot\dfrac{b}{(x-m)^2+b^2}$	b	m	The Cauchy distribution has two parameters, b and m.
Custom	N/A	Standard deviation	N/A	The custom distribution is used for specified expert distributions for pollutant effect (e.g., truncated parametric distributions or non-parametric distributions).

For the **Function** definition, commonly used mathematical operators (+, -, *, /) may be used. Other available "**Operators**" are listed to the right of the commonly used functional forms (e.g., ABS(x), EXP(x), LOG(x)). These operators are supported by the math and statistics library used within BenMAP-CE.

Under the heading "**Available Variables**", are temporary runtime variables which can be used in the health impact function definition. The values for these variables (except for the A, B, C constants) are expected to change as the code loops over grid cells and stratified population groups. Table 4-11 presents the available runtime variables that can be used in health impact function definition.

Table 4-11. Health Impact Function Available Runtime Variables

Variable	Type	Notes
Beta	Numeric	Beta coefficient
DELTAQ	Numeric	Change in pollutant concentration (baseline – control)
POP	Numeric	Population
Incidence	Numeric	Number of people with new (or newly diagnosed) adverse health effects within a given period of time.
Prevalence	Numeric	Number of people who already have a given adverse health condition (i.e., chronic illness).
Q0	Numeric	Control pollutant concentration
Q1	Numeric	Baseline pollutant concentration
A	Numeric	User-defined constant
B	Numeric	User-defined constant
C	Numeric	User-defined constant

The "**Population Variables (optional)**" box provides a list of setup variables which have been defined under the **Variable Datasets** (see **Section 4.1.7**).

4.1.7   Variable Data

Health Impact Functions and **Valuation Functions** may sometimes refer to socioeconomic variables for which BenMAP-CE does not automatically calculate values. For example, some valuation functions reference the median income within each area of analysis. Other functions apply to populations living below the poverty line in a given country. To facilitate this type of analysis, BenMAP-CE allows you to load datasets of socioeconomic **Variables**, which apply either globally or to a specific geographic area (i.e., they are associated with a **Grid Definition**).

4.1.7.1   Add Variable Data

To add **Dataset Variables** to BenMAP-CE (such as income and other miscellaneous variables that might be needed in the analysis), click on the **Manage** button below the **Variables Datasets** box in the **Modify Datasets** window. The **Manage Setup Variable Datasets** window will appear.
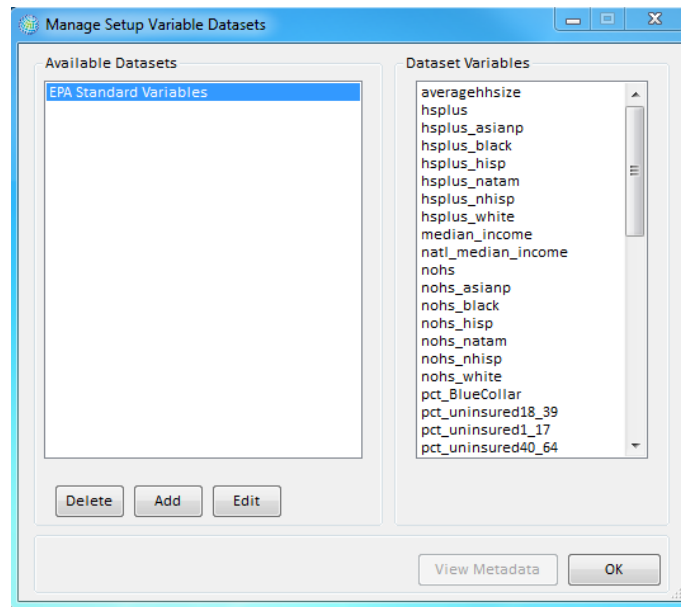
In this window you may **Add**, **Edit**, and **Delete** datasets. The section on the left under **Available Datasets** lists the variables datasets that are currently in the setup database. The section on the right under the **Dataset Variables** lets you view the variables in a selected dataset.

To add a **Variable** dataset click the **Add** button. This will take you to the **Setup Variable Dataset Definition** window. In this window you may add externally created variables through the **Load From File** button for any of your predefined **Grid Definitions**.
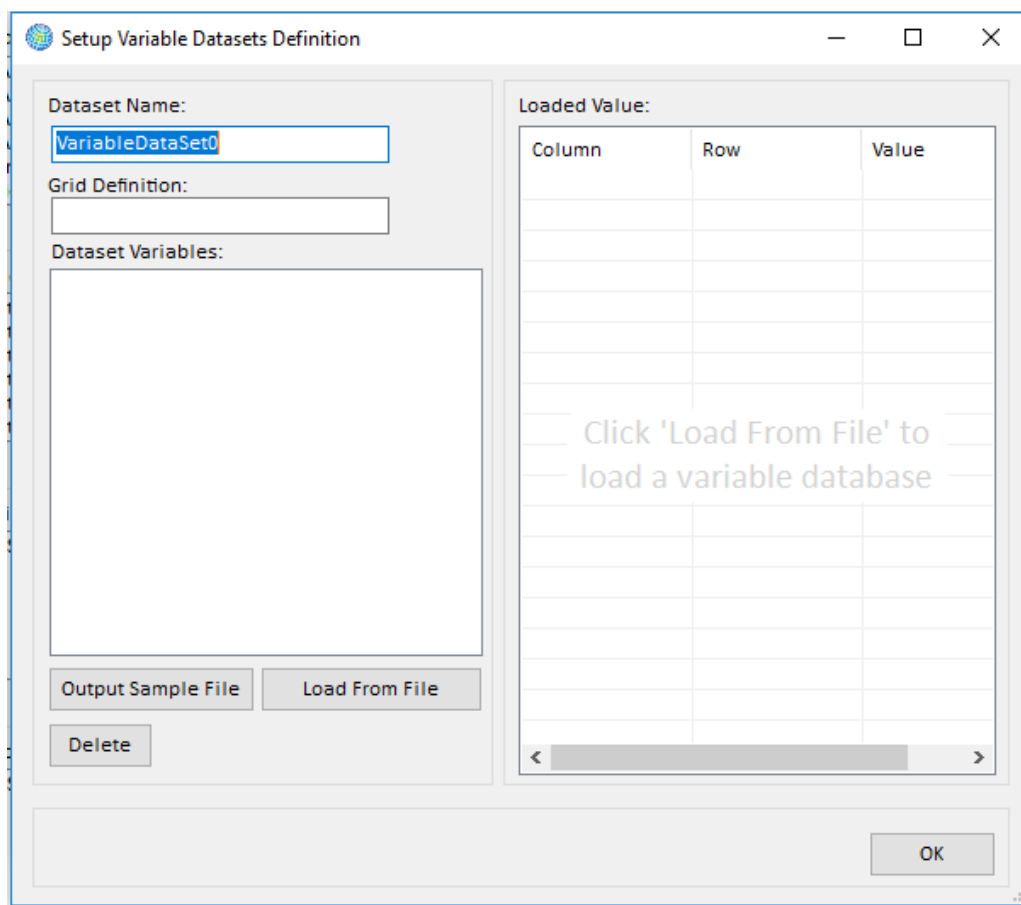
To start, type the name that you want to use for the **Variable Dataset** in the **Dataset Name** box. (This is a name that is internal to BenMAP-CE and used just for identification.)

To add an externally created **Variable Dataset**, click the **Load From File** button. This will bring up the **Load Variable Database** window. Here you need to choose the grid definition from the **Grid Definition** drop-down list that matches the level of aggregation in your variable data file. Remember that the **Variable Dataset** you import must use the same column/row index as the **Grid Definition**. Next, you may use the **Browse** button to find and select the desired **Database** (i.e., input file) and click **Open**. You can click the **Validate** button to ensure the file is properly formatted before importing. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings. You can also add metadata to the imported data (reference and description of the data file) by clicking the **View Metadata** button.
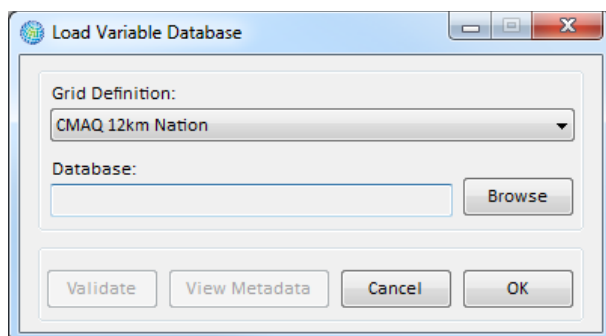
After choosing (and validating) the input file, click **OK**. This takes you back to the **Setup Variable Datasets Definition** window. This window displays the variables in the dataset and their values.
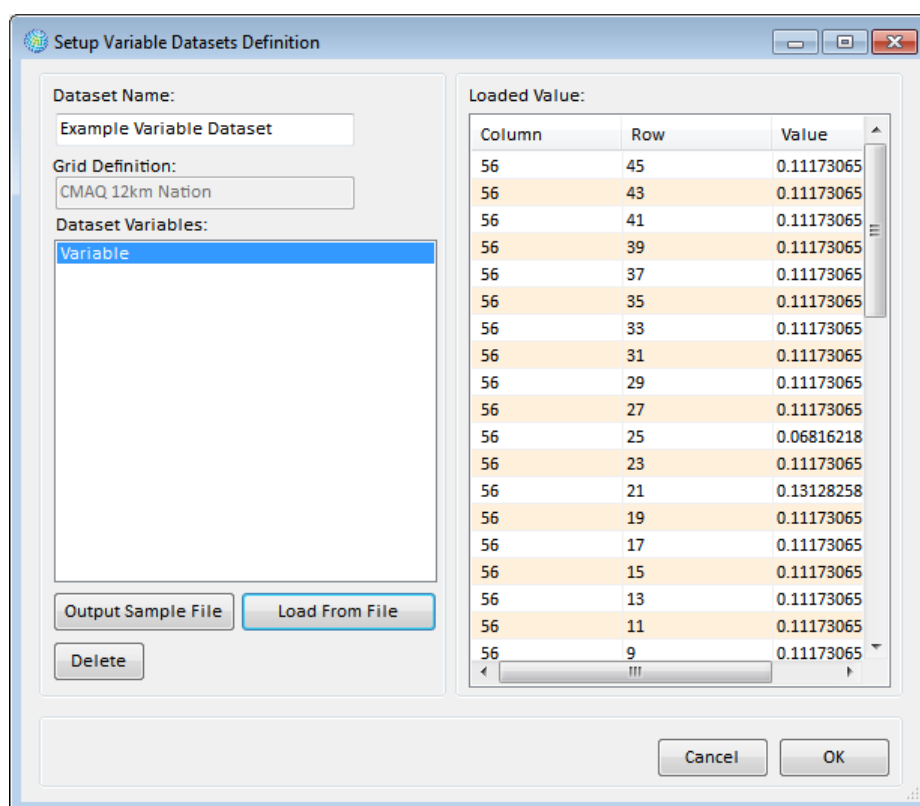
When finished adding variables, click **OK**. This will take you to the **Manage Setup Variable Datasets** window.

In the **Available Datasets** list box there is an entry for the dataset that you just added. And in the **Dataset Variables** list box are the variables in the highlighted dataset.
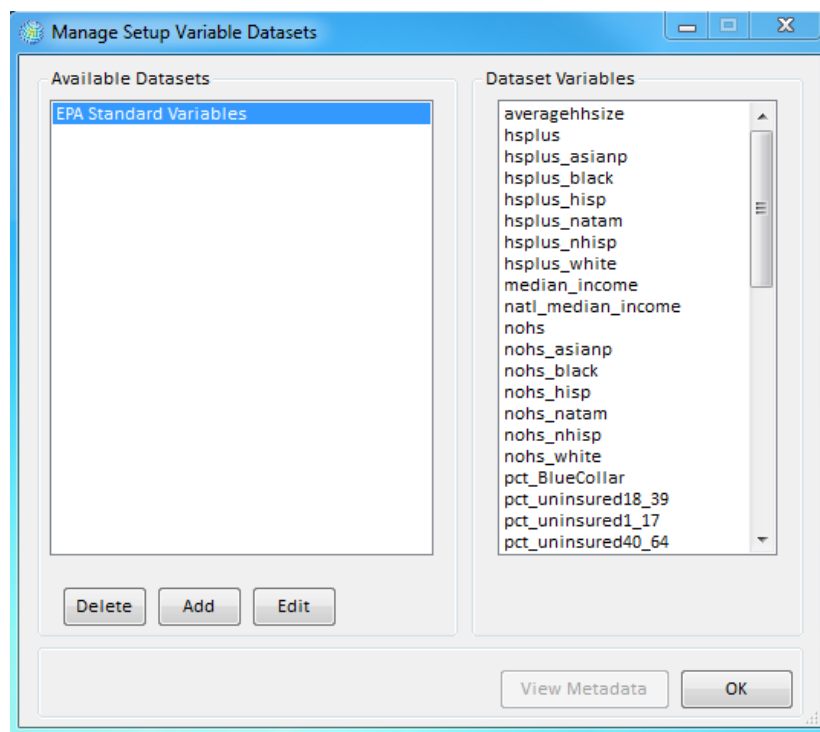
At this point you may click **Add** to load an additional dataset, click **Edit** to edit the selected dataset, click **Delete** to delete the selected dataset, or complete this variable management step by clicking **OK**.

Clicking **OK** returns you to the **Modify Datasets** window, where the entry for the variable dataset that you just entered should be visible under the **Variable Datasets** box.

4.1.7.2   Format for Variable Data

Table 4-12a presents the variables that can be used in variable datasets, and Table 4-12b presents a sample of what a dataset might look like. Note that if you are loading your own variable data, you can choose your own variable names (July 2018: this feature is currently disabled. Users should currently specify one variable per file).

Table 4-12a. Variable Dataset Variables

Variable	Type	Required	Notes
Column	Integer	Yes	The column and the row link the population data with cells in a Grid Definition.
Row	Integer	Yes	
Median_Income	Numeric (double)	Yes	Example: Median income value.
<Variable Name>	Numeric (double)	No	Additional variables may be specified.

Table 4-12b. Sample Variable Dataset

Column	Row	Median_Income
42	17	39871.19
42	29	43760.31
42	45	38186.77
42	91	42298.49
42	101	31261.07

4.1.8   Inflation Data

It may be desirable for the economic values generated by **Valuation Functions** to account for inflation and generate economic benefits using currency for years other than the year initially specified by your valuation data. To do this, you can load **Inflation Datasets** into BenMAP-CE, and then include a reference to your inflation data when developing valuation functions. (We give an example of this below.)

The **Valuation Functions** should have a consistent currency year, and this currency year has to be kept in mind when developing the inflation datasets. That is, whichever currency year is used for your valuation functions, then the inflation values for that year should be set to 1. For example, in the *United States* setup, the valuation functions have a dollar year of 2015, so the inflation datasets have a value of 1 for the year 2015. (Values for years earlier than 2015 are less than 1, and values for years after 2015 are greater than 1, because inflation has increased from one year to the next.) The *United States* setup in BenMAP-CE provides inflation factors for three different types of values:

- **All Goods Index** can be used to adjust the value of generic goods.

- **Medical Cost Index** can be used to adjust the value of medical expenses.

- **Wage Index** can be used to adjust the value of wages.

If a **Valuation Function** includes an estimate of wage income, for example, this value could be multiplied by the **Wage Index** adjustment factor to get the specified currency year. For example, in valuing work loss days, the *United States* setup uses a function like the following: *DailyWage*WageIndex*, where the *DailyWage* is specified in year 2015 dollars. In the **Inflation Dataset**, the **WageIndex** scales this **DailyWage** value up or down depending on the currency **Year** you have chosen. If the currency **Year** is 2015, then the **WageIndex** has a value of 1 and no change is made to the **DailyWage**. If the currency **Year** is specified to, say, 2020, then the **WageIndex** will have a value greater than 1 because of the inflation that has occurred between 2015 and 2020.

4.1.8.1.   Add Inflation Data

To add inflation data to BenMAP-CE, click on the **Manage** button below the **Inflation Datasets** box in the **Modify Datasets** window. The **Manage Inflation Datasets**

window will appear. In this window you may **Add**, **Edit**, and **Delete** datasets. The section on the left under **Available Datasets** lists the **Inflation Datasets** that are currently in the setup database. The section on the right under **Inflation Detail** presents the inflation factors in a selected dataset.

Click on the **Add** button. In **the Load Inflation Dataset** window, type in the name of the dataset in the **Inflation Dataset Name** box, and then click on the **Browse** button to the right of the **Database** box to choose the dataset that you want to import and click **Open**. You can click the **Validate** button to ensure the file is properly formatted before importing. BenMAP-CE's validation tool will review the file format (column names, required columns, and data types) and provide a report with any errors or warnings. In addition, you can add metadata to the file, to include references and descriptions, by clicking the **View Metadata** button. Click **OK**. The **Manage Inflation Datasets** window will appear. Here you may view the data that you just loaded.

Manage Inflation Datasets			
Available Datasets	**Inflation Detail**		
EPA Standard Inflators	Year	All Goods Index	Medical Cost Index	Wage Index
	1997	0.677166640758514	0.525123536586761	0.610798954963684
	1998	0.687714397907257	0.541911423206329	0.631290316581726
	1999	0.702903151512146	0.560937583446503	0.652543187141418
	2000	0.726530194282532	0.583769083023071	0.679082036018372
	2001	0.747203767299652	0.610629618167877	0.706370055675507
	2002	0.759017288684845	0.63928085565567	0.729130864143372
	2003	0.776315629482269	0.665022194385529	0.757174372673035
	2004	0.796989262104034	0.694121122360229	0.784466981887817
	2005	0.823991537094116	0.723443865776062	0.80721652507782
	2006	0.850571870803833	0.752542793750763	0.832246482372284
	2007	0.874797999858856	0.78579169511795	0.85651707649231
	2008	0.908386290073395	0.814915180206299	0.876966536045074
	2009	0.905154466629028	0.840763986110687	0.88731187582016
	2010	0.92000150680542	0.869466722011566	0.905452430248261
	2011	0.949041604995728	0.895928859710693	0.925136208534241
	2012	0.968681573867798	0.928756892681122	0.941727042198181
	2013	0.9828703999951935	0.951610743999481	0.959867537021637
	2014	0.998814404010773	0.974348187446594	0.981861174106598
	2015	1	1	1

Output Sample File
Delete   Add

View Metadata   OK

At this point you may add more data by clicking **Add**, or you may view and edit the Metadata for a specific dataset. This is done by selecting an **Available Dataset** and an entry in the **Inflation Detail** box and then clicking the **View Metadata** button. If you are satisfied with all import data, you can complete this step by clicking **OK**. Clicking **OK**

takes you to the **Modify Datasets** window, where you should see an entry for the inflation dataset that you just loaded under the **Inflation Datasets** box.

4.1.8.2   Format for Inflation Data

Table 4-13a presents the variables that can be used in **Inflation Datasets**, and Table 4-13b presents a sample of what a dataset might look like. Note that if you are loading your own inflation data, you can use different names than the ones specified below. Instead of specifying '*AllGoodsIndex*' you could have a variable called '*General Index*' — this is fine as long as you make sure that your valuation functions properly reference these inflation variables.

Table 4-13a. Inflation Dataset Variables in U.S. Setup

Variable	Type	Required	Notes
Year	Integer	Yes	The year of the data. Note that this will typically only include historical estimates.
AllGoodsIndex	Integer	No	Example: All goods inflation index value.
MedicalCostIndex	Integer	No	Example: Medical cost inflation index value.
WageIndex	Integer	No	Example: Wage inflation index value.
<Variable Name>	Integer	No	Additional indices can be specified.

Table 4-13b. Sample Inflation Dataset

YEAR	AllGoodsIndex	MedicalCostIndex	WageIndex
2005	0.75	0.62	0.71
2006	0.79	0.67	0.74
2007	0.81	0.72	0.76
2008	0.83	0.77	0.78
2009	0.86	0.8	0.81
2010	0.88	0.84	0.83
2011	0.91	0.87	0.86
2012	0.93	0.89	0.89
2013	0.94	0.92	0.92
2014	0.96	0.96	0.96
2015	1	1	1
2016	1.02	1.04	1.03

4.1.9 Valuation Data

BenMAP-CE allows the valuation estimates to vary by **Endpoint Group**, **Endpoint**, and **Age** (note that multiple estimates may be provided for a particular combination). BenMAP-CE allows the valuation function to be quite detailed, and allows an uncertain parameter (A) with a user-specified distribution. You can modify the valuation function with a number of constant values (B, C, and D) that might represent an adjustment for inflation, income growth, income elasticity, or, say, purchasing power parity. Finally,

BenMAP-CE has two fields to more clearly identify the valuation function (i.e., **Qualifier** and **Reference**).

When reviewing the economic literature to develop a valuation database or to simply add valuation functions to an existing database, it is important to have an economist assist. For an overview of valuation, see the Overview of Valuation section in Chapter 7: Aggregating, Pooling, and Valuing.

4.1.9.1   Add Valuation Data
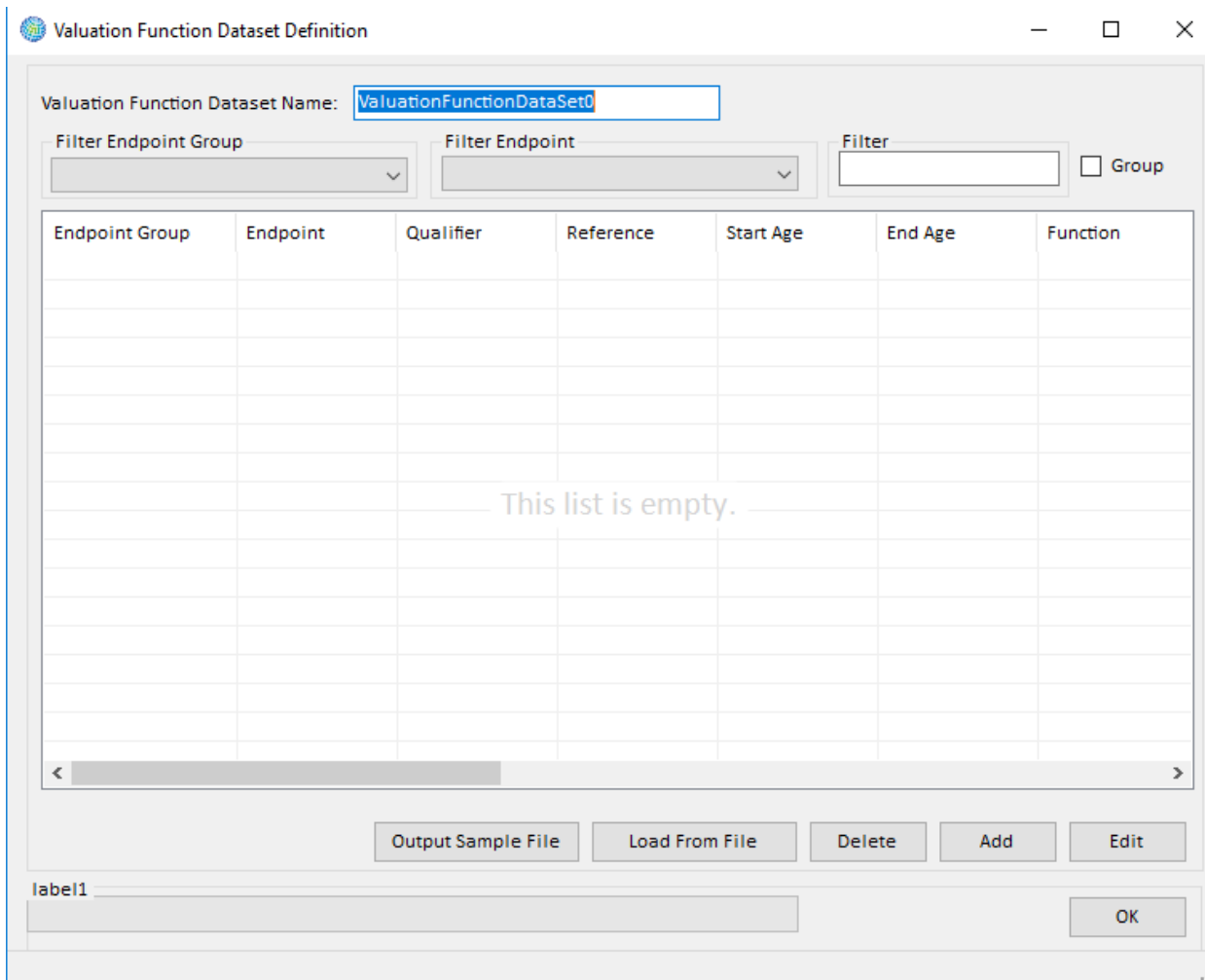
To add valuation functions to BenMAP-CE, click on the **Manage** button below the **Valuation Datasets** box in the **Modify Datasets** window. The **Manage Valuation Function Datasets** window will appear.
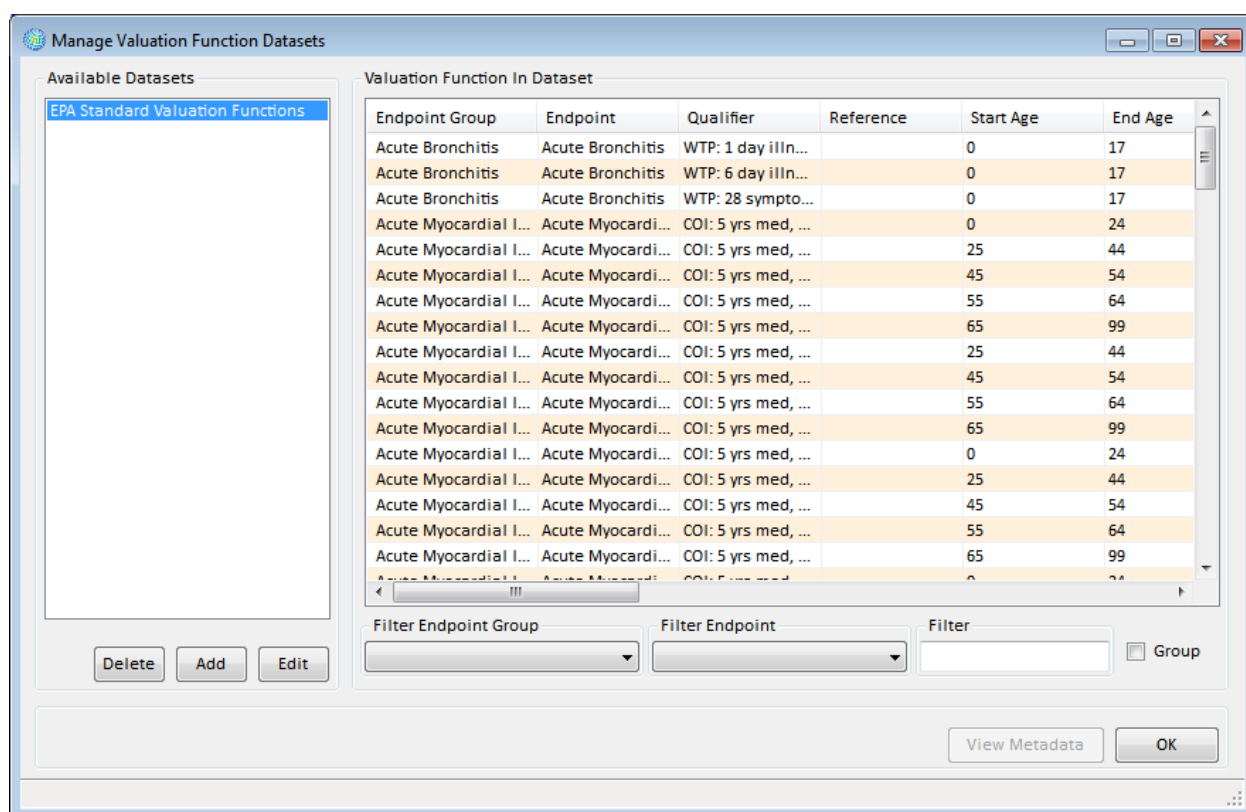
In this window you may **Add**, **Edit**, and **Delete** datasets. The section on the left under **Available Datasets** lists the valuation datasets that are currently in the setup database. The section on the right under the **Valuation Function In Dataset** lets you view the valuation factors in a selected dataset.

If the dataset is large, there are filters available to view a subset of the list by selecting a value from the **Filter Endpoint Group** and/or **Filter Endpoint** drop-down lists. Or, you can type a value in the **Filter** box to search for a particular word or phrase. You can also group data by **Endpoint Group** by selecting the *Group* option.
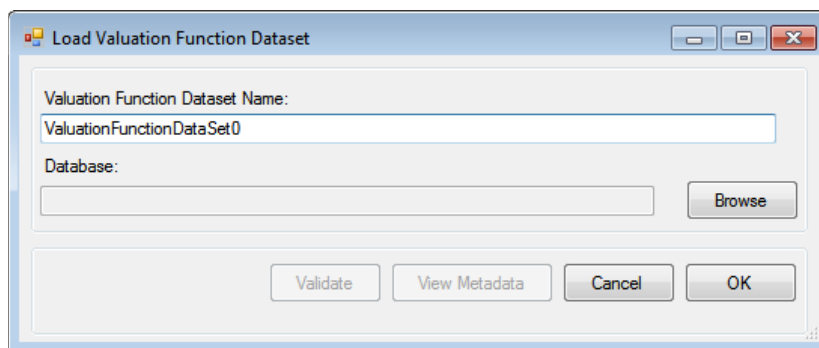
To add a dataset, click on **Add**. This will display the **Valuation Function Dataset Definition** window.

You may load valuation data with an externally-created data file, or you may add individual valuation functions from within BenMAP-CE. To import valuation functions, click on the **Load From File** button. This will bring you to the **Load Valuation Function Dataset** window.

In the **Load Valuation Function Dataset** window provide a name for the valuation function dataset. Use the **Browse** button to choose the valuation database and click **Open**. Once again, you can click the **Validate** button before the file is imported. In addition, you can create metadata, which includes references and descriptions of the file, by clicking on the **View Metadata** button. Click **OK** on the **Load Valuation Function Dataset** window. This will bring you back to the **Valuation Function Dataset Definition** window. Here you can view the valuation functions that you have in your database.

The columns within each of the list boxes can be rearranged in order to provide the most useful display. (Note that rearranging the columns is only for display and has no effect on the underlying Valuation Function dataset.) You can also change how you view the list by specifying a text string to **Filter** the records (e.g., Asthma), or create groupings of records by their Endpoint Group (click the **Group** option box).
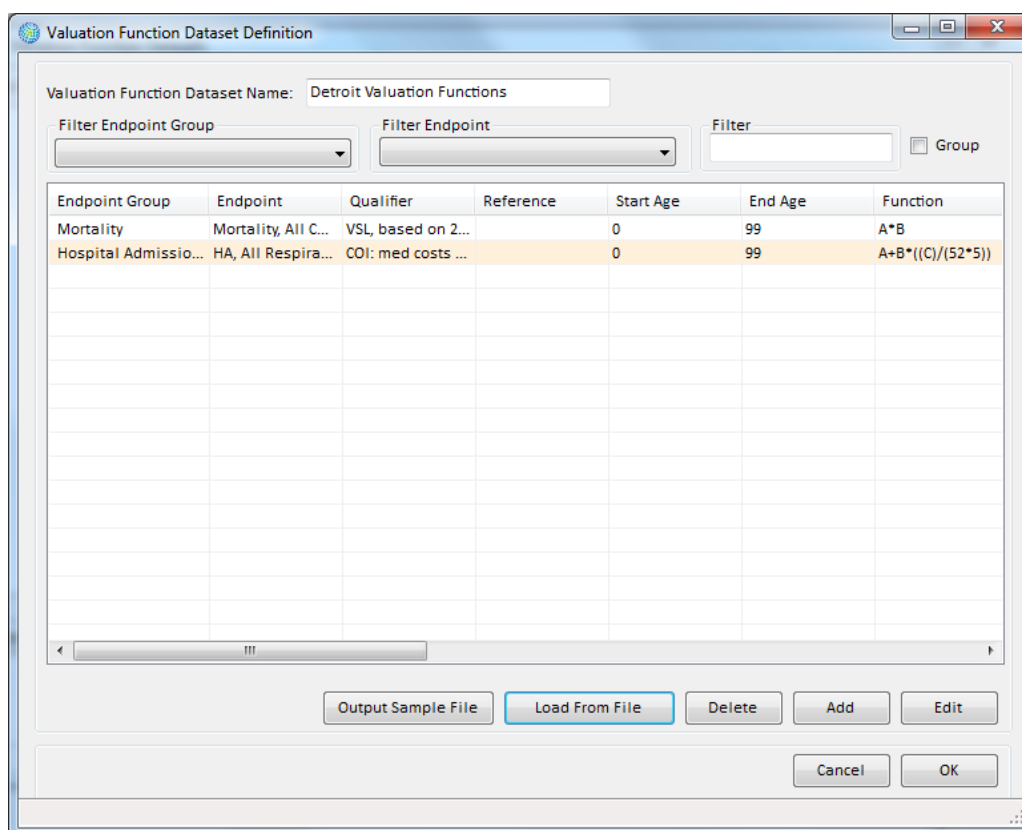
In the **Valuation Function Dataset Definition** window, you can also edit the functions already existing in your dataset by highlighting a particular **Valuation Function** and then clicking the **Edit** button.

If the dataset is large, there are filters available to view a subset of the list by selecting a value from the **Filter Endpoint Group** and/or **Filter Endpoint** drop-down lists. When you are finished, click **OK**.

From the **Valuation Function Dataset Definition** window you can also manually define a new **Valuation Function**. Click the **Add** button to open the **Valuation Function Definition** window where you can then create a new **Valuation Function**. (See Appendix M: Function Editor for additional information about the syntax for developing functions with this editor.)

After defining the new **Valuation Function**, click **OK**. This will take you back to the **Valuation Function Dataset Definition** window. When you are finished, click **OK**.

This will return you to the **Manage Valuation Function Datasets** window. From this form you can **Add**, **Delete**, or **Edit** the available datasets. In addition, you can view or edit the previously created metadata for a file by selecting an **Available Dataset** and **Valuation Function** and then click the **View Metadata** button. When you are satisfied with the inputs, click **OK**. The **Modify Datasets** window will appear. Here in the **Valuation Function** datasets box you should see any updates to the **Valuation Function** dataset that you just made.

4.1.9.2   Format for Valuation Data

Table 4-14 presents the variables that can be used in **Valuation Datasets**.

Table 4-14. Valuation Dataset Variables

Field Name	Type	Required	Notes
Endpoint Group	Text	Yes	If this doesn't reference an already-defined Endpoint Group, one will be added.
Endpoint	Text	Yes	If this doesn't reference an already-defined Endpoint for the Endpoint Group, one will be added.
Qualifier	Text	No	Provides additional information to identify a particular valuation function.
Reference	Text	No	Bibliographic reference, included to identify the source in the economic literature.
Start Age	Integer	No	Specifies the low and high ages, inclusive. For example, Start Age of '*0*' and End Age of '*1*' includes infants through the first 12 months of life and all one-year old infants.
End Age	Integer	No	
Point Estimate	Numeric (double)	Yes	Central estimate of the unit value.
Function	Text	Yes	The functional form, interpreted (executed) by BenMAP-CE when running an analysis.
A Description	Text	No	Description of variable A.
A	Numeric (double)	No	Mean of the A distribution.
A Distribution	Text	No	If A has no distribution, any value is acceptable. Otherwise, should be one of: *Normal*, *Triangular*, *Poisson*, *Binomial*, *LogNormal*, *Uniform*, *Exponential*, *Geometric*, *Weibull*, *Gamma*, *Logistic*, *Beta*, *Pareto*, *Cauchy*, *Custom*.
A Parameter 1	Numeric (double)	No	Parameter 1 of the A distribution (meaning depends on the distribution - for *Normal* distributions this represents the standard deviation).
A Parameter 2	Numeric (double)	No	Parameter 2 of the A distribution (meaning depends on the distribution - for *Normal* distributions this is not required).
Constant Description	Text	No	Description of variables, B, C, and D.
Constant Value	Numeric (double)	No	A constant value (denoted by B, C, and/or D) which can be referenced by the Function.

4.1.10   Income Growth Data

According to economic theory, willingness-to-pay (WTP) to avoid air-pollution related morbidity effects and premature mortality should grow as real income increases. BenMAP-CE allows users to adjust the WTP estimates to account for the growth in income over time. This adjustment is a combination of data on income growth and estimated income elasticity of demand, which measures the responsiveness of the quantity demanded of a good to the change in the income of the people demanding the good; this is distinct from elasticity of demand, which quantifies the change in demand for goods and services as a result of changes in price for those goods and services. This section describes how to load the data adjusting for income growth and how EPA developed these adjustment factors.

Note that the WTP estimates in the default valuation functions in the *United States* setup are assumed to be based on 1990 income, so the income growth adjustments are all relative to 1990. That is, the income growth data has a value of 1 in 1990, and because income has generally increased over time in the U.S., the income growth values are typically greater than 1 after 1990. (An exception is 1991, when incomes declined slightly in the U.S.)

If you load in your own valuation functions and/or income growth adjustment factors, be sure that you have carefully considered the income year. For example, if your valuation functions are based on income in the year 2005, then the income growth adjustment should have a value of 1.0 in 2005, because no adjustment is necessary. As you forecast into the future, under the assumption that incomes go up over time, then your income growth adjustment factors would have values greater than 1.0 for years past 2005, and would have values less than 1.0 for years prior to 2005.

4.1.10.1   Add Income Growth Data

To add income growth data to BenMAP-CE, click on the **Manage** button below the **Income Growth Adjustments** box in the **Modify Datasets** window. The **Income Growth Adjustment Dataset Manager** window will appear.
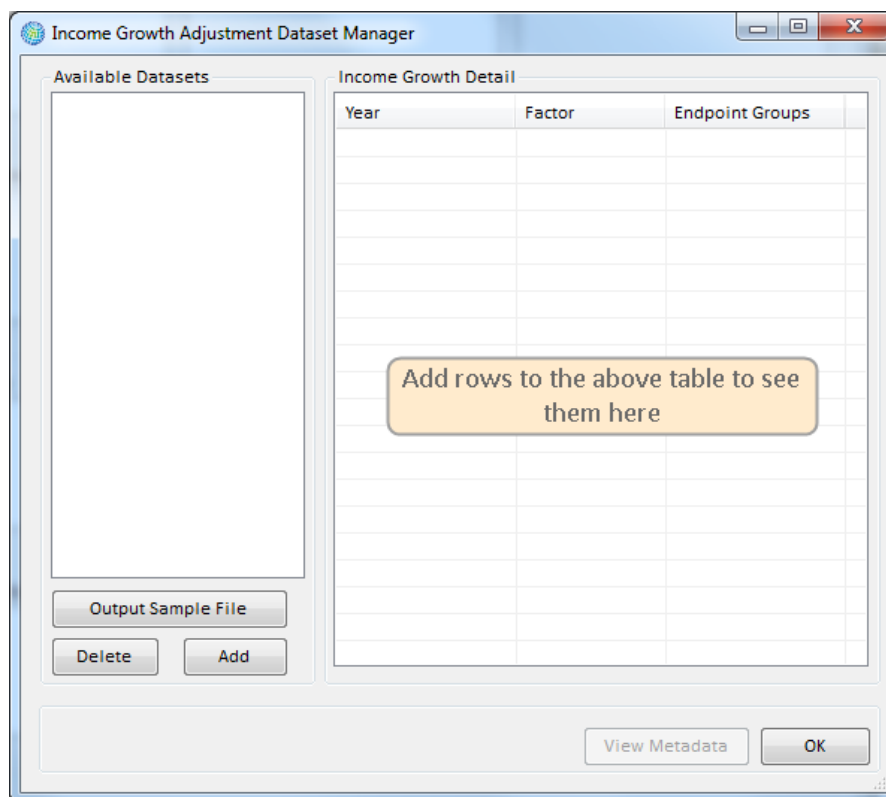
In this window you may **Add** and **Delete** datasets. The section on the left under **Available Datasets** lists the **Income Growth Adjustments Datasets** that are currently in the setup database. The section on the right under **Income Growth Detail** presents the income growth adjustment factors in a selected dataset.

Click on the **Add** button. In the **Load Income Growth Adjustment Factors Dataset** window, type in the name of the dataset in the **Income Growth Adjustment Dataset Name** box, and then click on the **Browse** button to the right of the **Database** box to choose the dataset that you want to import and click **Open**. To confirm the file has proper formatting, click the **Validation** button before importing the file. In addition, you can edit the metadata for the imported file, to include references and descriptions, by clicking the **View Metadata** button.

Click **OK**. The **Income Growth Adjustment Dataset Manager** window will appear. Here you may view the data that you just loaded. From this window, you can also view or edit metadata by selecting an **Available Dataset** and **Income Growth Detail** and then clicking the **View Metadata** button.

4.1.10.2   Format for Income Growth Adjustment Data

Table 4-15a presents the variables that can be used in **Income Growth Adjustment Datasets** and Table 4-15b presents a sample of what a dataset might look like.

Table 4-15a. Income Growth Adjustment Dataset Variables

Variable	Type	Required	Notes
Year	Integer	Yes	The year of the data. Note that this will include historical estimates as well as forecasts.
Mean	Numeric (double)	Yes	Mean income growth adjustment factor.
EndPointGroup	Text	Yes	Endpoint group (e.g., *Chronic Asthma*).

Table 4-15b. Sample Income Growth Adjustment Dataset

Year	Mean	EndpointGroup
1990	1	Acute Bronchitis
1991	0.99788719	Acute Bronchitis
1992	1.0010463	Acute Bronchitis
1993	1.00313532	Acute Bronchitis
1994	1.00725281	Acute Bronchitis
1995	1.00950849	Acute Bronchitis
1996	1.01338995	Acute Bronchitis
1997	1.01823652	Acute Bronchitis
1998	1.02308881	Acute Bronchitis
1999	1.02832949	Acute Bronchitis
2000	1.03276563	Acute Bronchitis
2001	1.0327363	Acute Bronchitis
2002	1.03402805	Acute Bronchitis
2003	1.03695309	Acute Bronchitis
2004	1.04123259	Acute Bronchitis
2005	1.04486656	Acute Bronchitis
2006	1.04742563	Acute Bronchitis
2007	1.04867125	Acute Bronchitis
2008	1.04677248	Acute Bronchitis
2009	1.04110742	Acute Bronchitis
2010	1.04366136	Acute Bronchitis
2011	1.0449264	Acute Bronchitis
2012	1.04712892	Acute Bronchitis
2013	1.04852092	Acute Bronchitis
2014	1.05097616	Acute Bronchitis
2015	1.05361962	Acute Bronchitis
2016	1.05460298	Acute Bronchitis
2017	1.05676126	Acute Bronchitis
2018	1.0586977	Acute Bronchitis
2019	1.05993581	Acute Bronchitis
2020	1.06099141	Acute Bronchitis
2021	1.06246126	Acute Bronchitis
2022	1.06402493	Acute Bronchitis
2023	1.06561422	Acute Bronchitis
2024	1.06718934	Acute Bronchitis
2025	1.0687362	Acute Bronchitis
2026	1.07025254	Acute Bronchitis

4.2 Export and Import Setups

BenMAP-CE allows you to export and import entire databases (all **Available Setups**), individual setups (e.g., *United States, China*), and parts of individual setups (e.g. all **Grid Definitions**, or individual **Health Impact Function** datasets). This functionality can be used to archive data, share data with other BenMAP-CE users, as well as to allow you to move databases between computers or between versions of BenMAP (currently only supported for versions 1.3 and 1.4). In particular, all of the steps involved in creating a setup can be done just once, after which the data can be exported and then imported on

other computers. See Chapter 9: Tools Menu for more information about the *Database Export* and *Database Import* tools.

4.3   Frequently Asked Questions

I've loaded new baseline incidence rates, but BenMAP-CE won't let me select it in the configuration stage.

When formatting these data for importing to BenMAP-CE, take special care to ensure that you have specified the health endpoints correctly. The baseline incidence rate must be associated with a specific health endpoint and endpoint group in BenMAP-CE. Be sure that you have recorded the endpoint group and endpoint exactly as it is recorded in BenMAP-CE. For example, if the baseline incidence rate is for asthma-related hospital admissions, be sure you have recorded the endpoint group as '*Hospital Admissions, Respiratory*' and the endpoint as '*HA, Asthma*'.

I've loaded a new grid and new population data into BenMAP-CE but I can't seem to use these new data.

Be sure to load the new grid definition first. Next, load the population dataset and be sure to select your new grid definition.

How do I generate a population dataset for a new grid definition?

You can generate a population dataset using a variety of approaches. The key is that you need to have a shapefile of your area of interest (e.g., Census tracts in a city) and you need to have census data matching your area of interest. One source for both a shapefile and the associated population data is the U.S. Census Bureau. (A variety of other agencies have census data, and you need to check around for your area of interest.) Another option for U.S. population data is to use the PopGrid software application, mentioned in Section 4.1.5 and described in Appendix J on Population Data for the U.S. Setup. Using PopGrid, you still need to have a shapefile for your area of interest.

Can I edit a population configuration?

No, you cannot edit a population configuration. You can only view a population configuration. If the population configuration does not match your data, you need to either create a new population configuration to match your data, or reshape your data so that it matches the population configuration.

Chapter 5
Creating Air Quality Surfaces

In this chapter...

- **Define air quality grids**
- **Create air quality surfaces using different methods.**
- **Learn how to structure input files.**
- **Learn how to interpolate monitoring data with Closest Monitor, Voronoi Neighborhood Averaging, or Fixed Radius.**
- **Learn about the Monitor Rollback feature.**

Chapter 5 Table of Contents

BenMAP-CE is not an air quality model and it cannot generate air quality data independently. Instead it relies on the air quality inputs given to it. To estimate population exposure to air pollution, BenMAP-CE uses air quality surfaces that it generates from input air quality data (modeling or monitoring data). Essentially, the air quality surfaces can be described as air quality grids (the structure) that have been populated with air pollution values (the data).

BenMAP-CE creates air quality surfaces to estimate the average exposure to ambient air pollution of people living in a "grid." These grids are either regularly shaped areas like those used by air quality models, or irregular shapes, like provinces, local government areas, cities, or nations. BenMAP-CE does not estimate personal exposure. Instead, the program calculates average pollutant concentration to which people are exposed in each grid cell. BenMAP-CE then uses these average values to calculate health impact functions.

To create air quality surfaces, BenMAP-CE uses a number of inputs, including modeling data or monitoring data. You may enter your own modeling and monitoring data, provided that the data are in a format recognized by BenMAP-CE.

To start the grid creation process, locate **Air Quality Surfaces** on the BenMAP-CE tree menu. Under this header, double-click **Pollutant**. On the selection window, click to select a pollutant from the left site and click "Add" button to add it to the right side. You may also click, hold and drag the pollutant of choice from the left side to the right side. Click **OK**.

To remove a selected pollutant, click the Remove button or double click the pollutant name to remove it from the left site.

Next, double-click **Baseline** on the BenMAP-CE tree menu to open the **Grid Creation Method** window.
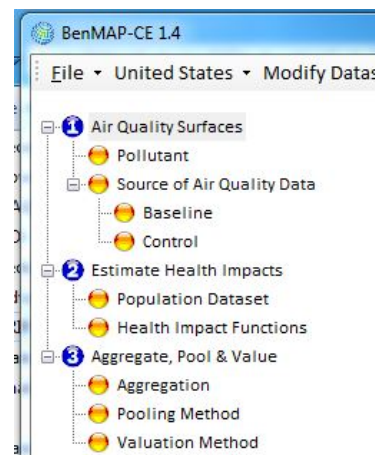
BenMAP-CE will then ask which **Grid Type** (previously loaded shapefile) to use and which of the following types of air quality data you wish to use:

- *Model Data*. Choose this option if you have air quality modeling data that you wish to use directly. Table 5-1 below describes the input format for modeling data.

- *Monitor Data*. Choose this option if you wish to use point source monitoring data (measured observations).

- *Monitor Rollback*. Choose this option if you want to reduce monitor levels by a specified amount.

- *Open *.aqgx file.* Choose this option to import a file that has already been created.
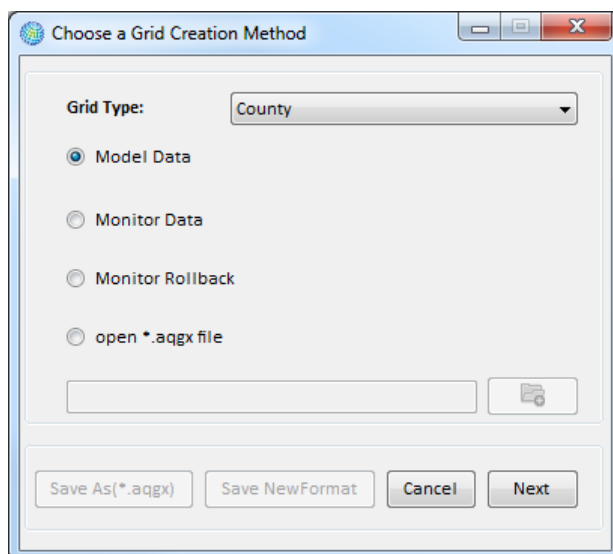
Select your **Grid Type** and then click **Next.** BenMAP-CE will direct you through the necessary steps for each option (described below).

5.1 Model Data

After choosing the *Model Data* option, use the **Generic Model Database** tab to load grid-definition-based modeling data (e.g., CMAQ or CAMx).[1] *Note: The **New Format Database** is designed to support a new format model (under development).*
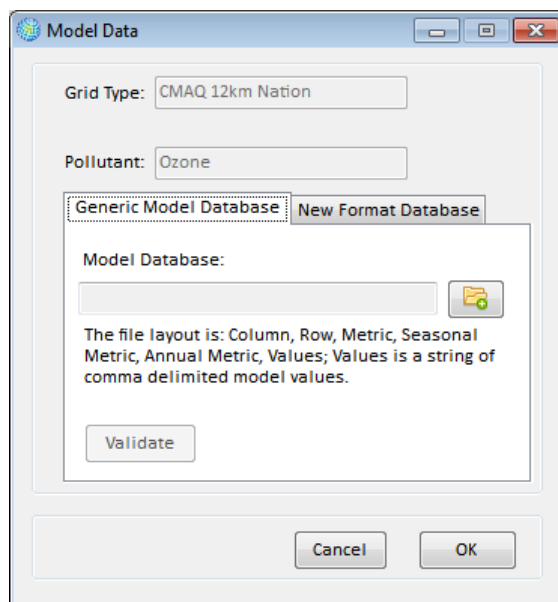
---

[1] Community Multi-scale Air Quality (CMAQ) Model is available at: http://www.epa.gov/amad/Research/RIA/cmaq.html or https://www.cmascenter.org/cmaq/. Comprehensive Air Quality Model with Extensions (CAMx) is available at: http://www.camx.com/.

The **Model Database** specifies the location of the air quality model results that you want to import. Table 5-1 presents the structure that these files must have, and Table 5-2 presents a sample data file with a variety of metrics. (For more information on air quality models, the EPA website has detailed descriptions of a variety of models at https://www.epa.gov/scram.)

Table 5-1. Air Modeling Data File Variables

Variable	Type	Required	Notes
Column	Integer	Yes	The column and the row uniquely identify each set of modeling values, and link the modeling data with cells in a Grid Definition.
Row	Integer	Yes	
Metric	Text	No	This variable is either blank (signifying that the Values are Observations, rather than Metric values), or must reference an already defined Metric (e.g., 1-hour daily maximum) for the appropriate Pollutant.
Seasonal Metric	Text	No	This variable is either blank (signifying that the Values are not Seasonal Metric values) or must reference an already defined Seasonal Metric for the Metric (e.g., mean of the 1-hour maximum values for the months of June through August).
Annual Metric	Text	No	This variable is either blank (signifying that the values are not an annual metric) or must be one of: *None*, *Mean*, *Median*, *Max*, *Min*, *Sum* (e.g., mean of the 1-hour maximum values for the year)
Values	Comma Separated Values (Text)[2]	Yes	If Metric is blank, values are supplied as a comma-delimited string of values for the year [e.g., 365 or 366 (leap year) values for daily data, 8760 or 8784 (leap year) values for hourly data]. If Metric is defined, but Seasonal Metric and Annual Metric are blank, 365 or 366 metric values. If Seasonal Metric is defined, but Statistic is blank, *n* seasonal metric values. If Annual Metric is defined, one annual statistic value for either the Metric (if Seasonal Metric is blank) or the Seasonal Metric. Missing values are signified with a period ('.').

---

[2] The list of comma-separated values must be surrounded by quotation marks, otherwise only the first value in the list will be used.

Table 5-2. Sample Air Modeling Data File

Column	Row	Metric	Seasonal Metric	Annual Metric	Values
24	101				.,.,.,.,.,.,.,.,.,.,.,.,.,5.90,.,.,11,.,.
24	102				19.80,.,.,3.5,.,.,14.70,.,.,.,.,.,4.90
24	103				11.5,.,.,1.5,.,.,10.20,.,.,2.5,.,.,2.6
24	104				12.60,.,.,4.60,.,.,7.30,.,.,.,.,.,9.20
24	105	DailyAverage			.,.,.,.,.,.,.,.,.,.,.,.,.,5.90,.,.,11,.,.
24	106	DailyAverage			19.80,.,.,3.5,.,.,14.70,.,.,.,.,.,4.90
24	107	DailyAverage			11.5,.,.,1.5,.,.,10.20,.,.,2.5,.,.,2.6
24	108	DailyAverage			12.60,.,.,4.60,.,.,7.30,.,.,.,.,.,9.20
24	109	DailyAverage	QuarterlyAverage		11.82,13.24,18.79,14.25
24	110	DailyAverage	QuarterlyAverage		12.31,13.89,18.27,21.19
25	101	DailyAverage	QuarterlyAverage		12.33,15.68,18.49,15.96
25	102	DailyAverage	QuarterlyAverage		13.52,13.69,19.04,22.43
25	103	DailyAverage		Mean	14.52
25	104	DailyAverage		Mean	16.41
25	105	DailyAverage		Mean	15.62
25	106	DailyAverage		Mean	17.17
25	107	DailyAverage	QuarterlyAverage	Mean	14.29
25	108	DailyAverage	QuarterlyAverage	Mean	18.97
25	109	DailyAverage	QuarterlyAverage	Mean	20.14
25	110	DailyAverage	QuarterlyAverage	Mean	16.46

5.2   Monitor Data

Using the *Monitor Data* grid creation option, you create an air quality grid directly from air pollution monitoring data. At the top left of the **Monitor Data** window, you will see the previously selected grid definition in the **Grid Type** field, and the previously selected pollutant in the **Pollutant** field. Below the **Pollutant** field of the **Monitor Data** window, you are asked to select an **Interpolation Method**. The interpolation method is used to move from point-based monitor data to grid cell based air quality data. That is, some grid cells will have many monitors in them, some will have just one, and some will have none. BenMAP-CE uses the interpolation methods to generate representative air quality metric values for each grid cell from monitor data for all of these cases.
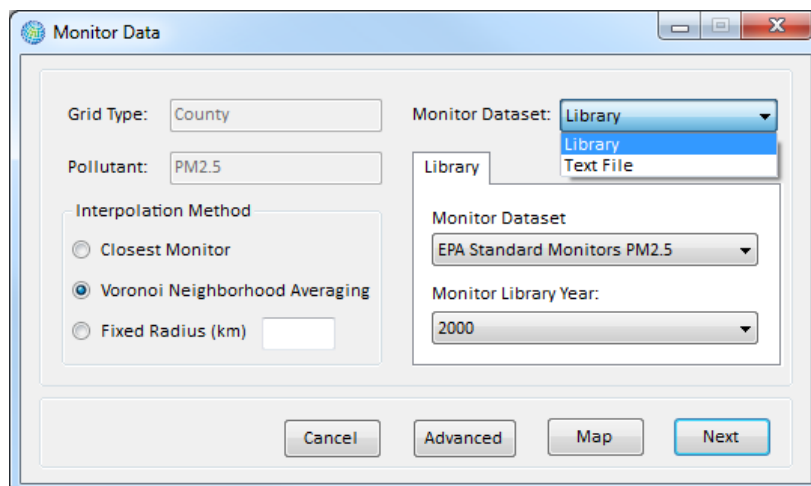
BenMAP-CE includes three **Interpolation Methods**. The *Closest Monitor* method simply uses the monitor closest to a grid cell's center as its representative value. The *Voronoi Neighborhood Averaging* method takes an inverse-distance weighted average of a set of the monitors surrounding a grid cell's center as its representative value. The *Fixed Radius* method averages all of the monitors within a fixed (user-specified) radius measured from the center of the grid cell. Each method is described below. For more detail, also see Appendix B on Air Pollution Exposure Estimation Algorithms.

On the right side of the **Monitor Data** window, you can specify a source for your monitor data. Using the **Library** tab, you may select from data that you have already loaded into BenMAP-CE. Choose the **Monitor Dataset** and **Monitor Library Year** from the drop-down lists.

If you want to load your own monitor data, choose *Text File* from the **Monitor Dataset** drop-down list. You can then browse to locate the data file you want to load. See Chapter 4 for more information about formatting monitor datasets.

At the bottom of the **Monitor Data** window, there is a **Map** button. The **Map** button opens the **Monitor Map** window, allowing you to preview the map that you are about to load.

5.2.1 Closest Monitor for Monitor Data

If you choose the *Closest Monitor* option, BenMAP-CE identifies the monitor closest to each grid cell's center, and then assigns that monitor's data to the grid cell. *Closest Monitor* interpolation can be modified by clicking on the **Advanced** button at the bottom of the window and typing in a **Maximum Neighbor Distance (in km)**.
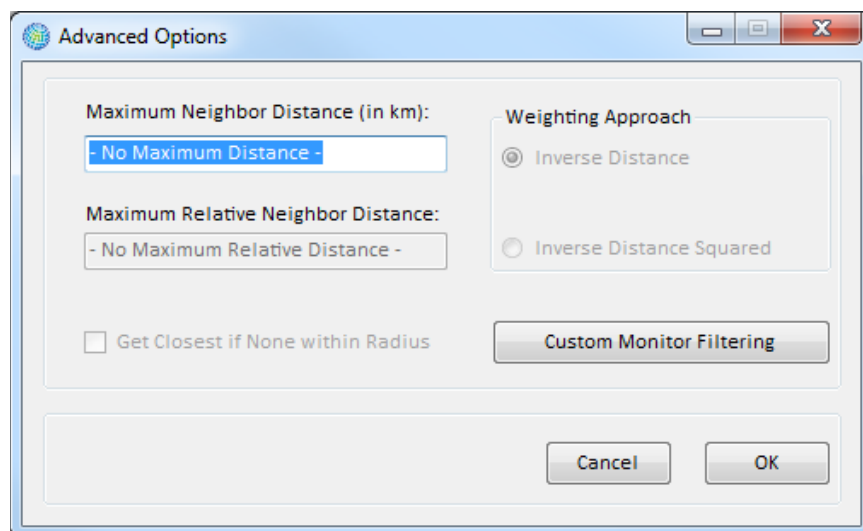
The **Maximum Neighbor Distance** specifies the maximum distance (measured in kilometers) that a monitor may be from the center of a grid cell. Cells without any monitors within this distance will not be included in the resultant air quality grid. The default setting is infinite (i.e., no limit to the monitor's distance from the center of the grid cell).

Note: The **Maximum Relative Neighbor Distance** and the **Weighting Approach** options are irrelevant (and are therefore disabled) when using the *Closest Monitor* method, since BenMAP-CE is only choosing a single monitor to assign to any given grid cell.

5.2.2 Voronoi Neighborhood Averaging (VNA) for Monitor Data

If you choose the *Voronoi Neighborhood Averaging* option, BenMAP-CE first identifies the set of monitors that "surround" each grid cell's center (these monitors are referred to as the grid cell's neighbors), and then BenMAP-CE calculates an inverse-distance weighted average of these neighboring monitors. In this section, we provide some examples of the different ways that BenMAP-CE calculates the average of the neighboring monitors. See Appendix B on Air Pollution Exposure Estimation Algorithms for an expanded discussion of VNA, including how the VNA algorithm actually chooses the neighbor monitors, as well as the different ways that it may be used.

VNA interpolation has three advanced interpolation options, which can be modified by clicking on the **Advanced** button at the bottom of the window:

- **Maximum Neighbor Distance (in km)** specifies the maximum distance that a monitor may be from the center of a grid cell, and still be included in the set of neighbor monitors used to calculate air pollution exposure at a particular grid cell. The default setting is infinite (i.e., no limit to the monitor's distance to the center of the grid cell).

- **Maximum Relative Neighbor Distance** specifies the maximum ratio for the distance of each included monitor to the distance of the closest monitor. The default setting is infinite.

- **Weighting Approach** specifies whether BenMAP-CE should use inverse-distance weighting for the monitors, or inverse-distance-squared weighting of the monitors. The default setting is inverse-distance weighting.

The following examples illustrate how varying these options affects the final average concentration estimate.

Example 1: Monitor Data VNA method

Default options

Consider the following example at a hypothetical rural grid cell, where there are relatively few monitors, and where the distance from a monitor to the center of a grid cell can be fairly large. With *VNA*, BenMAP-CE first identifies the set of "neighbor" monitors for each grid cell. The number of neighbors is usually in the range of about three to eight. In this case, assume that there are five monitors at distances of 25, 50, 100, 200, and 400 km from the grid cell, with annual $PM_{2.5}$ levels of 8, 13, 12, 18, and 15 µg/m^3, respectively. BenMAP-CE would calculate an inverse-distance weighted average of the monitor values as follows:

$$PM_{2.5}\ average = \frac{\frac{1}{25} \bullet 8 + \frac{1}{50} \bullet 13 + \frac{1}{100} \bullet 12 + \frac{1}{200} \bullet 18 + \frac{1}{400} \bullet 15}{\frac{1}{25} + \frac{1}{50} + \frac{1}{100} + \frac{1}{200} + \frac{1}{400}} = 10.68$$

Example 2: Monitor Data VNA method

Maximum Neighbor Distance = 75

Using the same example that we used above, let us say you have specified a **Maximum Neighbor Distance** of 75 km, and left unchanged the default options (infinite value) for **Maximum Relative Neighbor Distance**. BenMAP-CE would only consider the first two monitors, and would calculate an inverse-distance weighted average of the monitor values as follows:

$$PM_{2.5}\ average = \frac{\frac{1}{25} \bullet 8 + \frac{1}{50} \bullet 13}{\frac{1}{25} + \frac{1}{50}} = 9.67$$

Example 3: Monitor Data VNA method

> **Maximum Relative Neighbor Distance = 10**

Alternatively, if you left the **Maximum Neighbor Distance** at the default (infinite), but have set the **Maximum Relative Neighbor Distance** to 10, then BenMAP-CE would calculate the ratio of the distance for each monitor to distance of the closest monitor. In this case, the ratios would be 1 (=25/25), 2 (=50/25), 4 (=100/25), 8 (=200/25), and 16 (=400/25), and BenMAP-CE would drop the monitor with a ratio of 16. BenMAP-CE would then calculate an inverse- distance weighted average of the monitor values as follows:

$$PM_{2.5} \ average = \frac{\dfrac{1}{25} \bullet 8 + \dfrac{1}{50} \bullet 13 + \dfrac{1}{100} \bullet 12 + \dfrac{1}{200} \bullet 18}{\dfrac{1}{25} + \dfrac{1}{50} + \dfrac{1}{100} + \dfrac{1}{200}} = 10.53$$

Example 4: Monitor Data VNA method

> **Inverse-distance squared neighbor weighting**

In addition, you can specify an inverse-distance-squared weighting of the monitors. Let us say that you have left unchanged the defaults (infinite values) for **Maximum Neighbor Distance** and **Maximum Relative Neighbor Distance**, and specified that the **Weighting Approach** is *Inverse Distance Squared*. BenMAP-CE would then calculate an inverse-distance-squared weighted average of the monitor values as follows:

$$PM_{2.5} \ average = \frac{\dfrac{1}{625} \bullet 8 + \dfrac{1}{2,500} \bullet 13 + \dfrac{1}{10,000} \bullet 12 + \dfrac{1}{40,000} \bullet 18 + \dfrac{1}{160,000} \bullet 15}{\dfrac{1}{625} + \dfrac{1}{2,500} + \dfrac{1}{10,000} + \dfrac{1}{40,000} + \dfrac{1}{160,000}} = 9.26$$

Example 5: Monitor Data VNA method

> **Maximum Neighbor Distance = 75**
> **Maximum Relative Neighbor Distance = 10**
> **Inverse-distance squared Weighting Approach**

Finally, you could specify changes to all three options: a **Maximum Neighbor Distance** of 75 km, a **Maximum Relative Neighbor Distance** of 10, and a **Weighting Approach** of *Inverse Distance Squared* weighting. BenMAP-CE would then calculate the following average:

$$PM_{2.5} \ average = \frac{\dfrac{1}{625} \bullet 8 + \dfrac{1}{2,500} \bullet 13}{\dfrac{1}{625} + \dfrac{1}{2,500}} = 9.00$$

5.2.3   Fixed Radius for Monitor Data

If you choose the *Fixed Radius (km)* option, BenMAP-CE averages all of the monitor values within a fixed radius (measured in kilometers) that you specify. The way that the monitor values are averaged depends on the **Weighting Approach** that you choose after clicking the **Advanced** button. You can choose either *Inverse Distance* or *Inverse Distance Squared* weighting.

Note that the default option with the *Fixed Radius* approach is that BenMAP-CE will not calculate an average for a grid cell if there are no monitors within the fixed radius (distance) that you specify. In the **Advanced Options** window, if you select *Get Closest if None within Radius*, then for those grid cells without any monitors within the fixed radius, BenMAP-CE will choose the nearest monitor (regardless of distance) and apply that value as the "average".

5.2.4   Custom Monitor Filtering for EPA Standard Monitors

Custom Monitoring Filtering options apply only to the *EPA Standard Monitors* library in the *United States* setup; these are the only monitoring values that include all of the variables that BenMAP-CE needs in order to filter the data properly. This tool allows you to filter, map, and export your monitor data. You can reach the **Custom Monitor Filtering** tool by first choosing your pollutant, data source (e.g., monitor library) and year on the **Monitor Data** window.
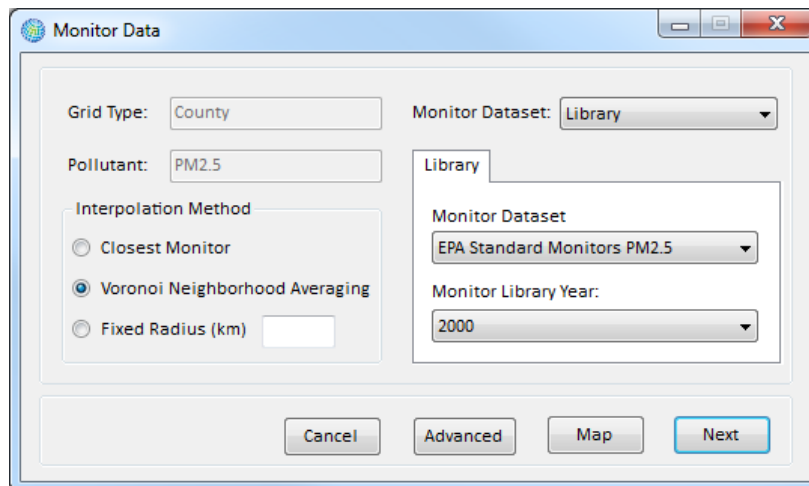
Click the **Advanced** button. This will take you to the **Advanced Options** window.
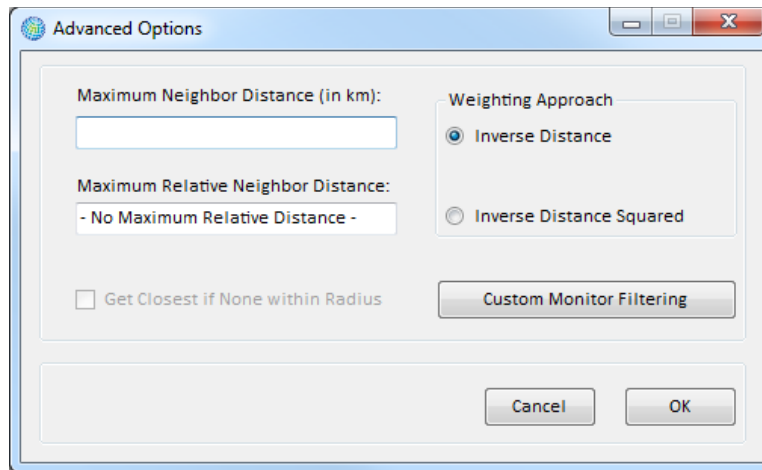
Click the **Custom Monitor Filtering** button. This will take you to the **Filter Monitors** window.
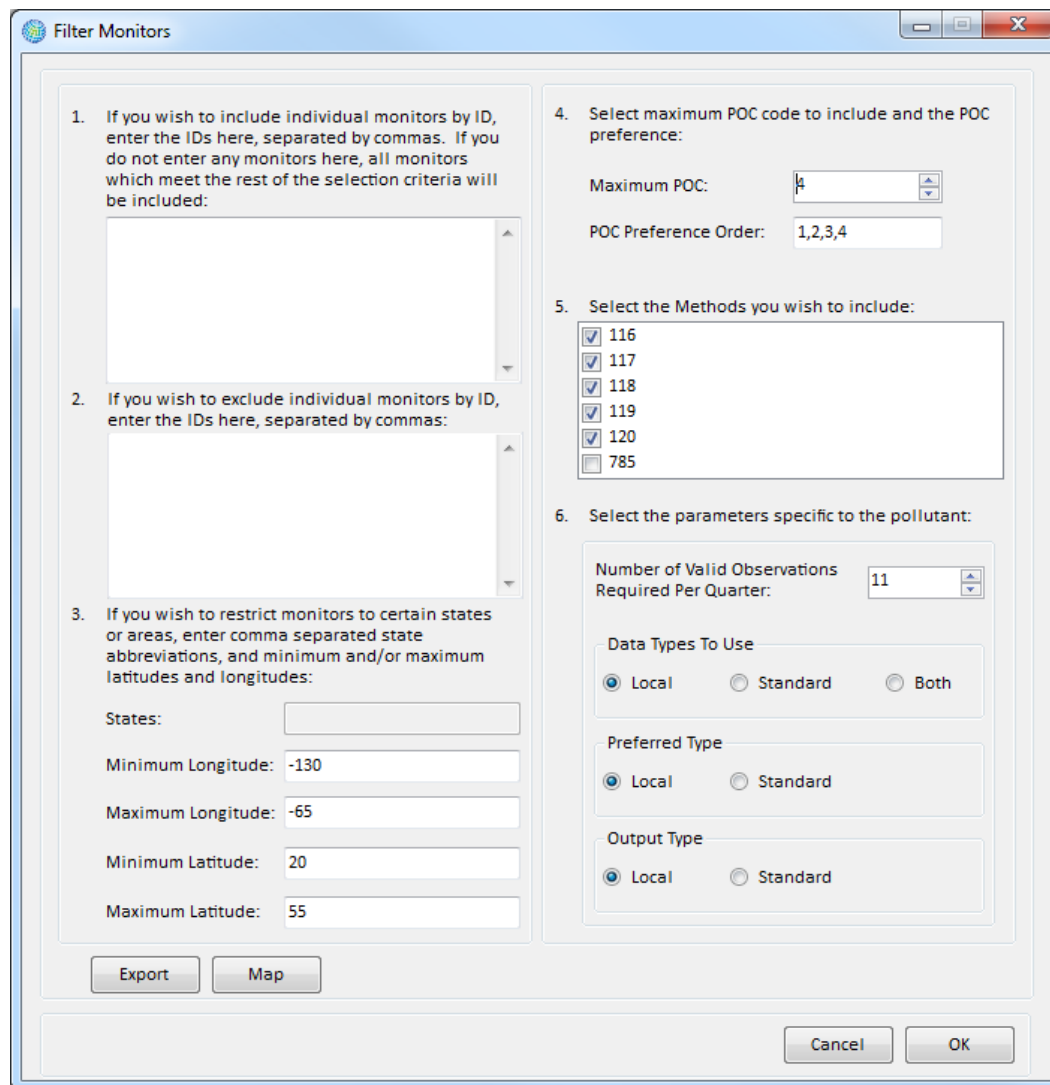
Note that the first five options are essentially the same for each pollutant, and the sixth option depends on the pollutant. The example above shows what the form looks like with *PM2.5* as the selected **Pollutant**.

1. Include specific monitors. Here you can specify particular monitor IDs that you want to include in your analysis. If this is left blank, then BenMAP-CE will include all monitors that meet the rest of the selection criteria.

2. Exclude specific monitors. Here you can exclude any particular monitor IDs from your analysis. Here again, if this option is left blank then BenMAP-CE will include all monitors that meet the rest of the selection criteria.

3. Restrict to particular states and/or latitude/longitude. You can choose monitors to include from particular states, by listing the two-character state abbreviation (e.g., CA = California). You can also choose monitors within particular latitude and longitude ranges. The default values for latitude (20 to 50) and longitude (-130 to -65) completely include the continental U.S. Here again, if this option is left blank then BenMAP-CE will include all monitors that meet the rest of the selection criteria.

4. Parameter Occurrence Code (POC)[3]. Sometimes, multiple monitors are collocated at the same site measuring the same parameter (e.g., to check precision). The **Maximum POC** specifies the highest POC value allowed in the data. The default is a value of 4. And to choose one monitor when more than one monitor is in the same location, the POC Preference Order specifies the preferred ordering of POC codes.

5. Methods.[4] The Method codes listed indicate U.S. EPA-defined methods for collecting and analyzing samples; these codes depend on the pollutant. In the case of PM$_{2.5}$, only federal-reference methods (FRM) are chosen by default -- specifically numbers 116 through 120. In the case of Ozone (O$_3$), all methods are checked by default.

6. Parameters Specific to the Pollutant. The default options vary by pollutant. Below, we have described the options that appear with *PM$_{2.5}$* and *ozone*.

- **PM$_{2.5}$ Monitor Filter:** The **Number of Valid Observations Required per Quarter** specifies the number of days of data needed. The default is to require 11 observations per quarter. The **Data Types to Use** options specify whether to use data at *Local* conditions (parameter code 88101), *Standard* conditions (parameter code 81104), or *Both*.[5] The default for PM$_{2.5}$ is to use data at *Local* conditions. When data at standard and local conditions are both available at the same monitor

---

[3] For information about data codes used in U.S. EPA's Air Quality System, refer to the *AQS Data Coding Manual*: https://www.epa.gov/sites/production/files/2015-09/documents/aqs_data_coding_manual_0.pdf
[4] For more information, refer to U.S. EPA's *List of Designated Reference and Equivalent Methods*: https://www3.epa.gov/ttnamti1/files/ambient/criteria/AMTIC%20List%20Dec%202016-2.pdf
[5] Particulate concentrations are expressed in either local conditions (volume is at temperature and pressure of the ambient sample) or at standard conditions (where the volume has been converted to standard conditions, typically 20° C at 760 mm Hg).

location, the **Preferred Type** allows you to choose which to use – the default is *Local*. The **Output Type** option is designed to allow you to make the data reasonably consistent when both local and standard condition data are used. The default is to use the *Local* output type, so *Standard* condition data will be converted to *Local*.

▪ **Ozone Monitor Filter:** The *ozone* specific options differ from $PM_{2.5}$ in large part because ozone is monitored hourly in the United States. The **Number of Valid Hours** specifies the number of hours needed for a particular day of monitoring to be considered "valid." BenMAP-CE counts the number of non-missing hourly values from the **Start Hour** through the **End Hour** and compares this number with that specified in the **Number of Valid Hours**.

The **Percent of Valid Days** specifies the percent of days between the **Start Date** and the **End Date** that need to be valid for the monitor itself to be considered valid. The default is 50 percent of the days between May 1 and September 30. The example below shows what the form looks like with *Ozone* as the selected **Pollutant**.

You can view a map of your data with the specified filter options by clicking the **Map** button on the bottom left side of the **Filter Monitors** window. This provides a brief preview of what will be shown once the data are loaded.

You can also export your data by clicking the **Export** button, which is found in the bottom left side of the **Filter Monitors** window. A **Save As** window will appear, allowing you to save the data as a CSV file.

5.3   Monitor Rollback

The *Monitor Rollback* option allows you to quickly test what the benefits would be from reducing historical monitor levels. BenMAP-CE has three methods to reduce, or "roll back," monitor data: *Percentage Rollback*, *Incremental Rollback*, or *Rollback to a Standard*. Each of these methods is depicted below. Note that with each of these methods you can use the same two interpolation algorithms (*Closest Monitor* or *VNA*) as you can use with *Monitor Data*.

Percentage Rollback reduces all monitor observations by the same percentage:
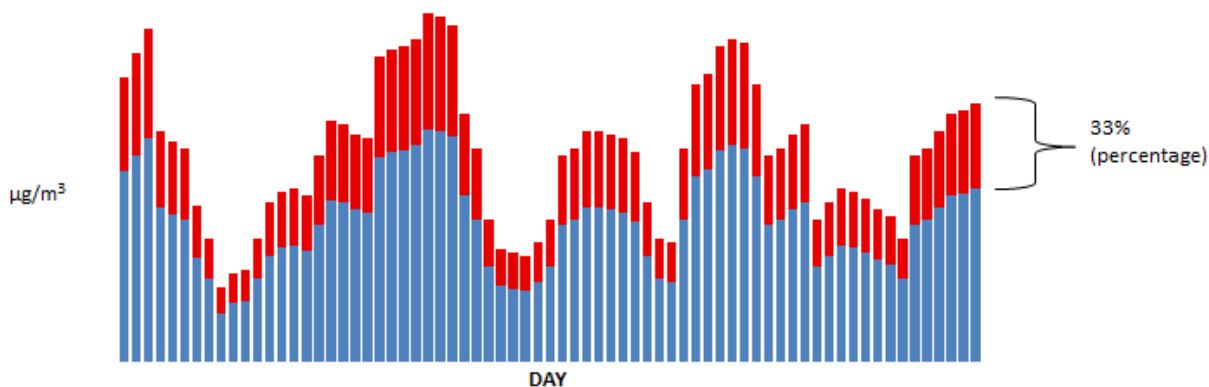
Incremental Rollback reduces all observations by the same increment:
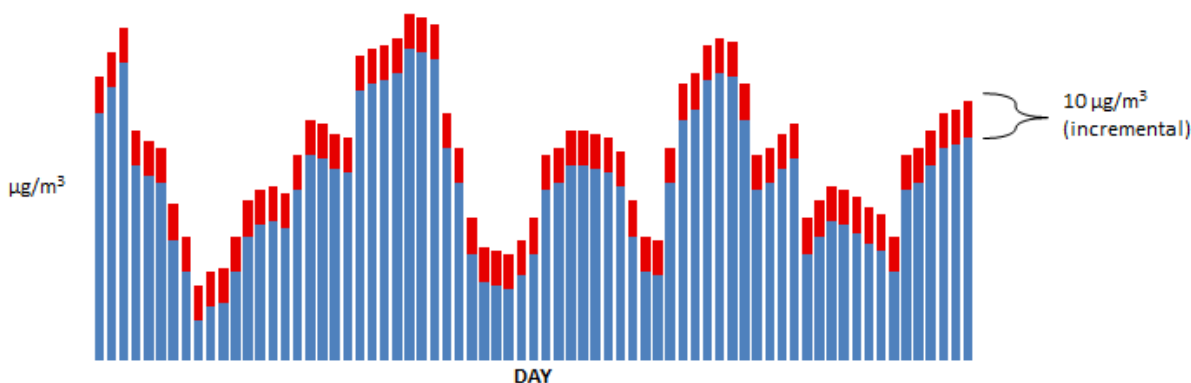
Rollback to a Standard lets you choose a standard, and then reduces any monitor observations exceeding the standard to the level of the standard:
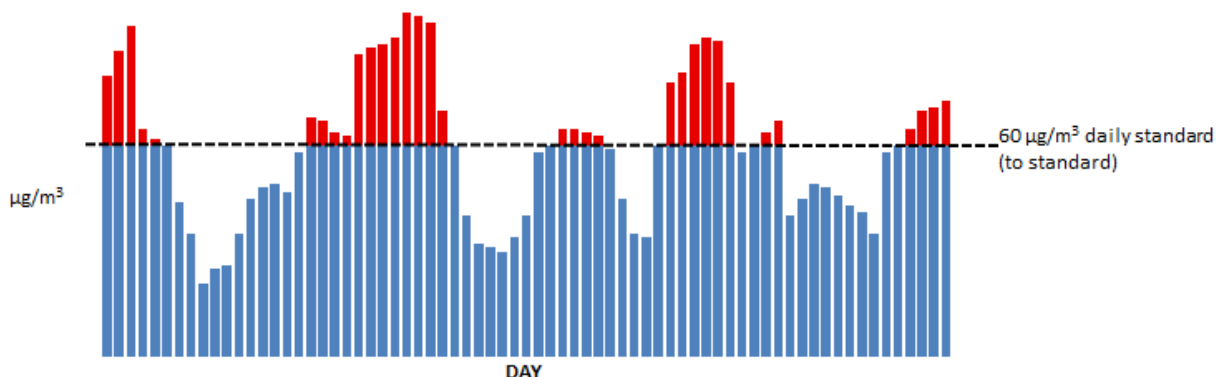
5.3.1   Example: A Single Rollback in One Region

To apply a monitor rollback, first click the **Create Air Quality Grids** button. On the **Air Quality Grid Creation Method** window, choose *Monitor Rollback.* Click **Next**.

There are three steps to the *Monitor Rollback* method.

1. Select Monitors. Choose the **Rollback Grid Type** from the drop-down list. This allows you to determine how detailed the rollback scenario may be. If the whole region (e.g., United States) will have the same type of rollback then you may simply choose a

grid outlining the area of interest. If you are interested in different rollbacks within a region, then you should choose a more finely detailed grid definition (e.g., states).

If you use data from an existing dataset, then choose the **Library** tab, and from the drop-down list choose the **Monitor Dataset** and the **Monitor Library Year**. If you want to use your own data, then choose the **Text File** tab. The file should have the monitor data format specified in Chapter 4: Loading Data.

When you have finished making your choices, click **Next**.
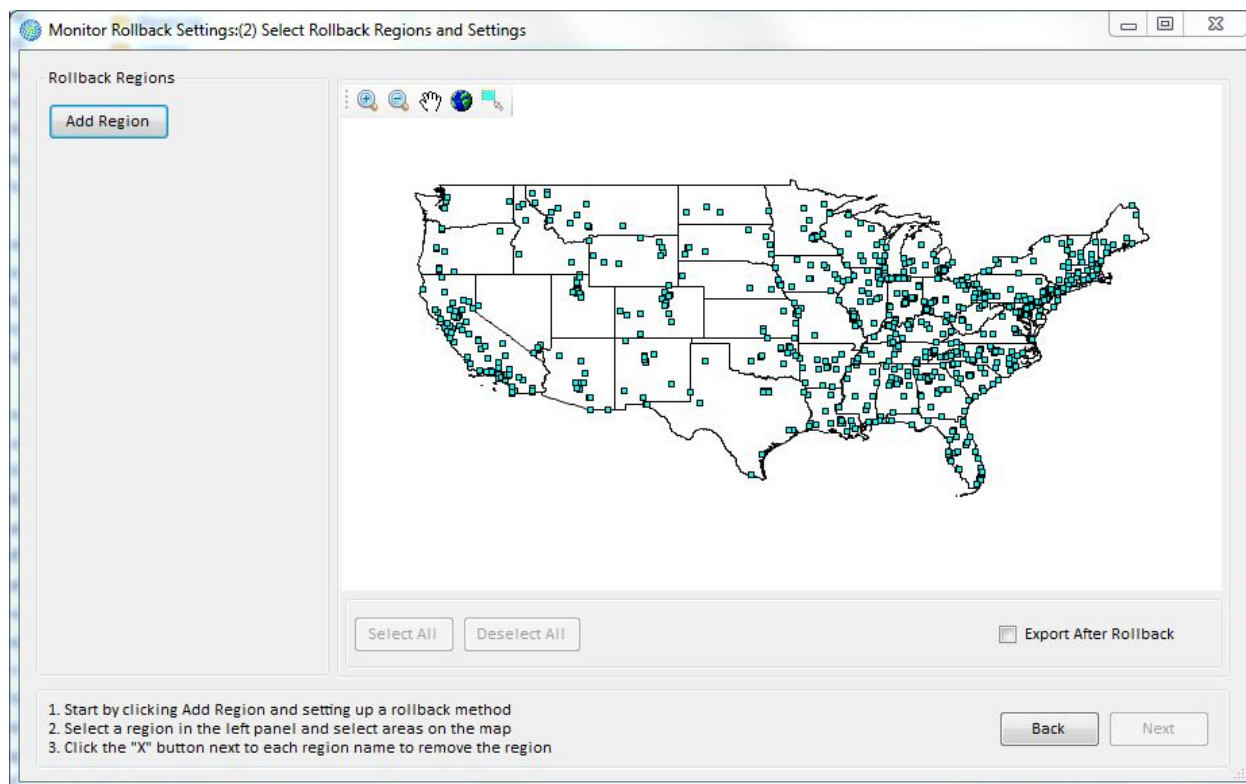
2. Select Rollback Regions and Settings. In this section, you can specify the type of the rollback method(s) that you would like to use, as well as the location of the monitors that you want to roll back.

Click the **Add Region** button to display the three rollback options: *Percentage Rollback*, *Incremental Rollback*, and *Rollback to a Standard*.
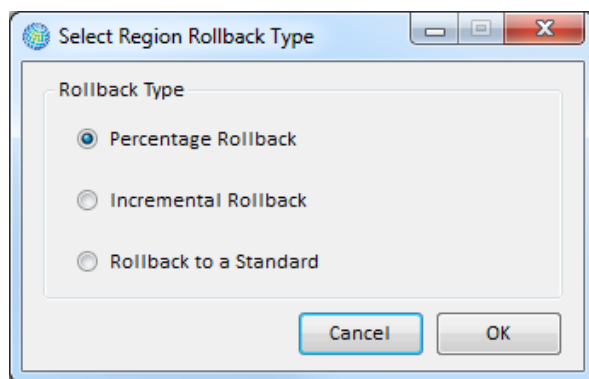
Choosing the **Rollback Type** and click **OK.** Then, specify the amount of the rollback and the region to which you want to apply it. You can click on the map to select and deselect the states (or other defined areas depending on your rollback grid type) to add to the region.

At the top of the map are five GIS toolbar icons, typically seen in mapping programs. The first four tools allow you to zoom in and zoom out, and to focus on the particular groups of grid cells that interest you. The fifth tool allows you to select grid cells graphically, by clicking or dragging a box over them.

In this example, we specified a 10 percent reduction, a background of 0 ppb, and applied it to all monitors in the state of California (by clicking on the particular state in the GIS window).
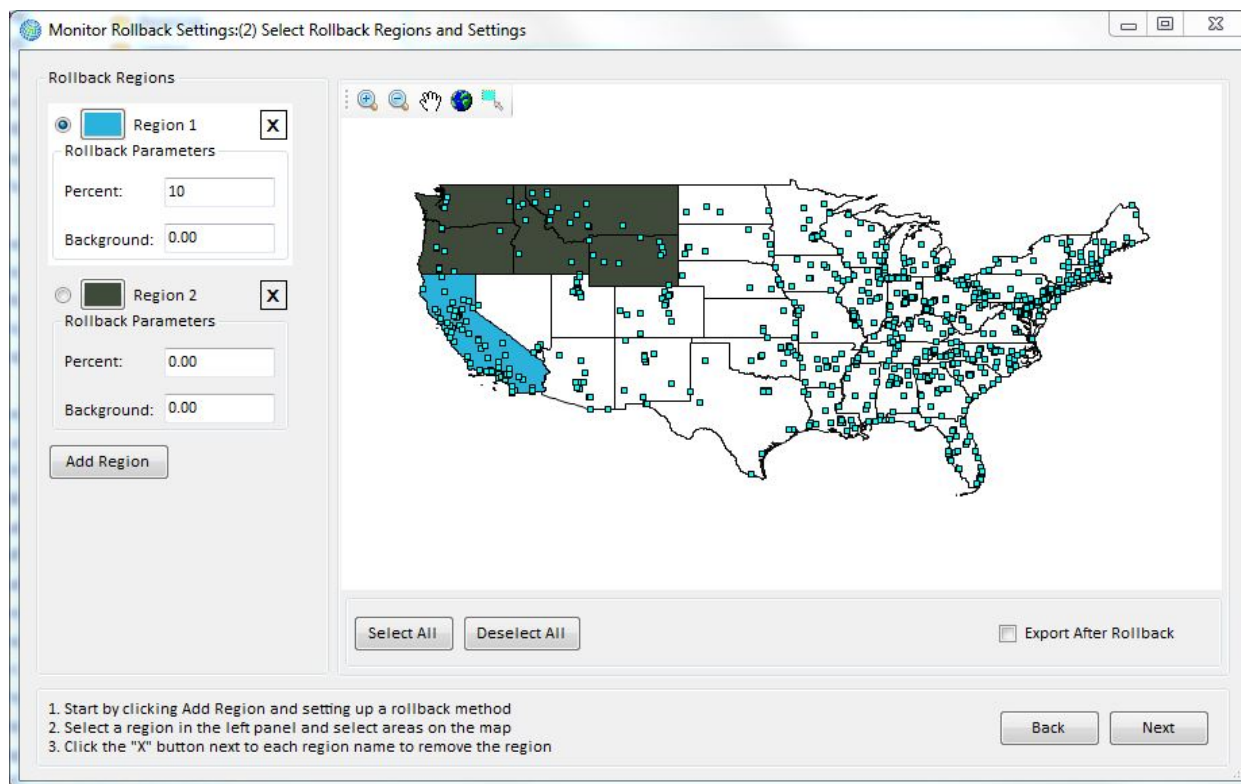
To apply the rollback to all the states, you can simply click on the **Select All** button. To clear the selections, click the **Deselect All** button. At any time, you can change the grid cells that you have selected. To remove a region from the selection, click the "X" button next to the region ID. This particular example is quite simple, so we will use a more complicated example below.

If you want to export your monitor data to a CSV file, click the option for *Export After Rollback*. The exported data will be formatted in the same specific format required of BenMAP-CE **Monitor Data** import files discussed in Chapter 4.

After defining the **Rollback Regions** and setting the **Rollback Parameters**, click on the **Next** button. BenMAP-CE will then perform the rollback you specified on the monitors in the grid cells that you have chosen.

3. Additional Grid Settings.

The third stage is similar to the *Monitor Data* grid creation method. As with monitor data, you need to specify the **Interpolation Method** (*Closest Monitor*, *VNA,* or *Fixed Radius*) and the **Grid Type**.

By checking the option for *Make Baseline Grid (in addition to Control Grid)* you tell BenMAP-CE to create a baseline grid at the same time as the control grid. The baseline grid uses the same parameters as the control grid, with the exception of the rollback. That is, the baseline uses the same monitor data, interpolation method, and the same

grid type. The two resultant grids will serve as both baseline and control scenarios and are automatically selected in the "Air Quality Surfaces" stage of the analysis.

Note that there is an **Advanced** button that lets you select the **Maximum Neighbor Distance (in km)**, **Maximum Relative Neighbor Distance**, and **Weighting Approach**. The specific availability of advanced features depends on the interpolation method that you choose. The **Advanced Options** window is described in more detail in Section 5.2.4. You can view a map of the inputs to the rollback grids that you are creating, as well as the grids themselves by accessing the **Monitor Map.** To do this, click on the **Advanced** button, then the **Custom Monitor Filtering** button, and then the **Map** button (on the **Filter Monitors** window).

5.3.2 Example: Combining Three Rollback Approaches in Different Regions

BenMAP-CE allows you to have different rollback approaches in different regions. In this example, we'll use the *United States* setup to combine the three rollback types: *Percentage Rollback*, *Incremental Rollback*, and *Rollback to a Standard*.

Start by clicking on the **Pollutant** button in the left-hand pane of the main screen, and choosing *Ozone*. Next, click **Baseline** under **Source of Air Quality Data**. Select *Monitor Rollback* as the **Grid Creation Method**. On the **Monitor Rollback: (1) Select Monitors** window, select the **Rollback Grid Type** (*State*), **Monitor Dataset** (*EPA Standard Monitors O3)* and **Monitor Library Year** (*2000*), and click **Next**.

On the **Monitor Rollback Settings: (2) Select Rollback Regions and Settings** window, click the **Add Region** button and select the *Percentage Rollback* method. Click **OK**. Enter *10* for the **Percent**. In the previous example, we included only one state in the rollback region. In this example we want to create three regions. Click on the three western-most states to add them to the first region. The states you have added to the region will fill in, as in the picture below.
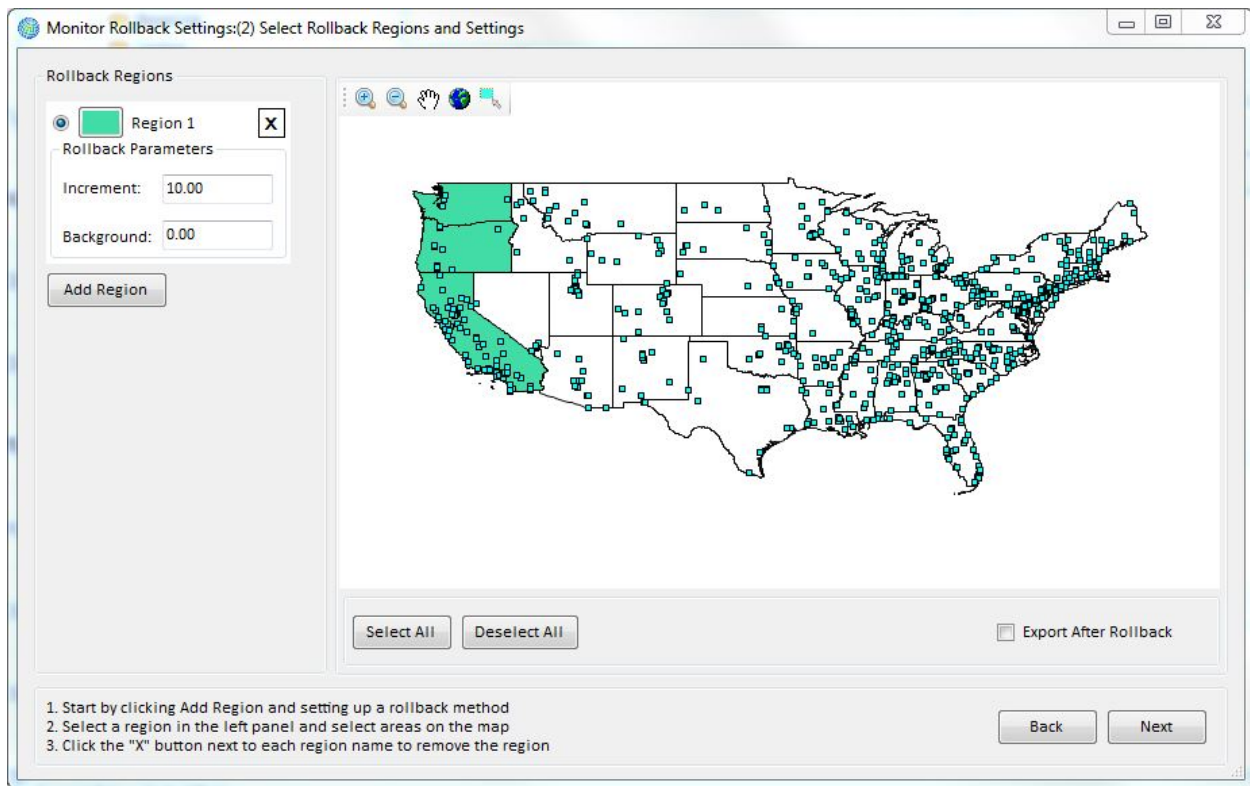
To add states with a second type of rollback, click on the **Add Region** button, choose the rollback type, and then click on the states to include in this second region, which BenMAP-CE denotes as Region 2. In this example, we have chosen an *Incremental Rollback* with an **Increment** of *5* and a **Background** of *0*, and applied it to the 14 next western-most states.
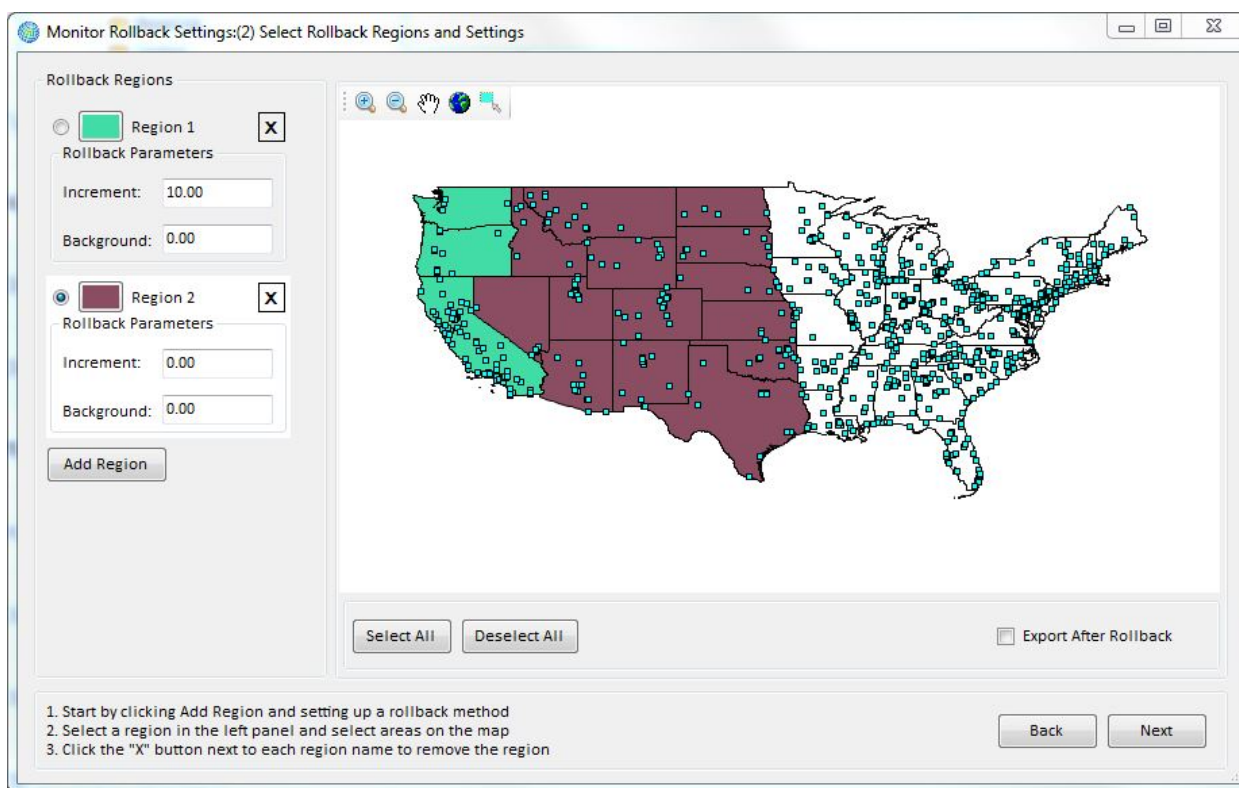
The map now depicts two rollback regions. We can toggle back and forth between each region by clicking on the button on the legend on the left side of the map. Any states that have not yet been included in a region may be added to an existing region, or we may create one or more regions for these remaining states. Note that once states have been included in a rollback region, they cannot be included in a different rollback region. In our example, the three western-most states are highlighted in gray.

If you want to add or remove states from a defined region, make sure you select the appropriate region by clicking on the button to the left of that region before clicking on the map to select or deselect the state(s).

To add a third rollback type covering the rest of the states, click again on the **Add Region** button, and then choose the rollback type. However, instead of individually choosing the states, simply click the **Select All** button. This will select all of the states that are not yet included in a region, and these remaining states will now become Region 3.

In this third region, we have chosen a *Rollback to a Standard*, which involves 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 **Daily Metric**, **Seasonal Metric**, **Annual Statistic**, **Ordinality**, and **Standard**. (Note: You will need to use the scroll bar to view more detail on the left side of the screen.)

In this step BenMAP-CE calculates the metric to be used to determine whether a monitor's values must be rolled back and, if so, how much (e.g., if **Metric** is *D8HourMax*, BenMAP-CE calculates the 8-hour daily maximum for each day at each monitor).

A monitor is considered in attainment if the nth highest value of a daily 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 *D8HourMax*, **Ordinality** is *4*, and **Standard** is *85*, a monitor will be considered in attainment if the fourth highest value of the eight-hour daily maximum is at or below 85 ppb.
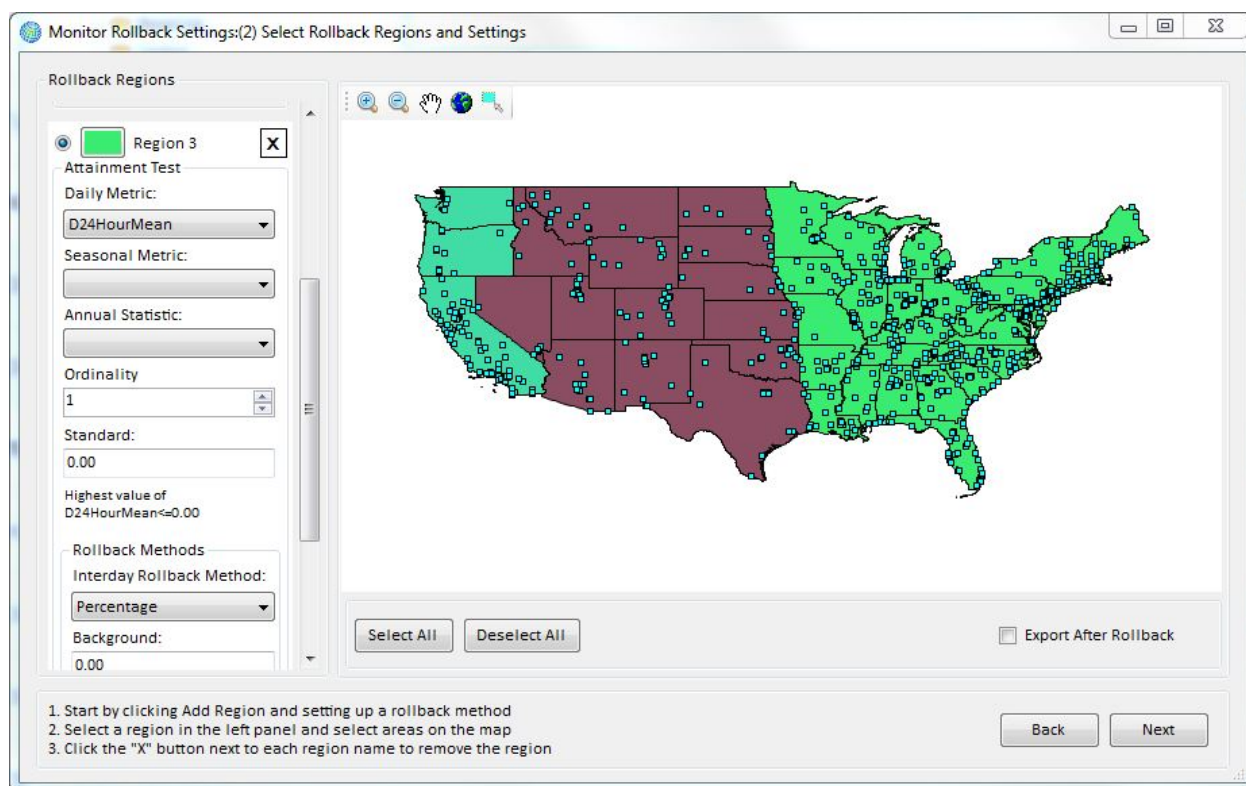
The **Attainment Test** can also be used for seasonal metrics (by choosing previously defined seasonal metrics from the drop-down list below **Seasonal Metric**), as well as for annual metrics (by using the drop-down list below **Annual Statistic**). For example, if you want the annual mean ozone level to stay below 60 ppb, then you would choose the daily 24 hour mean (*D24HourMean*) from the drop-down list below **Daily Metric**, choose *Mean* from the drop-down list below **Annual Statistic**, and set the **Standard** to *60*. (Note that in this case **Ordinality** cannot be chosen because there is only a single annual value.)

The **Rollback Methods** parameters determine the rollback procedures used to simulate out-of-attainment monitors coming into attainment:

- **Interday Rollback Method** (with associated **Background** level) – These are used to generate target values for the metric specified by the **Attainment Test**. Method types include *Percentage*, *Incremental*, and *Peak Shaving.*

- **Intraday Rollback Method** (with associated **Background** level) – These are used to adjust hourly observations to meet the target metric values generated in the previous step. Method types include *Percentage* and *Incremental.*

The methods involved for each can be somewhat complicated, so we have included a section in Appendix A: Monitor Rollback Algorithms which goes through several examples.

5.4   Open *.aqgx File

The final option for uploading an **Air Quality Grid** is to select the *Open *.aqgx File* option from the **Choose a Grid Creation Method** window.
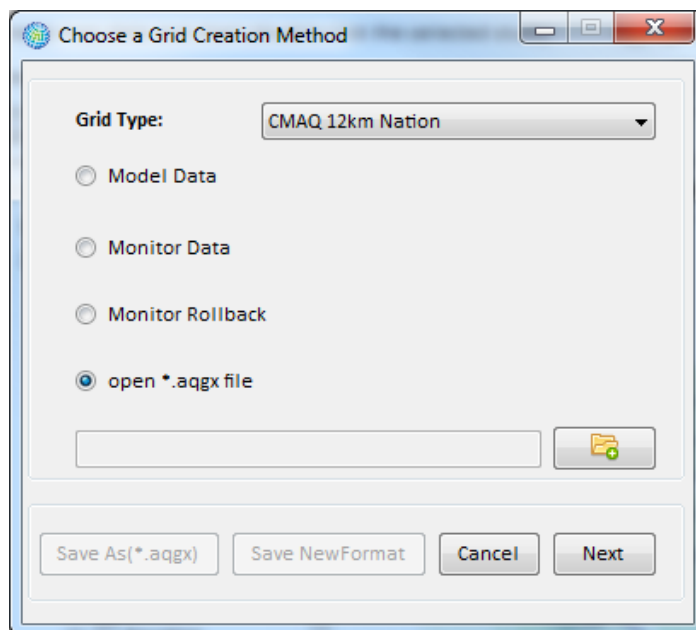
Choosing this option will activate the **Open File Browser** button located directly below this option. Click the **Open File Browser** button. This will cause an **Open** window to appear, allowing you to search for an **Air Quality Grid** (*.aqgx file) that has already been created. Select a file and click **Open**. The file path and name should appear in the box beside the **Open File Browser** button. Click **Next** to close the window and begin to create the map layer.

5.5   Frequently Asked Questions

How can I generate a map of an air quality grid and export it?

When viewing any of the displayed maps n the **GIS Map** tab (lower right frame of the main window), click on the GIS toolbar icon for **Save Shapefile** (looks like a 3.5-inch diskette). Follow the prompts to provide a name and location for the file. BenMAP-CE will export a set of files (.dbf, .prj, .shp, .shx) associated with the shapefile that you can use with any GIS viewer. To export the map as an image, click the **Export map image** icon (immediately below the **Save Shapefile** icon). This will use built-in DotSpatial GIS tools to allow you to save the map as a formatted image (.png) file.

For the Rollback to a Standard option, why are there Interday and Intraday rollback options?

The **Interday Rollback Method** option identifies the approach (e.g., *Percentage*) to reduce daily air pollution levels, in order to meet the specified standard. (In other words, there is more than one way to reduce daily pollution levels so as to meet the standard you have chosen, and BenMAP-CE lets you choose from among those approaches.) The **Background** level associated with the **Interday Rollback Method** specifies the bound, below which, BenMAP-CE will not make adjustments to daily levels.
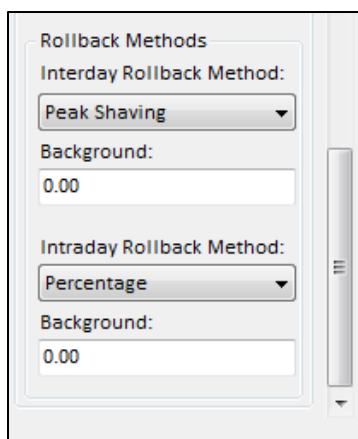
The **Intraday Rollback Method** option is only relevant for hourly air pollution data, like ozone measurements. This option specifies the approach (e.g., *Percentage*) used by BenMAP-CE to reduce hourly air pollution levels to reach the target metric values. That is, once you have chosen the approach to reduce daily air pollution levels, on any given day there is more than one way to reduce the hourly air pollution values to meet the targeted pollution level for that day. The **Background** level associated with the **Intraday Rollback Method** specifies the bound, below which, BenMAP-CE will not make adjustments to hourly levels.

The **Interday** and **Intraday** options are complicated. Appendix A on Monitor Rollback Algorithms explains these options in more detail and gives some numerical examples.

Can I use air quality grids based on different Grid Types in the baseline and control scenarios?

No. In any given analysis, you need to use the same Grid Type in the baseline and control scenarios.

Can I use air quality grids of the same Grid Type but based on different Grid Creation Methods?

Yes. In any given analysis, you may use air quality grids made with different methods. Air quality grids made with Model Data and Monitor Data may be used interchangeably, if desired. Similarly, air quality grids made with different interpolation methods may be compared. However, it generally is not recommended to create grids with different methods and use them in the same analysis.

Can I do an analysis with multiple pollutants?

You can currently only estimate impacts one pollutant at a time; however, BenMAP-CE allows you to aggregate the results of more than one pollutant. This is discussed in Chapter 7: Aggregating, Pooling, and Valuing.

Why does it take so long to generate an Ozone Air Quality grid if there are a lot of grid cells?

It can take a long time to create an air quality grid because the file being generated can be quite large. In some cases, air quality grids can be several hundred megabytes in size. (One reason the ozone files are large is that the definition of ozone has, by default, four metrics. If you do not need all of the default metrics for the health impact functions in your database, then delete the unneeded metrics and BenMAP-CE will run faster and generate smaller air quality grids. This is an advanced step, so do not do it if you are unsure.) The type of computer you use can also affect processing speeds. Refer to Chapter 1, Section 1.3 for recommended hardware specifications.

How do I access data in an Air Quality Surface?

You can access the data in an air quality grid by going to the **Tools** menu and choosing the **Export Air Quality Surface** option. (The **Tools** menu is available on the toolbar of the main BenMAP-CE window.) Locate the air quality grid from which you want to export air quality data and then give a name to your exported file. BenMAP-CE will generate a text file that you can then examine. This is discussed in detail in Chapter 3, Section 3.2.1 (Tools).

How do I perform a rollback to simulate attainment with an annual *and* daily PM₂.₅ standard?

Unfortunately, BenMAP-CE will roll back to either an annual standard or a daily standard—but not both. If this feature is of interest to you, please contact the BenMAP-CE developers at <u>benmap@epa.gov</u>.

Chapter 6
Estimating Incidence

In this chapter...

- **Get an overview of how BenMAP-CE estimates the incidence of health outcomes.**

- **Learn to create a health impact configuration.**

- **Learn about baseline and control scenarios.**

- **Learn the difference between Point Mode and the Monte Carlo analysis options.**

- **Learn how to run, save, and re-open a configuration.**

- **Learn how to view and export Incidence Results.**

Chapter 6 Table of Contents

To estimate changes in the incidence of adverse health effects from air pollution, you will need to create and run a BenMAP-CE configuration file (*.cfgx). A configuration is a reusable file that specifies the air quality grids, health impact functions, population data, and other parameters necessary for an analysis. It is a record of the choices you make in estimating the change in adverse health effects between a baseline and control scenario. The choices include the following:

- The pollutant and air quality grids for the baseline and control scenarios;

- The year for the analysis;

- The population dataset for the analysis;

- The health impact functions to be used in estimating adverse health effects; and

Whether the analysis will focus on a single "point" estimate (Point Mode), or a range of results that mirror the variability in the inputs to the health impact functions (Monte Carlo-generated percentiles).BenMAP-CE gives you flexibility in creating, editing, and saving configuration files. You can open an already existing configuration and proceed directly to estimate incidence. Or, you can create a new file, and then estimate incidence. In addition, you may save any edits made to existing configuration files. After calculating the change in adverse health effects, BenMAP-CE saves the results in a "configuration results" file with a *.cfgrx* extension. The results obtained from running a configuration are sometimes referred to as "raw" results because they represent the estimated change in incidence for each air quality grid cell in a given scenario; they have not been aggregated, pooled, or valued (see Chapter 7).

To load a previously saved configuration, click **Estimate Health Impacts**, which is beside **Step 2** in the BenMAP-CE tree menu (on the left side of the main window). An **Open Existing Configuration** window will appear, as shown below:
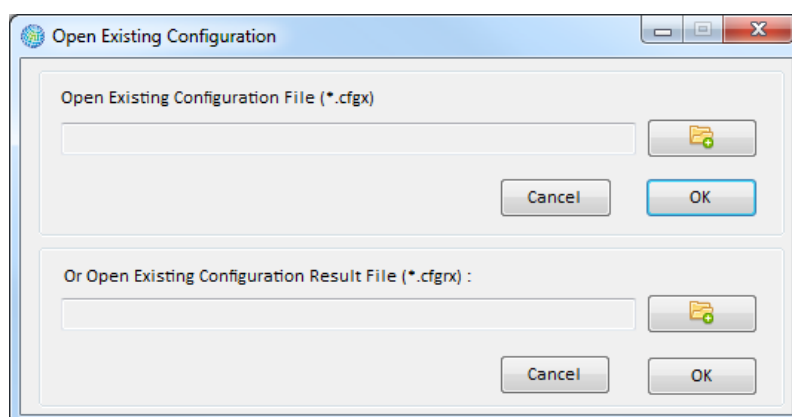
To open an existing Configuration file, click the top file browser button. To open an existing Configuration Result file, click the bottom file browser button. After selecting either of these options, search and select a file to load in the **Open** window. The file

path and name should appear in the box next to the file browser button. Once you have selected the desired option, and are satisfied with the file selection, click **OK**.

6.1   Introduction to Estimating Health Incidence

Health impact functions relate a change in the concentration of a pollutant to a change in the incidence of a health endpoint (i.e., premature mortality or work-loss days). It is typically derived from the estimated relationship between the concentration of a pollutant and the adverse health effects suffered by a given population in an epidemiology study. For example, the pollutant concentration being measured may be particulate matter ($PM_{2.5}$), and the population response may be daily premature deaths. For the purposes of estimating benefits, we are not interested in the health impact function itself, but rather the relationship between the change in concentration of the pollutant, and the corresponding change in the population-health response. We may want to know, for example, if the concentration of $PM_{2.5}$ is reduced by $10\mu g/m^3$, how many premature deaths will be avoided?

To estimate changes in health incidence, the first step is to calculate the change in pollution concentrations for a particular policy scenario, such as an air quality improvement produced by a set of emissions controls. The concentration change in a pollutant is the increment between the baseline scenario and the control scenario. This increment and a gridded population dataset are then used in health impact functions to calculate the change in health incidence that would result from this change in pollution. These functions are based on epidemiological studies and can be selected by the user. Typically, these health incidence results show the decrease in health incidence (e.g., the decrease in asthma, bronchitis, mortality) due to a decrease in pollution. In BenMAP-CE, the selected health impact functions are stored in configurations, which can be re-used over and over again.

Steps to Calculating Health Impacts

$$\Delta\,Y = Yo\,(1 - e^{-\beta\Delta\,PM})\,*Pop$$
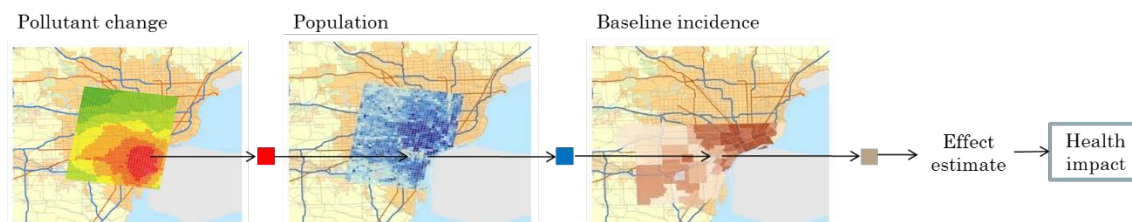
6.2   Create a Health Impact Configuration

There are three major steps to creating a new configuration. First, select a **Pollutant** and specify the **Source of Air Quality Data** for the **Baseline** and the **Control** layers using the BenMAP-CE tree menu on the left side of the main window. Next, choose the **Population Dataset** and **Health Impact Functions** that you want to use in your analysis. Finally, using the **Advanced** button on the **Health Impact Functions** window, specify whether BenMAP-CE will run in **Point Mode** or perform a Monte-Carlo analysis (default setting). (See **Section 6.2.3 Advanced Configuration Settings** for additional detail.)

Note: As you move through the analysis steps, the BenMAP-CE tree menu will update its "**stoplight**" colors to reflect progress. **Yellow** indicates an operation has not yet started. **Green** indicates that an operation has been successfully completed. **Red** indicates that an operation completed, but you may need to re-run this step.

6.2.1   Air Quality Surfaces – Pollutants and Air Quality Grids

Using the BenMAP-CE tree menu, specify the pollutant(s) of interest. Click the **Pollutant** tree menu item to open the **Pollutant Definition** window. Select the pollutant(s) of interest by clicking on an item in the **Pollutants** box (left side) and then clicking the **Add** button to move the pollutant(s) to the **Selected Pollutant** box. To deselect a pollutant and remove it from the **Selected Pollutant** box, click on the item in the **Selected Pollutant** box and click the **Remove** button. You can also add a pollutant by dragging it from the left-hand box to the right-hand box.
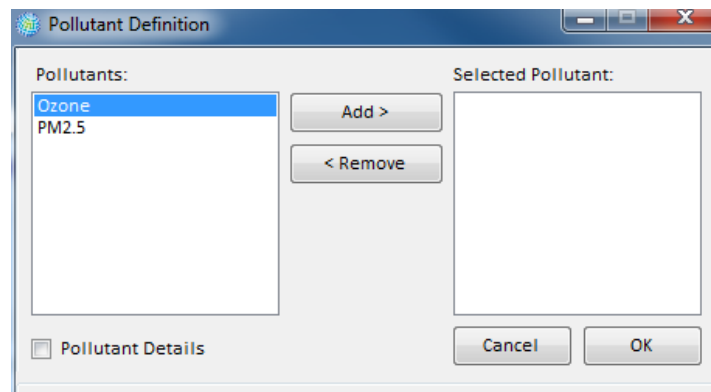
If you want to view or modify any pollutant details, click the box next to *Pollutant Details* and the window will expand to display details and options for the highlighted **Pollutant** (left box). Note that the details are based on the defined **Pollutant Definitions** (see Chapter 4: Loading Data).
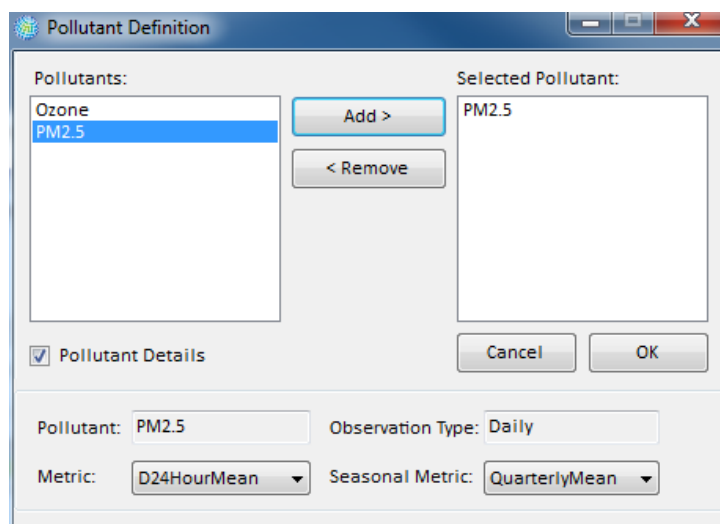
Once your pollutant is selected, click **OK** to close the form. The stoplight for **Pollutant** will change from yellow to green when the operation is successfully completed.

Next, select air quality grids for the **Baseline** file and **Control** file. You may choose existing air quality grids by double-clicking the **Source of Air Quality Data** from the tree menu. The **Open Existing AQ Data** window will be displayed. Select air quality grids, designated with an '.aqgx' extension, for **Baseline** and **Control**. To create a new air quality grid, follow the steps outlined in Chapter 5: Creating Air Quality Grids.
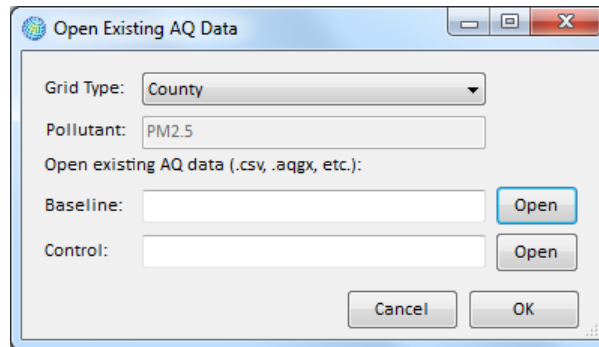
The **Baseline** file contains the air quality metrics for the scenario assumed to occur without any change in policy. The **Control** file specifies the air quality metrics assuming that some type of policy or change has been implemented. The air quality grids should be for the same pollutant, and should also be based on the same **Grid Type**.

If you choose a particular **Grid Type** (e.g., *County*) for the **Baseline** file, then the same grid type must be used in the **Control** file. Conversely, it would not be possible to use *County* grid-type in the **Baseline** and a *Tract* grid-type in the **Control** file.

The **Pollutant** specified in the air quality grids determines the suite of **Health Impact Functions** available for the configuration. Only functions associated with the specified **Pollutant** will be available for the configuration. Furthermore, if only certain **Metrics** associated with the pollutant are present in one (or both) of the air quality grids (see the information in Chapter 3 on monitor and model data formats for more information on how this can occur) only those **Health Impact Functions** associated with those **Metrics** will be available.

Once the **Baseline** and **Control** files are selected, click **OK** to close the form. The stoplights for each of these under **Source of Air Quality Data** will change from yellow to green when the operation is successfully completed.

6.2.2   Estimate Health Impacts – Population and Health Impact Functions

The second step in creating a health impact configuration is to select the **Population Dataset**, **Population Year** and **Health Impact Functions**. If you want to open an existing configuration, double-click the **Estimate Health Impacts** tree menu item and select the file to open. Follow instructions from **Section 3.1.2 Estimate Health Impacts**.

To continue creating a new configuration, double-click **Population Dataset** in the tree menu. Here, you can choose the **Population Dataset** and **Population Year** that will be

used in the analysis. The values in the drop-down list for **Population Year** depend on the range of values in your **Population Dataset**. (See Chapter 4: Loading Data.)
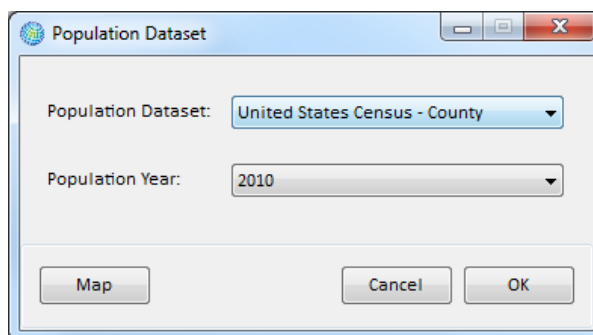
Clicking the **Map** option in the bottom left hand corner will open a separate **Population Map** window. (Warning: this procedure can require a significant amount of time to complete as it creates the population grid and loads the population growth weights). This provides a map of your selected population. If you want to view a different population subgroup based on race, gender, ethnicity, or age range, choose from the available drop-down list (this runs fairly quickly as it uses the same population grid already loaded). The map will refresh automatically.

To change the **Population Dataset** or **Population Year**, click on their respective drop-down list and click **Draw** to update the map and display new results. Warning: This requires BenMAP-CE to recalculate the grid.

Click the close ('×') button at the top of the **Population Map** window to close it.

Clicking **OK** on the **Population Dataset** window will close the window and change the stoplight color to green in the tree menu.

Next, double-click **Health Impact Functions** on the tree menu to display the **Health Impact Functions** window.
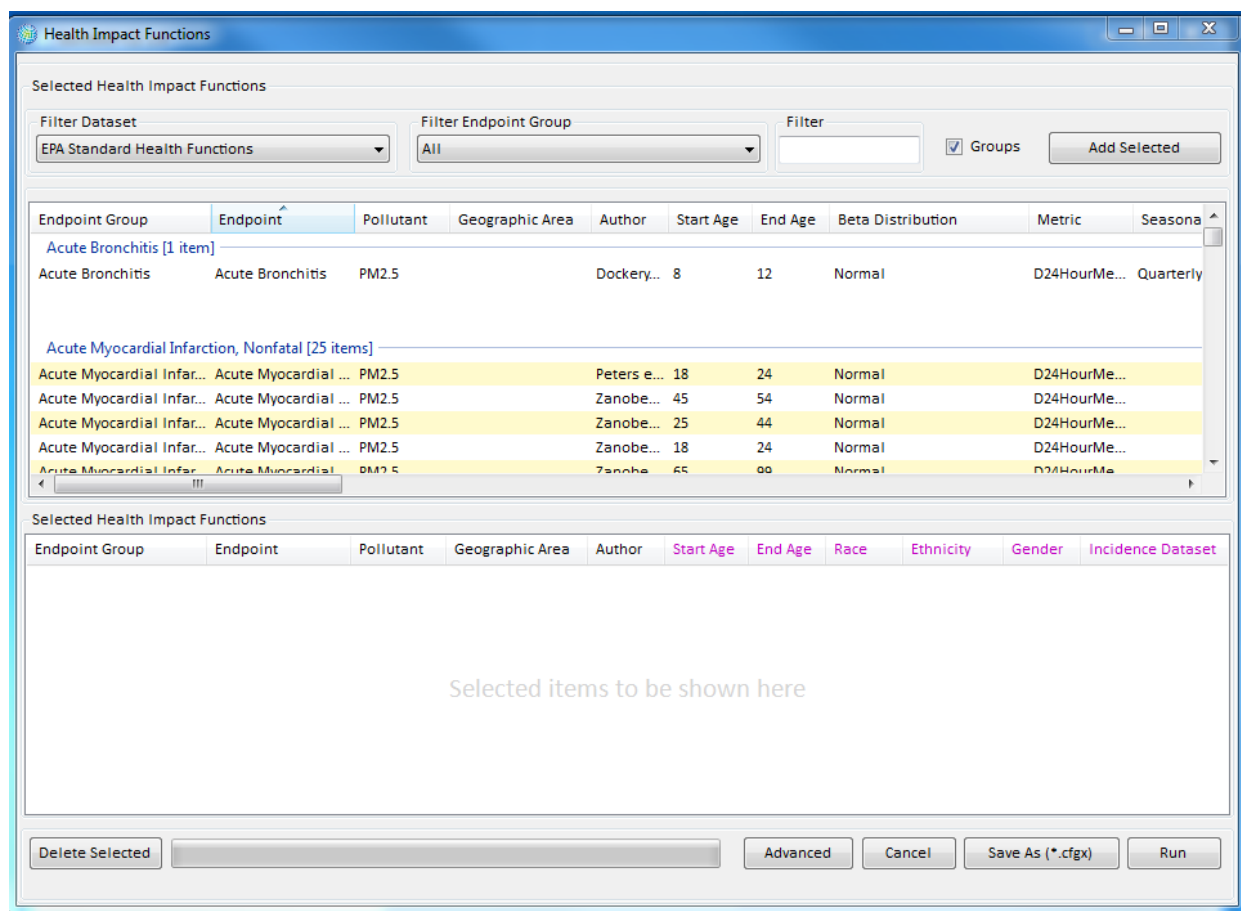
The **Health Impact Functions** window is split into two display frames. The upper frame presents the **Available Health Impact Functions**, which you may select, and the lower frame shows the **Selected Health Impact Functions** that you have already selected. Both frames have a tree structure and the ability to change the order of the fields for easy viewing of the functions. Above the **Available Health Impact Functions** frame, there are a series of buttons that give you the option to select different datasets and filter options for each one.

To add studies to your configuration, simply click to select the health impact functions of interest in the top frame and drag them down to the lower frame. You can do this for blocks of health impact functions by selecting the *Groups* option, clicking on the header of an endpoint group, and dragging the entire endpoint group into the lower frame. And, you can use the options to **Filter Dataset**, **Filter Endpoint Group**, and **Filter** (by keyword) to filter the list according to your preference. Once you are satisfied with the filter, click the **Add Selected** button to apply the selection to the list of selected functions.

If you want to delete some of the health impact functions that you added to your configuration, just highlight the studies of interest and hit the **Delete** key on your keyboard (or click the **Delete Selected** button on the form).

BenMAP-CE displays information for selected studies in two broad categories: **Function Identification** (column headings in black text) and **Function Parameters** (column headings in pink text). The **Function Identification** includes information such as the **Endpoint Group**, **Endpoint**, **Metric**, **Location**, and other variables. This identification information is useful when distinguishing between multiple health impact functions. The **Function Parameters** include those variables that you may directly edit: **Race**, **Ethnicity**, **Gender**, **Start Age**, **End Age**, **Geographic Area, Incidence Dataset**, **Prevalence Dataset**, and **Variable Dataset**.

Clicking on a column header will sort the health impact functions according to that variable. For example, clicking the **Start Age** column will sort the functions by youngest to oldest **Start Age**. You can add the same study to the selection multiple times, and then make edits, in order to be able to calculate the impact of changes in these variables of the function. For example, you can perform an analysis using multiple versions of the same function with different age ranges. To edit the default **Start Age** and **End Age**, just highlight the appropriate cell and type in the desired age values. Keep in mind that these ages represent inclusive age bounds, so if you type in '*10'* and '*12'* this will include all children ages ten, eleven, and twelve years old. If you want to apply a single age year (e.g., only children who are eleven years old), then type the same year (e.g., '*11'*) in both

the **Start Age** and the **End Age**. Note that the accuracy of the populations calculated for these age ranges will depend on the specificity of the population data present in your selected population dataset. Use the drop-down lists in the **Incidence Dataset**, **Prevalence Dataset**, and **Variable Dataset** fields.

New in version 1.4, the **Geographic Area** field allows users to assign health impact functions to either the entire area of analysis ("Everywhere"), to a specific subset of that area (defined by a specific grid definition), or to all areas outside of one or more specified geographic areas ("Elsewhere"). To access these options, select the drop-down list under the **Geographic Area** header. Please note that users must specify whether grid definitions can be linked to health impact functions when defining new grid definitions or modifying existing grid definitions.

6.2.3   Advanced Configuration Settings

Clicking the **Advanced** button at the bottom of the **Health Impact Functions** window gives you access to a number of options for tailoring your assessment.

The Point Mode and Monte Carlo options allow you to generate an average incidence estimate, or a range of results that mirror the uncertainty in the inputs to the health

impact functions. By default, BenMAP will run in Monte Carlo mode. If you select the option for *Run in Point Mode*, BenMAP-CE 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.

With the Monte Carlo (default) option, you can generate a number of estimates that mirror the variability in the inputs to the health impact functions. The Monte Carlo option allows you to generate specific percentiles along the estimated incidence distribution. For example, if you specify 20 **Percentiles** (default value), then BenMAP-CE will generate incidence estimates of the 2.5th percentile, 7.5th percentile, and so on, up through the 97.5th percentile. The number of points suggested in the drop down menu for **Percentiles** varies between 10 and 100. The greater the number of chosen points, the greater the amount of time BenMAP-CE will need to process the results. The relationship between the number of points and time to process is essentially linear, so a doubling of the number of points would double the processing time.
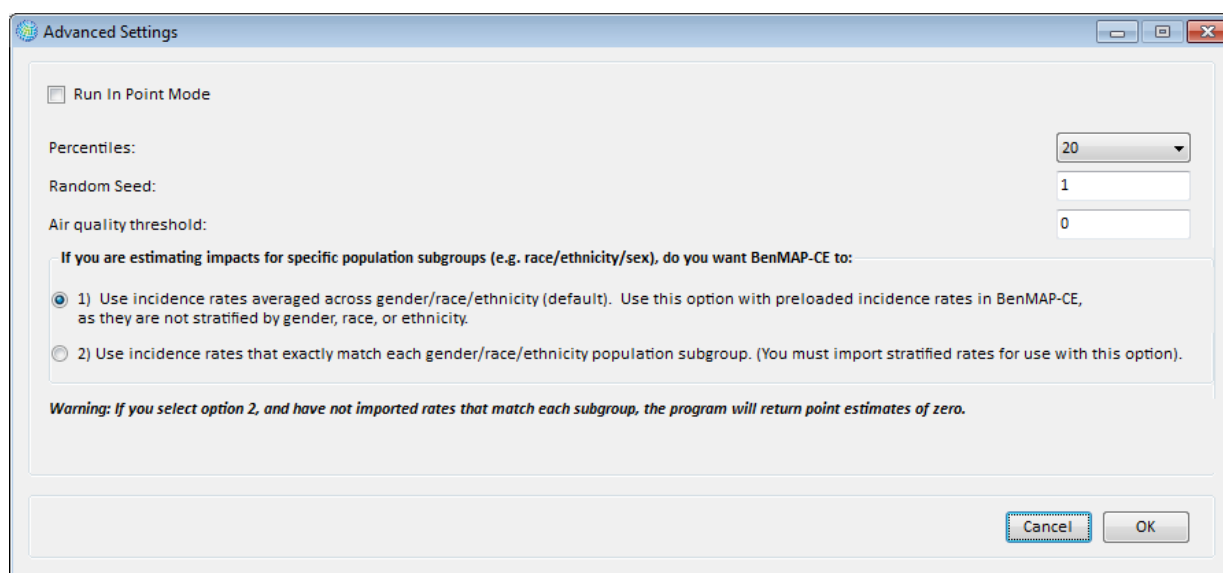
If you choose to *Run in Point Mode*, the field for **Percentile** points for Monte Carlo sampling is disabled and will be ignored (treated as zero). However, with the default Monte Carlo option, the program will still report a point estimate. As discussed in Chapter 7 on Aggregation, Pooling, and Valuation, by choosing the Point Mode, you limit your ability to pool the results. You cannot conduct fixed effect/random effects pooling, or any other procedure that depends on knowing the distribution, or the range of variability of the incidence estimates.

The **Air Quality Threshold** indicates the minimum air quality value that BenMAP-CE will use to quantify health impacts. That is, air quality metrics below the threshold will be replaced with the threshold value. With a threshold of zero, there is no impact on the estimates generated by the health impact functions. However, as the threshold increases, then it will have a progressively larger impact on the incidence estimate. The **Air Quality Threshold** option allows you to explore the impact of any given threshold

on the incidence estimate. This can also be useful for scenarios where you might want to know the incidence associated with changes in air quality occurring only above a standard.

If you are estimating impacts for specific population subgroups (e.g., stratification by gender, age, race, or ethnicity), BenMAP-CE provides two options for calculating baseline incidence:

1) You can use incidence rates averaged over gender, race, and ethnicity (default). (This option <u>must</u> be selected for the rates supplied with BenMAP-CE, as they are not stratified by gender/race/ethnicity).

2) You can use incidence rates that match the gender, race, and ethnicity selected with your health impact function. (You <u>must</u> import stratified rates for use with this option). If you select this option and have not imported rates that match each subgroup, the groups without incidence rates will not be included in the calculations and will have point estimates of zero.

It is recommended that users select this second option if your incidence/prevalence data include overlapping groups. For example, the asthma exacerbation prevalence rate for ages 5-17 has two race groups, one is "ALL" with value 0.107 and the other is "BLACK" with value 0.177**.**

If the first option (above) is selected, BenMAP will use the arithmetic mean of both values (0.142) as the prevalence rate in its calculations, regardless of whether users select "ALL" or "BLACK" as the race category—or left the race category blank.

If the second option is selected, and if users select "ALL" as the race category (or leave it blank), BenMAP will first check if this dataset has overlapping race groups to avoid double counting of individuals in incidence or prevalence calculations. In this example, BenMAP will use the "ALL" prevalence rate of 0.107, which exactly matches the name of the selected race. Thus, it is recommended that users select the second option unless the user is certain that incidence or prevalence rate categories do not overlap.

After making selections for calculating impact functions, BenMAP-CE allows you to save the configuration for future use. Click the **Save As (*.cfgx)** button and specify a file with a *.cfgx* extension.

6.3   Run Health Impact Configuration

To execute the calculation of incidence for the health impact functions in the configuration, click the **Run** button on the bottom right-hand corner of the **Health Impact Functions** window. BenMAP-CE will require that you specify a file in which to save the results, with a *.cfgrx* extension.
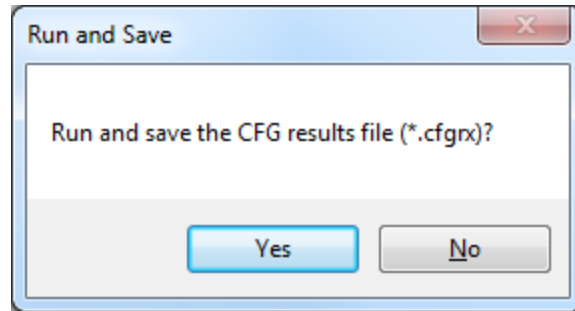
6.4 View and Export Health Impact Results

The **Health Impact Results** report gives you the opportunity to examine the incidence results of each health impact function applied at the grid-cell level, or temporarily aggregate them to, say, the state or national level. The configuration results files (*.cfgrx*) contain "raw" health impact estimates that you have not yet aggregated, pooled, or valued.

To begin, click the **Health Impact Results** tab in the upper portion of the main window. To display incidence results for a single study (i.e., health impact function), double-click on the study of interest (or select the study and click the **Show Results** button). Your results will be generated and displayed in a results table (on the **Data** tab). If you choose a single study, you will also have options to view results on the **GIS Map** tab and in a simple bar chart (view in **Chart** tab.) To change the selected study, double-click on a different study choice or select the study and click the **Show Results** button to refresh the display.
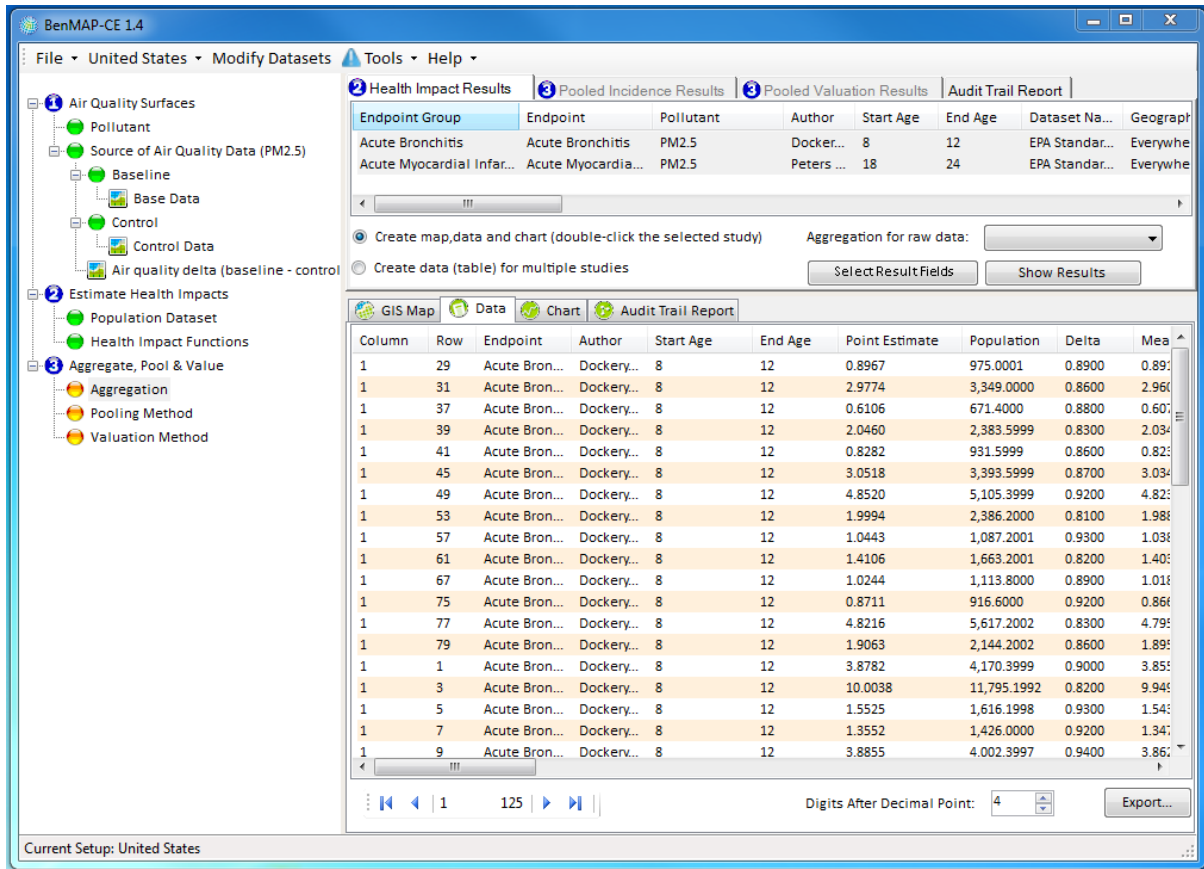
The Health Impact Results tab provides a simple tool to aggregate the raw incidence results. You may select an aggregation level from the **Aggregation for raw data** drop-down list. If you do so, the **Data** table will refresh to display the incidence results aggregated to the selected level.

If you check the option to *Create data (table) for multiple studies,* results will only be available in the **Data** tab. For example, you might want to select three different studies and view aggregated results at the national level. As you modify your choices, the display will be updated accordingly.

On the **Data** tab, clicking the **Select Result Fields** button opens a **Configuration Results Report** window that allows you to choose the columns that will appear in the results table.

- **Grid Fields** permit the inclusion of *Column* and *Row* fields, which can be helpful in identifying the grid-cell of a particular line in the report. For example, when results have been temporarily aggregated to the national level.

- **Health Impact Function Fields** permit the inclusion of various fields which describe or define a function (e.g., *Endpoint Group, Endpoint, Pollutant, Metric,*

Author, Year, Start Age, End Age, Gender, Race, Ethnicity, Beta, Beta Distribution).
These fields can be helpful in identifying the health impact function associated
with a particular line in the report.

▪ **Result Fields** permit the inclusion of fields associated with results of this
analysis (e.g., *Point Estimate*, *Population*, *Delta*, *Mean*, *Percentiles*).
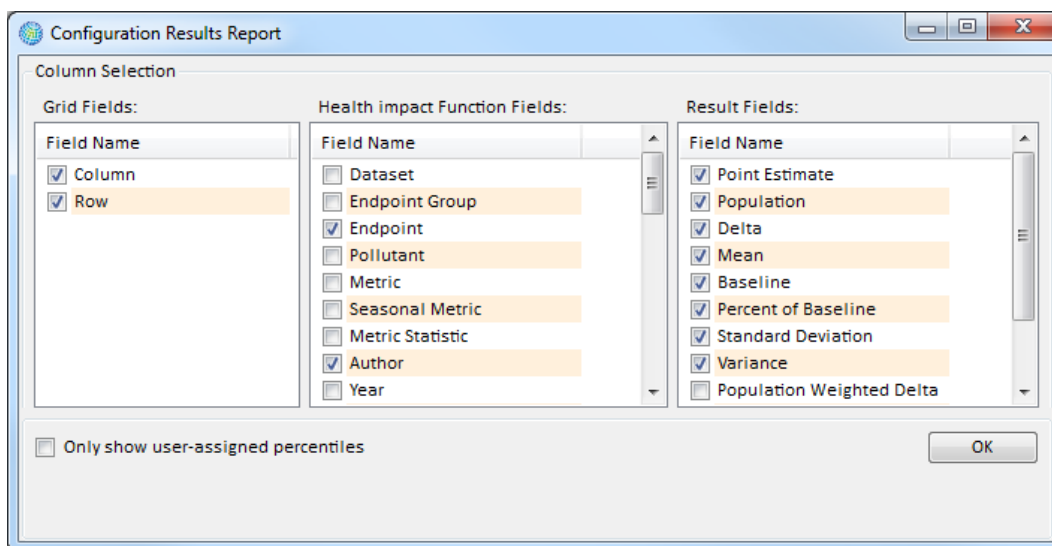
At the bottom of the **Data** tab, there is also an option to specify the number of digits that
appear after the decimal point (click the up or down arrows to edit the number beside
the **Digits After Decimal Point** field, or type a number directly in the box).

Table 6-1 provides a summary of the optional fields that have not been previously
described in this section.

Table 6-1. Selected Variables in the Reports Based on the APVR file

Variable	Description
Column	The column of the grid cell of the result. For grid cell level results, this is the column of the grid cell. For county and state level results, this is the state FIPS code. For national results, this is always 1.
Row	The row of the grid cell of the result. For grid cell level results, this is the row of the grid cell. For U.S. county results, this is the county FIPS code. For U.S. state and national results, this is always 1.
Dataset	Specifies the dataset from which a health impact function was chosen.
Population	Population provides the number of persons used in the health impact function calculation.
Delta	The difference between the baseline and control scenarios for the metric used in the health impact function. Calculated by subtracting the metric value in the control scenario from the metric value in the baseline scenario.

Variable	Description
Point Estimate	The point estimate for the result from the health impact function. The point estimate is generally based on the mean estimate of the "Beta" from the health impact function.
Mean	Mean of the points in the Monte Carlo-generated distribution for this result. The mean is set to missing if the Point Mode option is chosen.
Baseline	Estimate based on the baseline function, which typically estimates health impacts due to all causes (not just air pollution-related causes).
Percent of Baseline	Estimates the percentage change in health impacts (e.g., hospital admissions) due to the change in air quality from the baseline to the control scenario. Calculated by dividing the Point Estimate by the Baseline.
Standard Deviation	Standard deviation calculated based on the points in the Monte Carlo-generated percentiles for this result.
Variance	Variance calculated based on the points in the Monte Carlo-generated percentiles for this result.
Percentiles	The number of percentiles depends on the number of points in Monte Carlo-generated percentiles for this result.

Once all of the options have been selected for your report, you can export the **Health Impact Results** data. First select the **Data** tab then click the **Export** button. This will bring up a window allowing you to name the file you want to save. Note that by default BenMAP-CE will export the file to the CFGR folder. Carefully name the file that you are generating so that you will recognize it in the future!

6.5 Frequently Asked Questions

How do I know which health impact functions to use? Which functions does EPA use?

One option regarding the choice of health impact functions is to work with someone, say another BenMAP-CE user, who is familiar with the epidemiological literature and develop your own set of health impact functions. Reviewing the epidemiological literature can be time consuming, though in some situations, this might be the best option. For example, it would be worthwhile to develop health impact functions to estimate the impacts of carbon monoxide exposure, for which BenMAP-CE does not have pre-installed functions.

Another option is to use the ozone and $PM_{2.5}$ configurations used by EPA. These are available on the BenMAP-CE website (http://www2.epa.gov/benmap/benmap-community-edition). These functions are derived from the epidemiology literature described in the appendices to this user manual. If desired you can edit this configuration and then save it under a different file name—it is always a good idea to keep the original version, so you can go back to it if needed!

How do I edit or add other health impact functions?

To edit or add health impact functions you need to go to **Modify Datasets** window available from the BenMAP-CE main menu. See the health impact function section in Chapter 4: Loading Data for details on how to do this.

How do I learn more about the population data in BenMAP-CE?

Appendix J describes the population data for the *United States* setup in detail.

Why did I not get results for a given geographic area that I wanted in my analysis?

Check to see if your air quality grids mapped properly.

How do I determine what the Column and Row refer to in the result table?

The **Column** and **Row** are variables designed to uniquely identify each grid cell in the grid definition. In the case of the *U.S. County* grid definition, the **Column** refers to the *state FIPS code* and the **Row** refers to the *county FIPS code*. One way to get a good sense of the **Column** and **Row** variables is to create a map (discussed in the next chapter) and then view where particular **Column** and **Row** variables occur in the map.

Chapter 7
Aggregating, Pooling, and Valuing

In this chapter...

- **Get an overview of valuation, discounting, and pooling.**
- **Configure an Aggregation, Pooling, and Valuation (APV) file.**
- **Sort and pool incidence results.**
- **Learn the differences between the pooling methods.**
- **Assign economic values to incidence results.**
- **Aggregate incidence results and valuations.**
- **Run, save, and re-open an APV configuration.**
- **View and export APV results.**

Chapter 7 Table of Contents

This section presents an introduction to valuation, discounting, and pooling. Most BenMAP-CE users find this portion of the program the most complex to understand and use. You may find yourself referring to this chapter frequently.

Once you have created a configuration results file with incidence results based on your two air quality grids (refer to Chapter 6), you can use the Aggregate, Pool, and Value feature to combine the incidence results and place an economic value on the combined results. You have two options.

- **Create a New Configuration to Aggregate, Pool, and Value results.** You can create a new type of configuration, termed an Aggregation, Pooling, and Valuation (APV) Configuration. This allows you to (1) specify the geographic level at which you want to report your results, (2) specify how you might want to combine or "pool" the incidence results, and (3) specify how to assign an economic value to the health incidence results. These selections can be saved in an APV Configuration file (*.apvx*) and used to calculate results, which are stored in an APV Results file (*.apvrx*).

- **Open Existing Configuration for Aggregation, Pooling, and Valuation.** You can load an existing APV Configuration file, edit the configuration, save it with the same or a different name, and then proceed to calculating the results.

7.1   Introduction to Valuation, Discounting, and Pooling

Valuation generally refers to placing a monetary value on estimated health incidence. In the example below, we discuss U.S. dollar values and provide a brief introduction to discounting, which has to do with placing less weight on things occurring in the future than on things occurring today. Finally, we discuss pooling, which has to do with combining comparable results.

7.1.1   Overview of Economic Valuation

Improvements in ambient air quality generally lower the risk of developing an adverse health effect by a fairly small amount across a large population. A lower risk for everyone means that fewer cases of the adverse health effect are expected, although we cannot predict which people would be spared. Therefore, the health benefits conferred on individuals by a reduction in pollution are actually reductions in the risk of having to endure certain health problems. Monetizing the benefits of a reduction in air pollution involves estimating society's willingness to pay (WTP) for these reductions in risk, or the observed Cost of Illness (COI) for an effect, and is typically referred to as valuation. BenMAP-CE uses valuation functions to estimate the monetized benefits of reducing air pollution.

These benefits (reductions in risk) may vary across the population (and could be zero for some individuals). Likewise, the WTP for a given benefit is likely to vary from one individual to another. In theory, the total social value associated with the decrease in risk of a given health problem resulting from a given reduction in pollution

concentrations is generally taken to be the sum of everyone's WTP for the benefits they receive.

7.1.1.1   Monetizing Benefits

Epidemiological studies allow us to estimate the number of cases of an adverse health effect that would be avoided by a given reduction in pollutant concentrations. If we have an estimate of the average individual's WTP for the risk reduction conferred upon him, we can derive from that an estimate of the value of a statistical case avoided. Suppose, for example, that a given reduction in pollutant concentrations results in a decrease in mortality risk of 1/10,000. Then for every 10,000 individuals, one individual would be expected to die in the absence of the reduction in pollutant concentrations (who would not be expected to die in the presence of the reduction in pollutant concentrations). If the average individual's WTP for this 1/10,000 decrease in mortality risk is $100, then the value of a statistical life is 10,000 × $100, or $1 million. In general, the ex-ante WTP for a risk reduction of x can be converted into an ex-post value of a statistical case avoided by dividing the average individual's WTP for the risk reduction of x by x (e.g. $100/0.0001 = $1,000,000). The same type of calculation can produce values for statistical incidences of other health endpoints.

Sometimes those economic values come from contingent valuation studies, in which study participants are queried about their WTP to avoid a specific adverse health effect. When estimates of WTP are not available, it can be approximated by other measures, most notably COI measures.

An individual's WTP to avoid an adverse health effect will include, at a minimum, the amount of money he or she would have to pay for medical expenses associated with the illness. Because medical expenditures are to a significant extent shared by society, via medical insurance, Medicare, etc., the medical expenditures actually incurred by the individual are likely to be less than the total medical cost to society. The total value to society of an individual's avoidance of an adverse health effect, then, might be thought of as having two components: (1) the COI to society, including the total value of the medical resources used (some portion of which will be paid by the individual), 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.

The COI approach attempts to estimate the total value of the medical resources used up as well as the value of the individual's time lost as a result of the illness. Because this method does not include the value of avoiding the pain and suffering resulting from the illness (a potentially large component), it is generally believed to underestimate the total economic value of avoiding the illness, perhaps substantially.

The contingent valuation method (and conjoint analysis) attempts to elicit from people what they would be willing to pay (WTP) to avoid the illness. Because of the distortion in the market for medical goods and services, whereby individuals generally do not pay

the full value of the medical care, this method too is likely to understate the total economic value of avoiding the illness.

Although the COI and WTP are the two most common methods, other methods have been used in certain circumstances. The method the benefit analyst chooses to value a particular health endpoint will depend in part on what data are available. The unit values currently available for use in BenMAP-CE are data or estimates that have been collected or generated by researchers and can be readily obtained in publicly available databases or in the open literature. When reviewing the economic literature to determine the appropriate valuation functions to use, it is important to have an economist assist.

7.1.1.2   Valuing Reductions in Premature Mortality

The economics literature discussing the value of changes in fatality risks is extensive and provides a basis for monetizing benefits when the number of deaths avoided as a result of an air quality improvement can be calculated, but the literature on certain issues regarding the appropriate method for valuing reductions in premature mortality risk is still developing. 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-CE currently offers a variety of options reflecting the uncertainty surrounding the unit value for premature mortality. See the Appendices H and I for more detail on the valuation functions available in BenMAP-CE.

Monetary estimates of changes in premature mortality risk are often expressed in terms of the Value of a Statistical Life (VSL). This term is easily misinterpreted and should be carefully described when used in benefit analysis. VSL is the aggregate dollar amount that a large group of people would be willing to pay for a reduction in their individual risks of dying in a year, such that we would expect one fewer death among the group during that year on average. The basic assumption underlying the VSL approach is that equal increments in fatality risks are valued equally. For similar reasons, the VSL approach is only appropriate for marginal changes in the risk of death and should not be used to value more significant changes. Because changes in individual fatality risks resulting from environmental regulation are typically very small, the VSL approach is usually acceptable for these types of benefit analyses.

The U.S. EPA National Center for Environmental Economics provides answers to frequently asked questions regarding the economic value of mortality risk on its website: https://www.epa.gov/environmental-economics/mortality-risk-valuation. You may wish to consult this site as you have questions regarding how U.S. EPA derives VSL and applies it in an environmental benefits analysis.

7.1.2 Overview of Discounting

What is discounting?

In general, people prefer current consumption to future consumption. In other words, a $1 today is worth more today than a $1 tomorrow is worth today, and that dollar continues to decrease in value as you go further out into the future. (This concept is also referred to as the social rate of time preference or the time value of money. This is a different concept than inflation, which is a general increase in the price level of goods and services.) Discounting is the process of converting that future dollar into a value that can be compared to the value of a dollar today. The discount rate expresses this process in quantitative terms. The higher the discount rate, the faster value decreases over time. For example, $1 twenty years from now is worth $0.55 today at a 3% annual discount rate, but worth only $0.26 at a 7% annual discount rate.

A basic discounting function is as follows:

Present Value = Future Value / $(1+r)^t$

where r is the discount rate and t is the time period (usually years).

Example: $1 twenty years from now at a 3% annual discount rate is worth $0.55 today

Present Value = $1.00 / $(1 + 0.03)^{20}$ = 1 / $(1.03)^{20}$ = 1 / 1.806111 = 0.553676 = $0.55

Why do we discount benefits?

The benefits of reductions in air pollution may need to be discounted for three key reasons:

1. Today's society values benefits that occur today more highly than benefits that will occur in the future. Therefore, we must discount in order to compare those future benefits with current benefits.

2. For a cost-benefit analysis, benefits estimates in a future year need to be comparable to the cost estimates for that same year (which are also discounted). 3. Discounting can be used to compare the future *streams* of benefits and costs. The core BenMAP-CE program estimates changes in adverse health effects based on changes in air quality for one specified analysis year, even though certain health benefits may occur after the analysis year.[1] Discounting can be used to compare the future benefits with benefits occurring during the analysis year.

---

[1] The PopSim tool estimates the change in population mortality risk over a multi-year period, but it is not yet possible to estimate the economic value of these impacts in BenMAP-CE.

Under which scenarios would I need to discount benefits?

Health benefits may occur three different ways after the analysis year specified in BenMAP-CE.

1. Certain health endpoints accrue medical expenses or lost earnings for multiple years. The future medical expenses would need to be discounted to compare with expenses occurring in the analysis year.

2. Pollution exposure and the resulting health effects do not occur within the same year (a.k.a. a cession lag). The monetized benefits of future health effects would need to be discounted to compare with the benefits of health effects that occur during the analysis year.

3. In some analyses, you may want to estimate a stream of benefits occurring over multiple analysis years instead of just one analysis year. In this scenario, you would need to discount the future benefits occurring in each year analyzed back to the present year in order to present the cumulative total estimate of benefits (i.e., the net present value of a stream of benefits).

When would we not discount benefits?

In many instances, it is not necessary to discount the benefits estimates generated by BenMAP-CE. If the health effect and the monetized value of all the medical expenses, lost earnings, and suffering occur entirely in the analysis year, then you may not need to discount your benefits. For example, school loss days occur within the analysis year, and all monetized expenses occur within the analysis year. It is important that you understand the assumptions within the health and valuation functions before you decide whether you need to discount. (If your analysis year for your benefits estimates does not match the analysis year for your costs estimates, you may need to discount in order to compare your benefits with your costs even if you meet the criteria listed above.)

Which discount rate should I choose?

Selecting a discount rate is challenging and is one of the most contentious methodological issues encountered in economic analyses of environmental policies. Because environmental regulations frequently have differing streams of costs and benefits over time, the selected discount rate may determine whether the benefits of a regulatory action exceed the costs. In addition, selecting a higher discount rate may result in a smaller benefits estimate because the future benefits are worth much less than they would be if a lower discount rate was selected. For benefits that occur well into the future, the issue of intergenerational equity further complicates the selection of the discount rate. (In the context of environmental policy, intergenerational equity refers to the fairness of the distribution of the costs and benefits of a long-lived policy when those costs and benefits are borne by different generations. Most criteria

pollutants are not considered to have intergenerational equity issues, but the issue frequently arises in analyses of climate and mercury.)

There are various economic arguments in support of and in opposition to various discount rates. To comply with OMB and EPA's recommendations, EPA currently uses discount rates of 3% and 7% for benefit analyses. For more details, see EPA [1999; 2000] listed in Chapter 1, Section 1.7 (Sources for More Information).

Which health endpoints accrue medical expenses or lost earnings for multiple years, and how do I discount them?

BenMAP-CE includes health and valuation functions for several chronic health effects, including $PM_{2.5}$-related chronic bronchitis and non-fatal acute myocardial infarctions (AMIs, or heart attacks).

- Chronic bronchitis is assumed to last from the initial onset of the illness throughout the rest of the individual's life. BenMAP-CE currently includes one WTP function as well as two COI functions representing the two discount rates for chronic bronchitis.

- Technically, AMIs are discrete, acute events, not chronic conditions. However, heart attacks cause chronic follow-up health effects that accrue medical expenses over time, similar to chronic conditions. You can discount the economic value of these chronic effects through the valuation function in BenMAP-CE. AMIs are assumed to accrue costs over five years. Although WTP functions for AMIs are not available, BenMAP-CE currently includes several COI functions that incorporate the direct medical costs and the opportunity cost (lost earnings) for specific age groups at two discount rates.

See Appendix G for details on the discounting assumptions within the valuation functions.

Should I discount the health incidence as well as the valuation?

You should not discount the health incidence for any of the scenarios mentioned above. Changes in the lag assumptions do not change the total number of estimated deaths, for example, but rather the timing of those deaths. If you discounted the health incidence along with valuation, you would essentially be discounting twice.

Which health endpoints do not occur in the same year as exposure?

In many cases, the health effect from exposure to air pollution occurs shortly after exposure, but there can be a significant lag between exposure and the health effect. The cession lag can be a matter of hours or days, but some health effects may lag exposure by much longer. If exposure and the health effect do not occur within the same year, it is necessary to discount those benefits back to the analysis year. The only health function currently in BenMAP-CE that falls into this category is $PM_{2.5}$-related premature

mortality. Discounting PM-related premature mortality is controversial because the lag structure is unknown, but scientific literature on similar adverse health effects and new intervention studies suggest that premature mortality probably would not occur in the same year as the exposure. (See: Roosli M, Kunzli N, Braun-Fahrlander C, Egger M. 2005. "Years of life lost attributable to air pollution in Switzerland: dynamic exposure-response model." International Journal of Epidemiology 34[5]:1029-35.)

EPA's Science Advisory Board recommends future research to support the development of defensible lag structures and provided a lag structure that could be assumed until additional research has been completed. See Chapter 5 of the PM Regulatory Impact Analysis for more detail on assumed lag structures for $PM_{2.5}$-related premature mortality (http://www.epa.gov/ttn/ecas/regdata/RIAs/finalria.pdf). Some example lag structures from the PM RIA are shown in Figures 7-1 and 7-2 below. Currently, BenMAP-CE does not have the capability to do this type of discounting, so you must discount outside of BenMAP-CE.

Note: Discounting is not necessary for ozone-related premature mortality because it occurs within the analysis year.
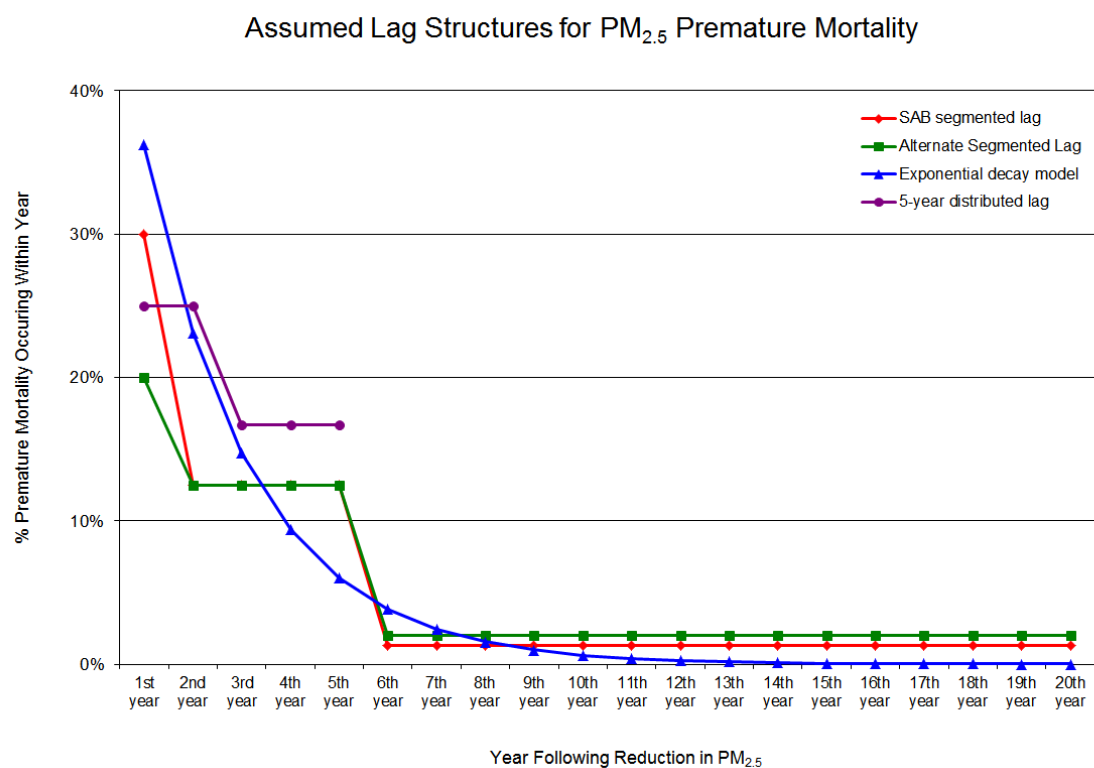
Figure 7-1. Graphical representation of assumed lag structures analyzed in EPA's PM RIA as sensitivity analyses
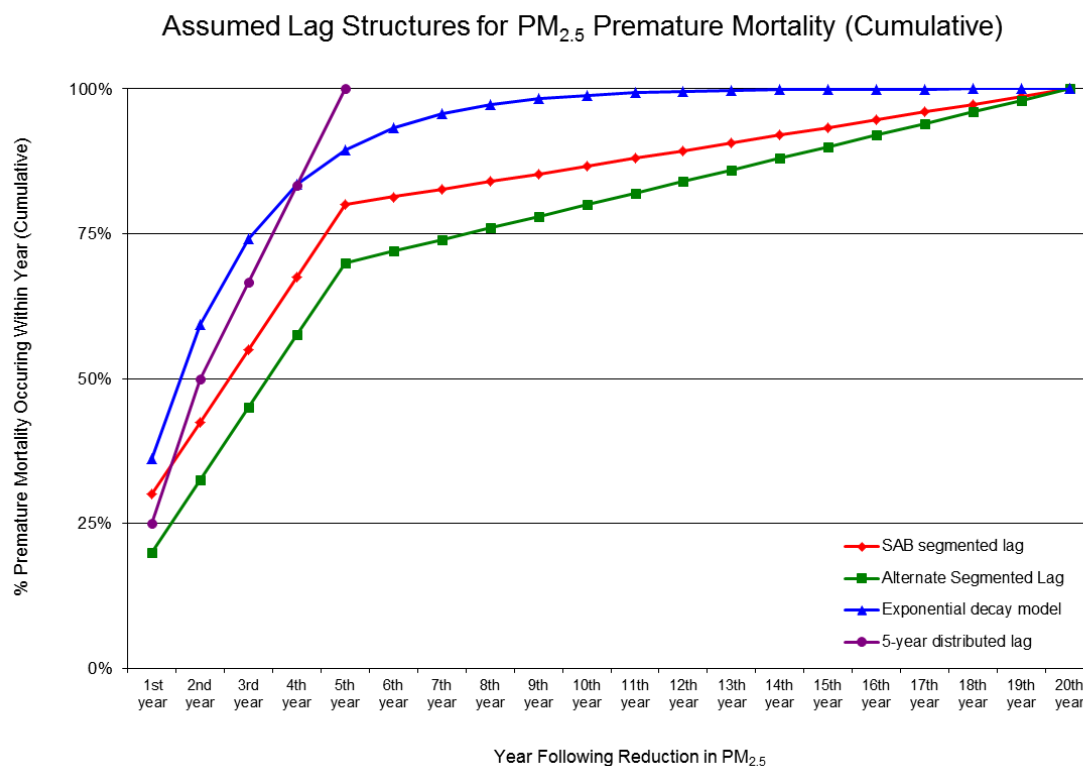
Figure 7-2. Graphical representation of cumulative assumed lag structures analyzed in EPA's PM RIA as sensitivity analyses

7.1.3 Overview of Pooling

For many of the health endpoints (e.g., respiratory hospital admissions), BenMAP-CE contains many different functions from different studies that you could choose to include in your configuration. Combining data from several comparable studies in order to analyze them together is often referred to as meta-analysis. For a number of reasons, it is often impractical or impossible to combine the original data sets. Combining the results of studies provides a second-best way to synthesize information. This is referred to as pooling.

BenMAP-CE allows users to pool the estimated incidence changes predicted by several studies for the same pollutant-health endpoint group combination (e.g., $PM_{2.5}$-related cardiovascular hospital admissions). It also allows the pooling of the corresponding study-specific estimates of monetary benefits.

Why would you want to pool results?

There are two good reasons to pool across study results, one practical and one methodological. Pooling allows you to

- Combine or aggregate multiple study estimates into a single estimate. This combined estimate is easier to report.

- Certain types of pooling—including random effects techniques—account for heterogeneity in the risk estimates reported in the epidemiological literature used to construct the health impact functions you used to calculate incidence.

However, as we discuss below, pooling may not be such a good option if

- You don't know a great deal about the studies used to quantify health impacts; you'll need to know a lot about epidemiological studies used to construct the health impact functions in order to pool properly.

- You think it is important to convey the variability across incidence estimates for a given health endpoint.

BenMAP-CE allows you to pool in five different ways:

- Addition

- Subtraction

- User-assigned weights

- Random Effects

- Fixed Effect.

The examples that follow demonstrate each of these methods (Fixed and Random Effects are combined in one example.) In each of the examples below, the distribution of estimated health impacts are represented by a normally distributed probability density function (PDF) shaped as a bell curve. In each PDF, the mean health impact estimate is represented by a dashed line. While we illustrate these examples using a normally distributed PDF for ease of presentation, BenMAP-CE can accommodate several other types of distributions (e.g. Weibull, Triangular).

Example 1: Pooling by Addition

You might want to use the *Pooling by Addition* option if you would like to aggregate two outcomes that are non-overlapping. In the example below, we have estimated ischemic heart disease hospital admissions and dysrhythmia hospital admissions. You'll see that each endpoint is associated with unique, and non-overlapping, International Classification of Disease 9th edition (ICD-9) code (Slee 1978). Therefore, it's ok to add the two estimates, because doing so would not double-count impacts.

Addition allows us to combine non-overlapping estimates of a common health endpoint

The sum of ischemic heart disease and dysrhythmia is provides a better overall characterization of the effects of air pollution on cardiovascular outcomes than either endpoint alone.
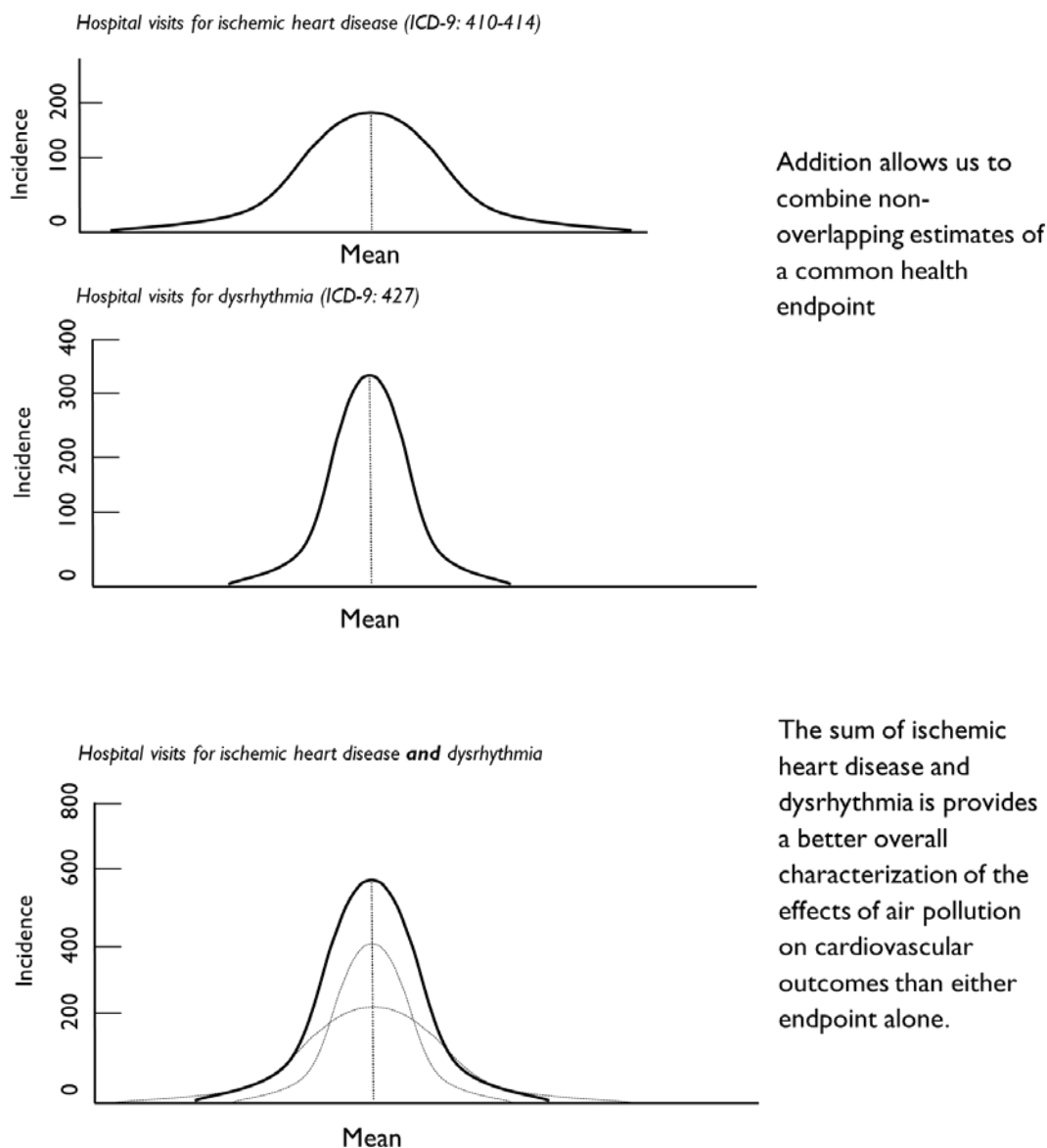
Example 2: Pooling by Subtraction

In this example, suppose that you have estimated the incidence of total cardiovascular hospital admissions using one health impact function. Suppose also that you estimated the incidence of total cardiovascular hospital admissions, *less stroke*. In this instance, you could subtract the second incidence estimate from the first incidence estimate to yield the number of stroke hospital admissions.

Hospital visits all cardiovascular outcomes (ICD-9: 390-459)
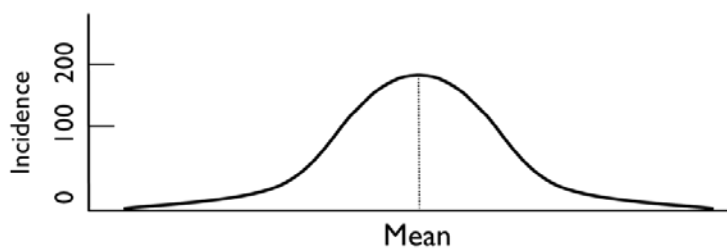
Subtraction allows us to "net out" the incidence of a health endpoint from two or more studies

Hospital visits for all cardiovascular outcomes except stroke (ICD-9 390-440)
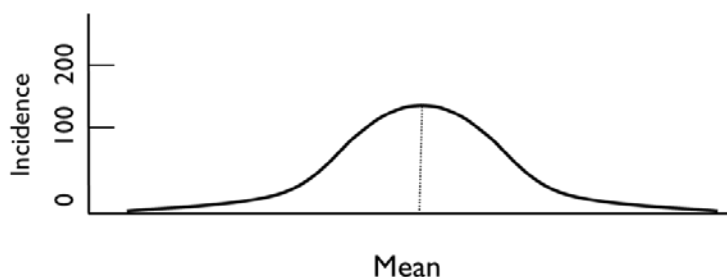
In this example, the only difference between these two studies is that study one includes all cardiovascular outcomes, while study two excludes strokes

Hospital visits for stroke (ICD-9: 440-449)
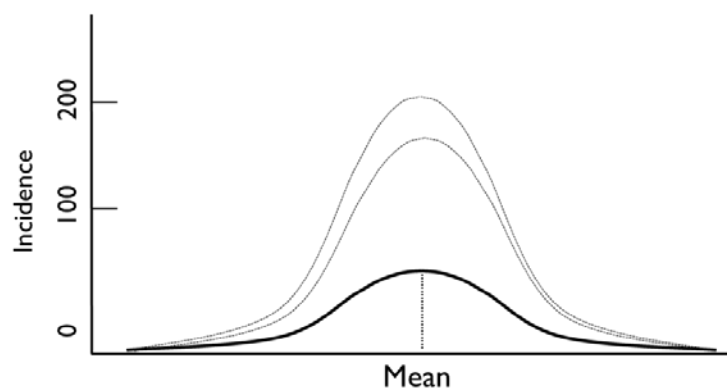
Subtracting the results of study two from study one yields an estimate of stroke

Example 3: Pooling with User-Assigned Weights

In this case, you might have estimated the change in incidence using two different health impact functions for the same health endpoint and would like to combine them using weights that you specify.

Peng et al. 2009 Multi-City Study of Cardio Hospital Admissions

25% weight
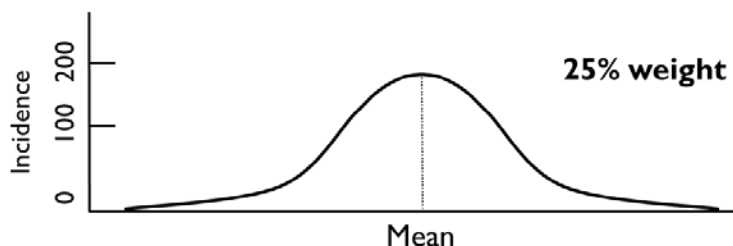
Some studies examine a common health endpoint and share a similar methodology, but may differ slightly in the populations examined

Bell et al. 2008 Multi-City Study of Cardio Hospital Admissions
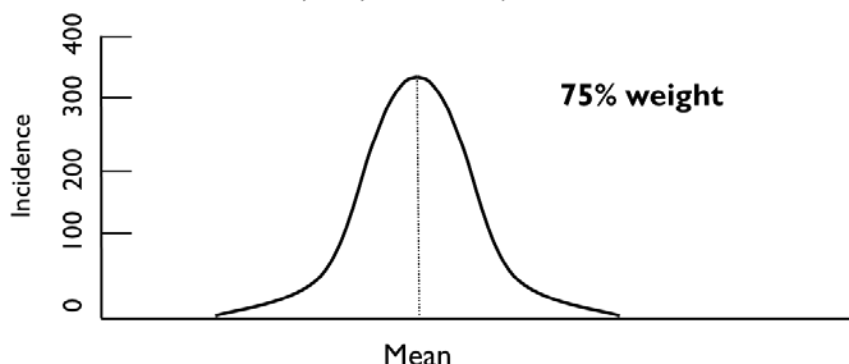
75% weight

Users may wish to combine these study estimates together using equal weights

Pooled estimate of Peng & Bell
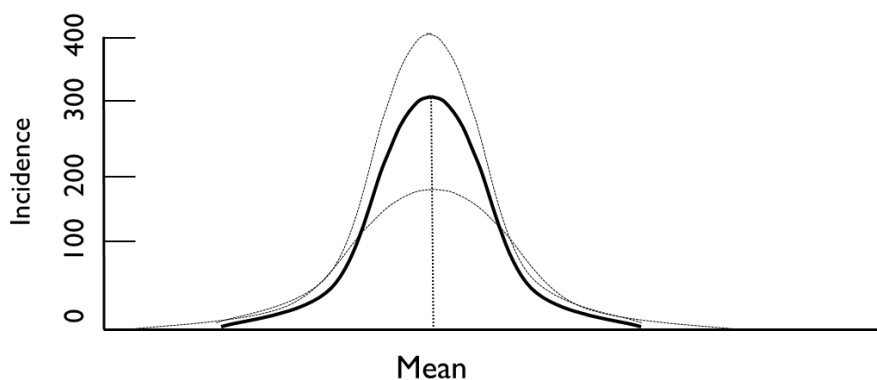
The pooled value reflects a weighted average of the two studies

Example 4: Pooling Using the Fixed Effect and Random Effects

The Fixed Effect and Random Effects pooling techniques are among the most complicated and are best applied only when you understand clearly the assumptions inherent in the method and its suitability to the incidence estimates. The example below describes the procedure for performing this technique.

For the risks of a given health outcome there is a **true** but **unknown** distribution

The individual studies in the literature report individual risk estimates from that distribution

Random effects pooling accounts for heterogeneity in the individual risk estimates to generate a single mean risk estimate

Adapted from: Mosteller and Colditz (1996); Charles Poole EPID 731

The Random-Effects model assigns each study a weight based on two factors:
1. The spread of estimates reported by each study (i.e. the variance)
2. How much that spread of estimates differs from spread reported by the other studies

Adapted from: Mosteller and Colditz (1996); Charles Poole EPID 731

Pooling: Random-Effects
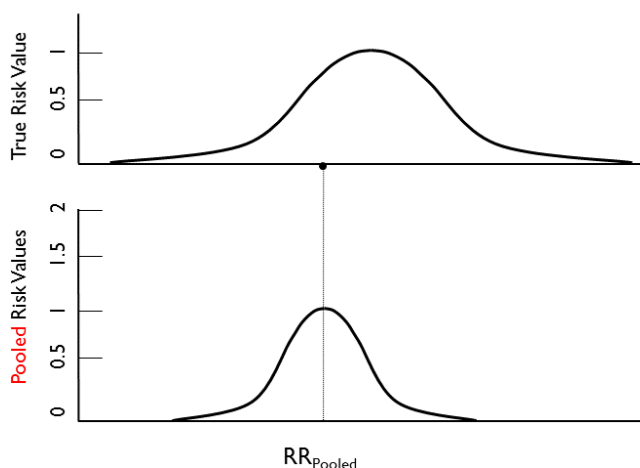
The Random-Effects model assigns each study a weight based on two factors:

1. The spread of estimates reported by that study (i.e. the variance)
2. How much that spread of estimates differs from the other studies

Finally, the Random-Effects model calculates a weighted average of the studies

Table 7-1 summarizes the different types of pooling approaches, and Appendix I provides a detailed discussion of the approaches. Note that some pooling methods are only available in Monte Carlo mode. This is because these pooling methods attempt to combine distributions of results into new distributions, and no distributional information is available in Point Mode. The Pooling Method column will thus have different values in its drop-down list depending on the mode used to generate the incidence results being pooled.

Table 7-1. Pooling Approaches for Incidence and Valuation Results

Pooling Approach	Description of Pooling Approach	Availability	
		Point Mode	**Monte Carlo**
None	No pooling performed.	Yes	Yes
Sum (Dependent)	Results are summed assuming they are perfectly correlated. In Point Mode, this is just a simple sum. In Monte Carlo mode, BenMAP-CE chooses the first point from each result in the pooling and does a simple sum to generate the first point in the pooled result, and so on for all of the points in the distribution of results.	Yes	Yes
Sum (Independent)	Results are summed assuming that they are independent. A Monte Carlo simulation is used. At each iteration, a random point is chosen from the distribution for each result, and the sum of these values is put in a holding container. After some number of iterations, the holding container is sorted low to high and binned down to the appropriate number of percentile points.	No	Yes

Pooling Approach	Description of Pooling Approach	Availability	
		Point Mode	Monte Carlo
Subtraction (Dependent)	Results are subtracted assuming they are perfectly correlated. All subsequent results are subtracted from the first result (the highest result in the display - to reorder results, simply click and hold a result and then drag it to its new position). In Point Mode, this is a simple subtraction. In Monte Carlo mode, BenMAP-CE chooses the first point from each result in the pooling and does a simple subtraction to generate the first point in the pooled result, and so on for all of the points in the distribution of results.	Yes	Yes
Subtraction (Independent)	Results are subtracted assuming that they are independent. A Monte Carlo simulation is used. At each iteration, a random point is chosen from the distribution for the first result, and then random points are chosen from the distribution for each subsequent result and subtracted from the first. The result is put into a holding container. After some number of iterations, the holding container is sorted low to high and binned down to the appropriate number of percentile points.	No	Yes
User defined Weights	Weights are specified by the user. In Point Mode, the new result is generated by a simple weighted sum of the input results. In Monte Carlo mode, the results are combined using the user specified weights with the '*Round Weights to Two Digits*' Advanced Pooling Method. Note that the weights you enter need not add up to one - BenMAP-CE will normalize them internally. Also note that BenMAP-CE initializes all the weights to 1/n, where n is the number of results being pooled.	Yes	Yes

7.2   Create Aggregation, Pooling, and Valuation (APV) Configuration

Once you have run a configuration result file (see Chapter 6), you can begin creating your APV configuration. You will start with selecting the aggregation levels for incidence and valuation results, then move on to pooling and valuation. These processes are described in detail below.

7.2.1 Selecting Aggregation Levels

Double-clicking **Aggregation** from the BenMAP-CE tree menu item opens the window which lets you choose the level of aggregation for the incidence and valuation results.
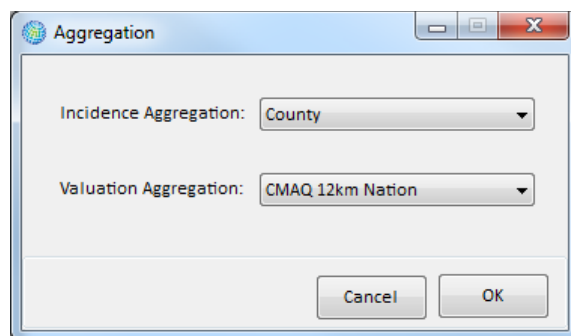
7.2.2   Pooling Incidence Results

To begin pooling incidence results, double-click **Pooling Method** from the BenMAP-CE tree menu. In the top half of the **Select and Pool Incidence Estimates** window, you will find a list of **Available Incidence Results**. The results are represented by the health impact functions from which they were created. You may click on the **Select study fields** button to display additional information about each health impact function. This is recommended when using the geographic areas functionality (i.e., applying region-specific air pollution effect coefficients).
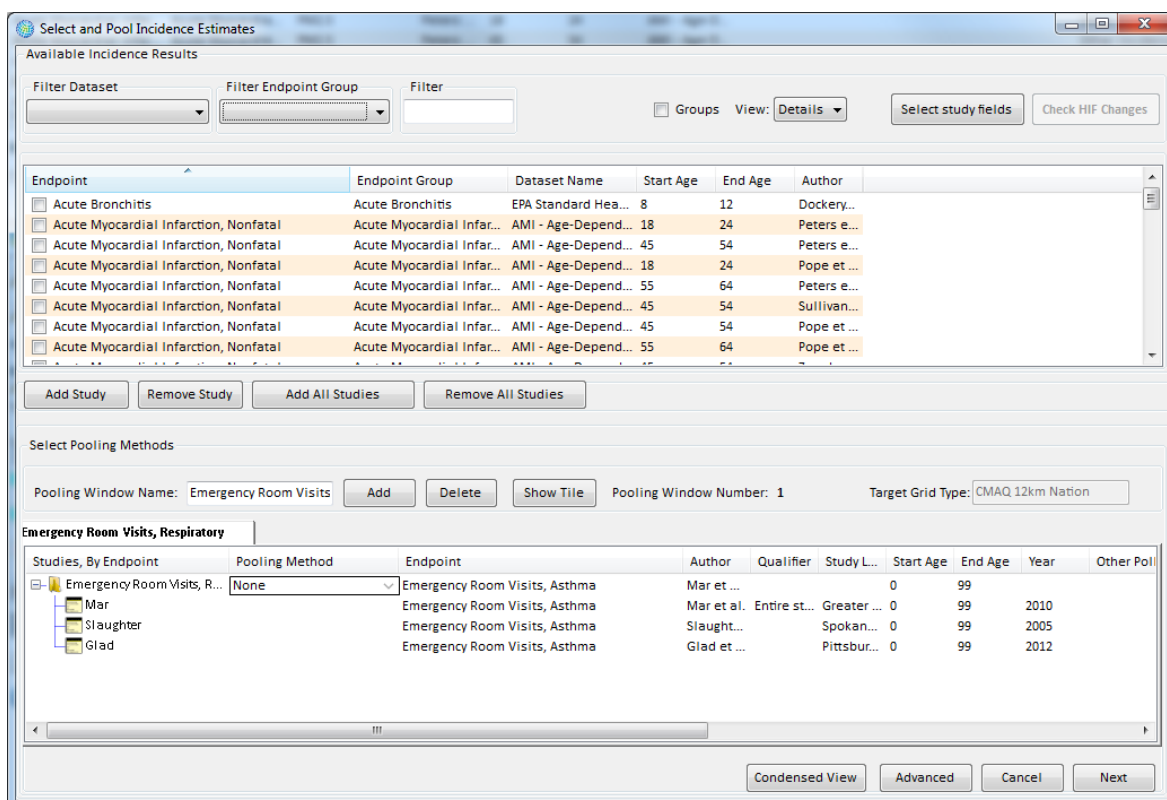
There are several steps to pooling your incidence results:

Step 1. Add incidence results to the pooling window

Using the available incidence results, you can drag individual incidence results to the pooling window, or select the result(s) using the checkboxes next to each study, and click **Add Study** to move them to the pooling window. You do not have to add all of your incidence results to the pooling window, but note that only those results showing in the pooling window will be included in the pooled incidence or valuation results. As with health impact functions, in Chapter 6, there is a filter method above the **Available Incidences Results** for your convenience.

The bottom half of the pooling window consists of one or more tabs, each with an associated "pooling window", where health impact functions can be added and specifications for how to pool them (if any) can be made. In version 1.4, by default, BenMAP-CE will automatically create a pooling window for each Endpoint Group for which one or more Health Impact Functions were selected in the Health Impact Function selection screen, and the Endpoint Group will be indicated on the tab. For example, there could be one pooling window for Mortality and one for Asthma Exacerbation. Health impact functions should be added to the pooling window corresponding to that function's Endpoint Group. Additional pooling windows can be added if needed, as discussed below.

Incidence results are displayed in the pooling window in a tree structure determined by (1) the order of the columns, and (2) the values of the identifying variables of the **Health Impact Functions** from which the incidence results were generated (**Endpoint Group**, **Endpoint**, etc.).
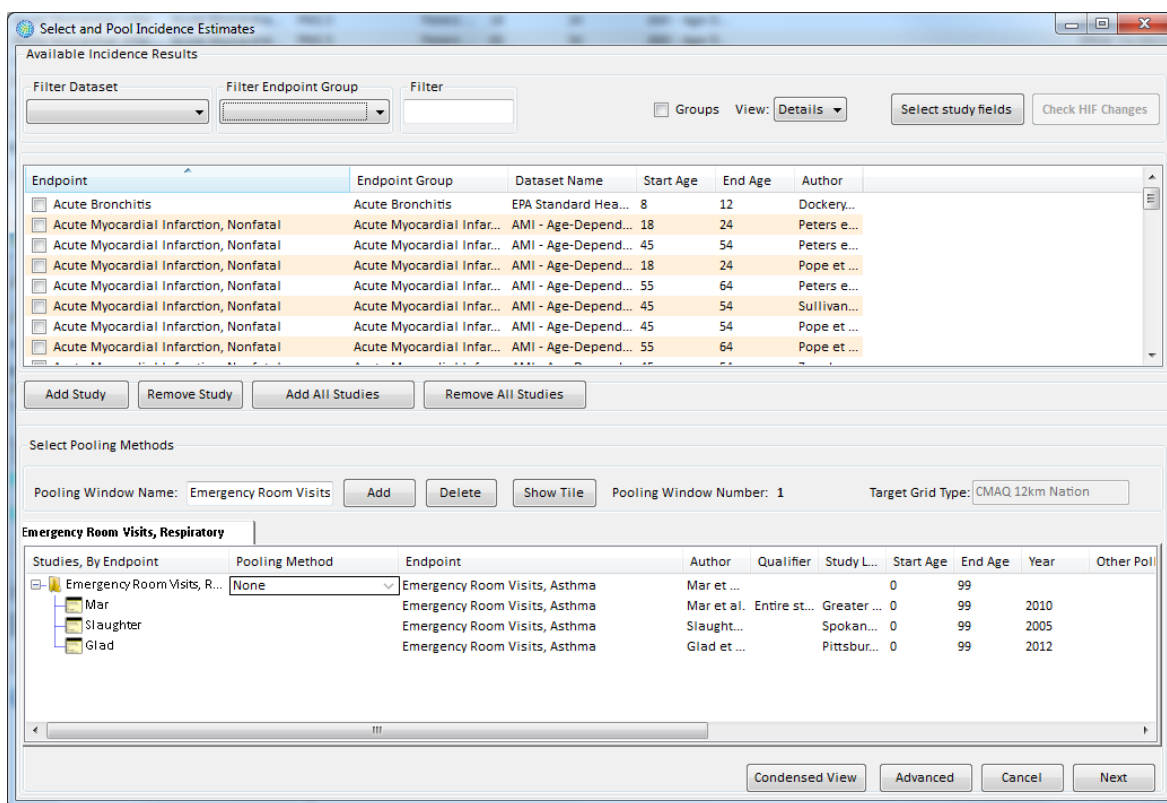
Each line in the pooling window represents a node in the tree structure, with each node representing either an individual incidence result or a collection of incidence results which have common values for their leftmost identifying variables. The tree structure is generated by comparing the leftmost values of the incidence result's identifying variables. High level nodes in the tree are formed when results have common values for identifying variables, and branches in the tree occur when the values differ.
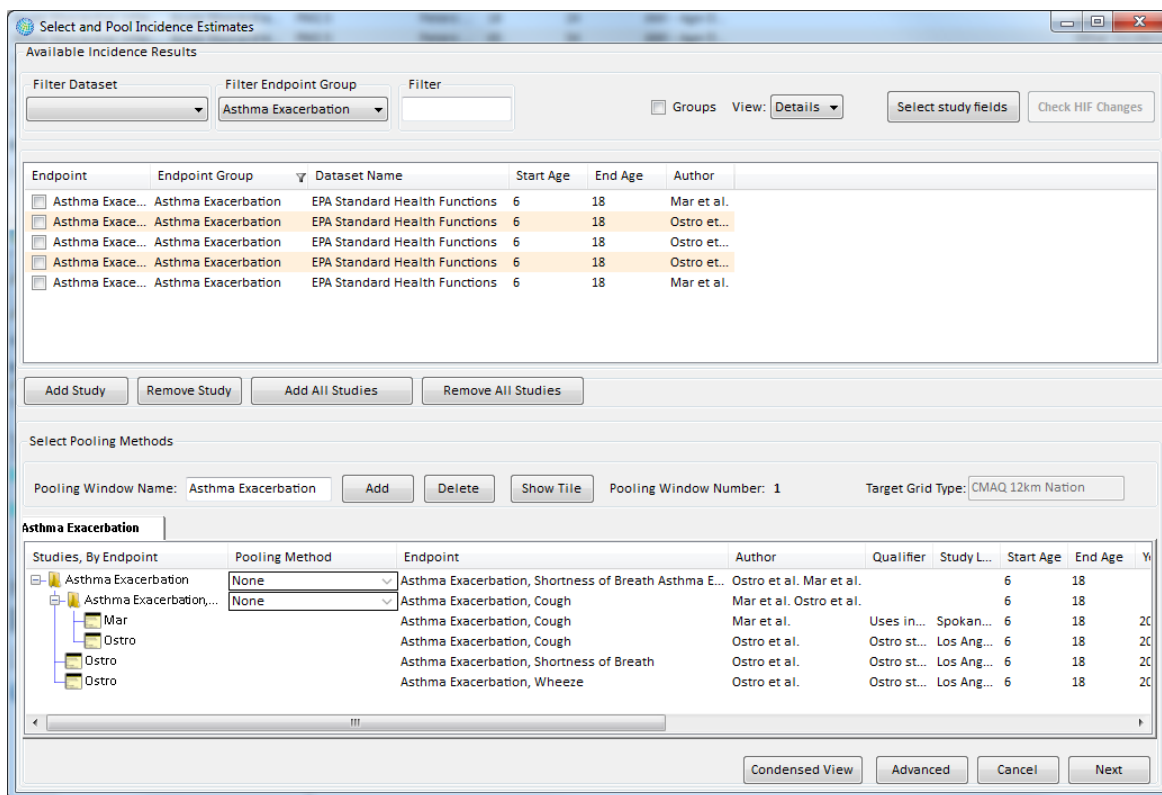
In the above example, four incidence results have been dragged into the pooling window. Each of the four health impact functions has **Endpoint Group** '*Asthma Exacerbation*'. Thus, the top line, or root of the tree structure, represents all four incidence results. A branch then occurs in the tree structure, because two studies have **Endpoint** '*Asthma Exacerbation, Cough*', while two others have **Endpoint** '*Asthma Exacerbation, Wheeze*' and '*Endpoint Asthma Exacerbation, Shortness of Breath*'. A further branch occurs within **Endpoint** '*Asthma Exacerbation, Cough*' when **Author** of the two incidence results differs. Once a node has only a single incidence result, no further branching can occur.

Step 2. Sort results

After dragging incidence results into the pooling window, you can rearrange the order of the columns (variables), and thus change the tree structure. To do this, click on a column and hold the button down as you drag it to its new location. Note that the **Pooling Method** is always the first column after the **Studies, By Endpoint** column. All the other columns can be moved. To see how the order of the columns in the pooling window affects the tree structure, consider the following example:

This example uses exactly the same incidence results as the previous example, but with the **Author** column (variable) immediately after the **Pooling Method** column.

In the **Tile View** (accessed by selecting **Tile View** from the **View Menu**), you can sort studies by endpoint tiles by selecting the appropriate criterion in the **Sort By** drop-down menu.
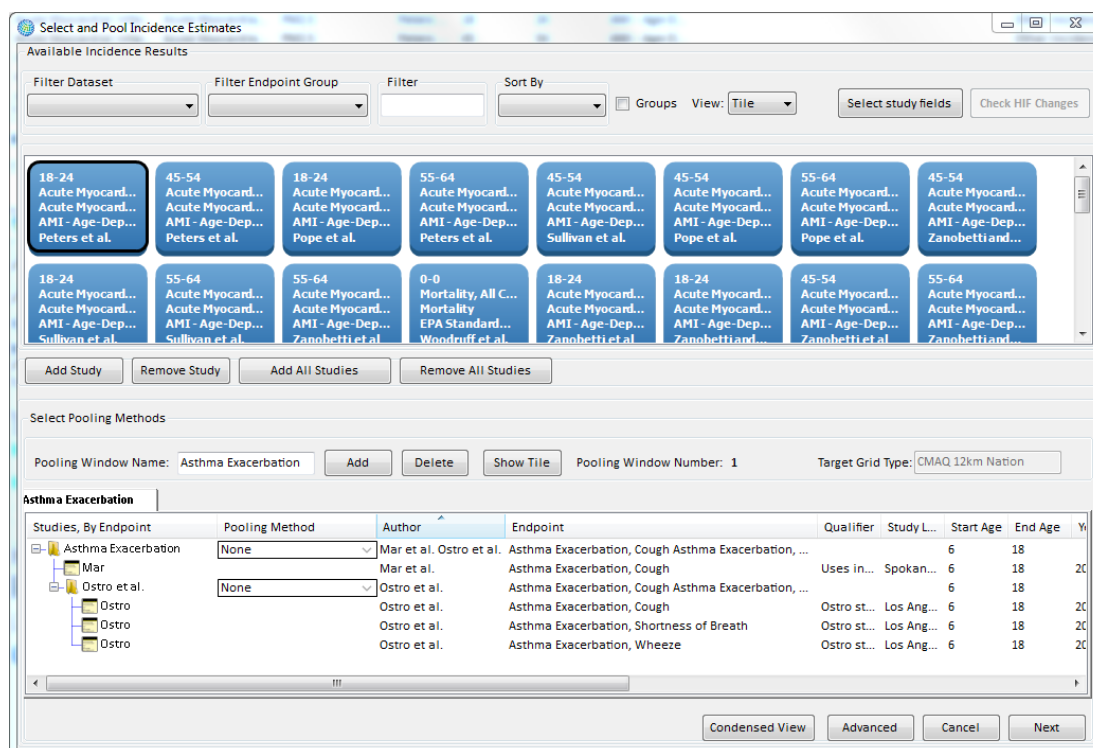
Step 3. Select pooling methods

Once the tree structure is set up in the pooling window, you are ready to select your pooling methods. Essentially each **Pooling Method** involves a different method of combining input incidence results to generate new incidence results. Results can be pooled any time a branch occurs in the tree structure — that is, any time two or more results share common values for their leftmost variables. BenMAP-CE helps you to identify these spots by inserting a value of *None* in the **Pooling Method** column at each spot where pooling is possible.

When a row that represents a group of incidence results is selected, an arrow appears at the right of the **Pooling Method** column, indicating the availability of a drop-down menu where you can select the desired **Pooling Method.** If you select **User Defined Weights** from the **Pooling Method** menu, a **Weight** column will appear to the right of the **Pooling Method** menu. All studies are assigned equal weight by default unless you update them. The sum of all **User Defined Weights** must equal 1.

If a **Pooling Method** is selected, you can view the calculated weights before executing the pooling by pressing the **Preview** button. In the case of the **Random or Fixed Effects** pooling method, the **Preview** button will reveal whether random or fixed effects are being applied.

Step 4. Create additional pooling windows if needed

Within a given pooling window, you can have only one ordering of the columns (variables). As we have seen, however, the ordering of the columns determines the structure of the tree used to pool results. It may thus sometimes be necessary for analyses to have multiple tree structures to handle the various pooling trees they require. To facilitate this, BenMAP allows additional pooling windows to be added and deleted. The pooling windows are displayed in a tabbed format.

To open a new pooling window, simply click on the **Add** button next to the **Pooling Window Name**. You may do this as many times as needed to accommodate different sort orders. You can add the same incidence results to as many different pooling windows as you like.

As needed, you can also delete a pooling window by clicking on the window you wish to delete and clicking the **Delete** button.
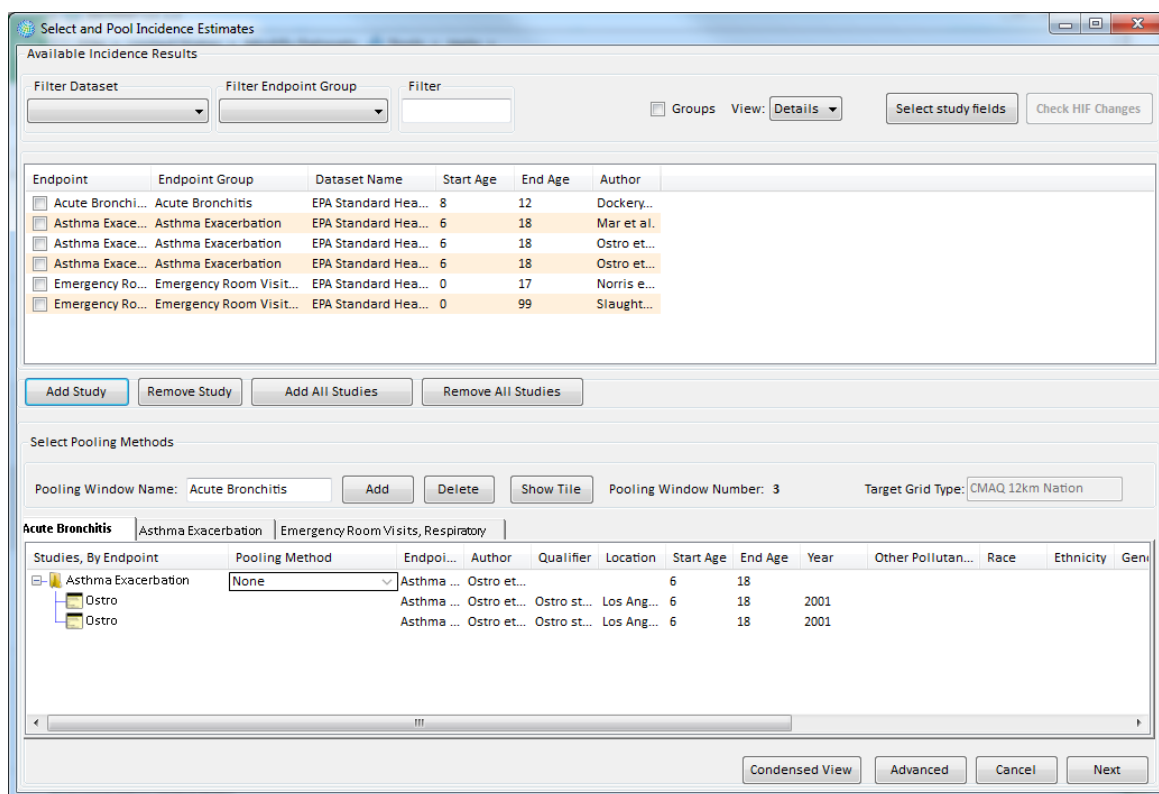
Example: Simple Sorting & Pooling

> If you add a single incidence result to the pooling window, you will see just one line, and therefore no opportunities to pool. This is shown in the example below:

> If you add a second incidence result to the window whose health impact function has the same **Endpoint Group**, but a different **Author**, you will then have a tree with two items in it. The tree branches at the point where the two health impact functions vary - at the **Author** column.

> Note that a pooling method can now be selected for the two incidence results, since a branch has appeared. If we desired to pool these two incidence results, we would end

up with a pooled result representing two '*Asthma Exacerbation, Shortness of Breath*' incidence results.

If you now add two more incidence results to the window whose health impact functions have the same **Endpoint Group** but different **Endpoints**, you will see the following:

Now you have many pooling options. Setting aside the issue of which pooling method to choose, there are four different pooling options at this point (including doing nothing), since we have two places where we can choose to pool or not to pool.

If you choose to pool at the Studies, By Endpoint corresponding to **Endpoint** '*Asthma Exacerbation, Shortness of Breath*' you would end up with three results (one pooled and two unpooled) instead of four individual incidence results.

If you choose to pool at the Studies, By Endpoint for the **Endpoint Group** '*Asthma Exacerbation*' (where the **Pooling Method** field says '*None'* in the above image*)*, you will end up with a single result representing all four of the original incidence results.

If you pool at both spots:

- First, the '*Asthma Exacerbation, Shortness of Breath*' results are pooled to a give a single '*Asthma Exacerbation, Shortness of Breath*' result.

- Next, the three separate **Endpoint** results are pooled to give a single '*Asthma Exacerbation*' result.

These same principles apply no matter how many incidence results are being pooled, and regardless of which pooling methods are selected.

Example: Multiple Pooling Windows

There are many different ways to pool your incidence results. Sometimes you may want to look at the same results in different ways, or you may just have many results that need to be sorted by different variables. In these cases, you can open up multiple pooling windows by clicking on the **Add** button.

For example, you might want to pool all results of health impact functions by a particular **Author**, rather than pooling all results of health impact functions of a particular **Endpoint**. The examples below show the same set of incidence results, first sorted by **Author**, then sorted by **Endpoint**. In order to change the sorting levels of the health impact functions, move the first category on which you wish to sort to the first position after the **Pooling Method** column. In this example, you would drag the **Author**

column ahead of **Endpoint** column. As you can see, after this change the pooling options are very different.

Sorted by Author:

Sorted by Endpoint:

If you use two different pooling windows, each sorted as shown above, you can create results pooled by **Author**, and results pooled by **Endpoint**.

7.2.3   Valuing Pooled Incidence Results

After you have specified your incidence pooling options, click on the **Next** button and select valuations and valuation pooling options from the **Select Valuation Methods, Pooling and Aggregation** window. This window should look quite similar to the **Select and Pool Incidence Estimates** window, with tree views on the left side representing the valuation methods available, and various pooling windows on the right side representing the selected valuations and pooling options.

For each pooling window you created on the **Select and Pool Incidence Estimates** window, there is a corresponding pooling window in the **Select Valuation Methods, Pooling, and Aggregation** window. You will notice that the number of incidence estimates available in the valuation pooling window will reflect any pooling choices you made in the **Incidence Pooling and Aggregation** window. For example, if in the **Select**

and Pool Incidence Estimates window you pooled 5 incidence estimates into a single incidence estimate, you will see a single incidence estimate in the **Valuation Methods Pooling and Aggregation** window.

The columns present in the **Select Valuation Methods, Pooling, and Aggregation** window are determined by the incidence results left after all incidence pooling has occurred. There will be exactly enough columns in each pooling window to represent the "least" pooled incidence result. That is, the columns will be in the same order they were in the **Select and Pool Incidence Estimates** window, but the only columns present will be those up to the level of the pooled incidence result with the most columns left over after all pooling has occurred. Here is an example:
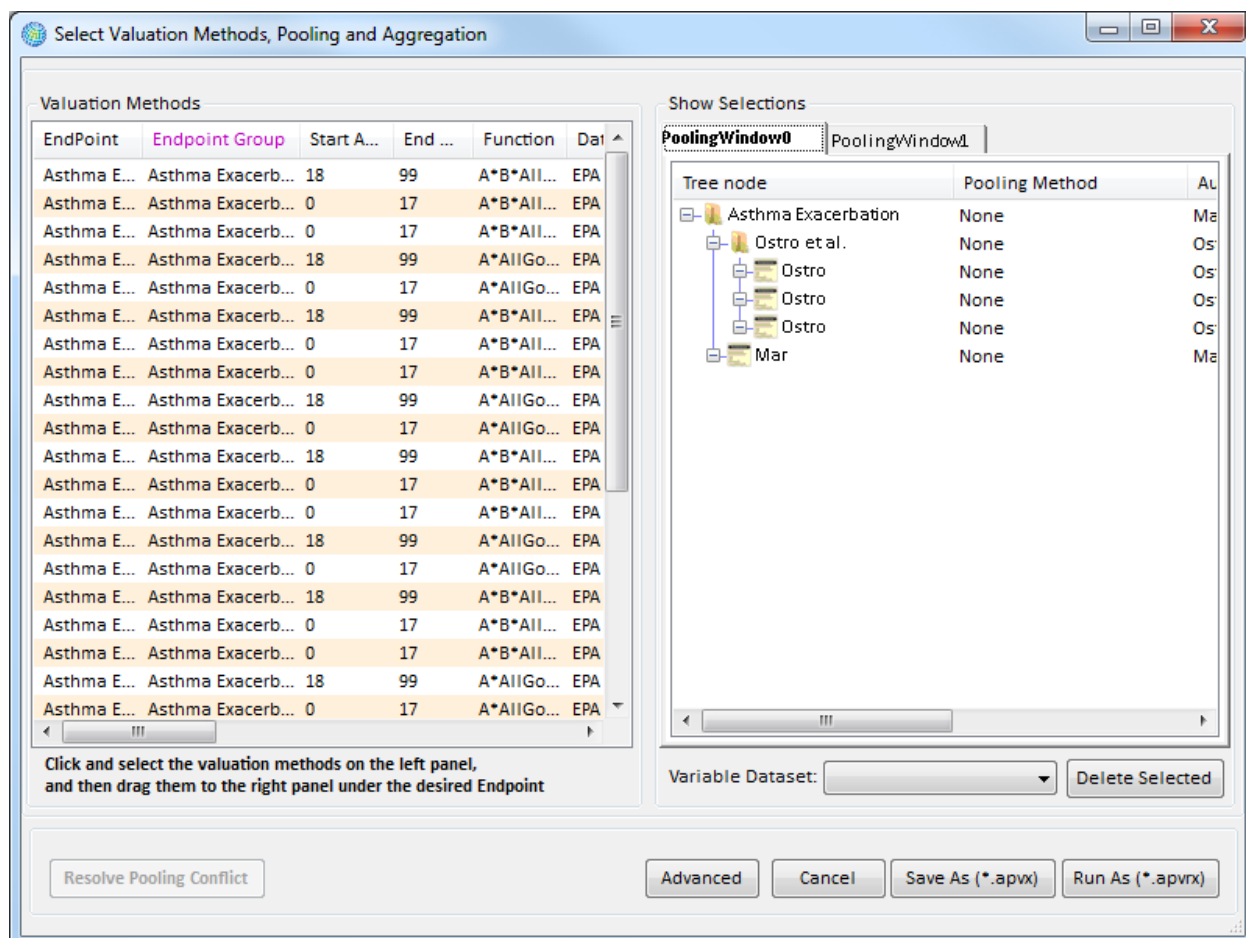
There are several steps to take in the **Select Valuation Methods, Pooling, and Aggregation** window:

Step 1. Select your valuation methods

Valuation Methods are specific to **Endpoint Groups**, and sometimes to **Endpoints** as well. The only **Valuation Methods** which appear on the left-side of the window are those which have the same **Endpoint Group** values as the pooled incidence results which are available to be valued. To select a **Valuation Method,** select it from the table

and drag-and-drop it onto the appropriate incidence result in the pooling window. Note that BenMAP-CE will only allow you to drop **Valuation Methods** onto incidence results which have the same **Endpoint Group** value. For example, BenMAP-CE will not allow you to drop a '*Mortality'* valuation on a '*Hospital Admissions'* incidence result. Note also that you can only drag-and-drop individual **Valuation Methods**, not entire groups of them. For explanations of the various valuation methods, see Appendix G.

If you have added any of your own valuation methods, as described in the Valuation Data section of Chapter 4: Loading Data, you can drag-and-drop them in the same way as the *EPA Standard Valuation Functions* shown in these examples.
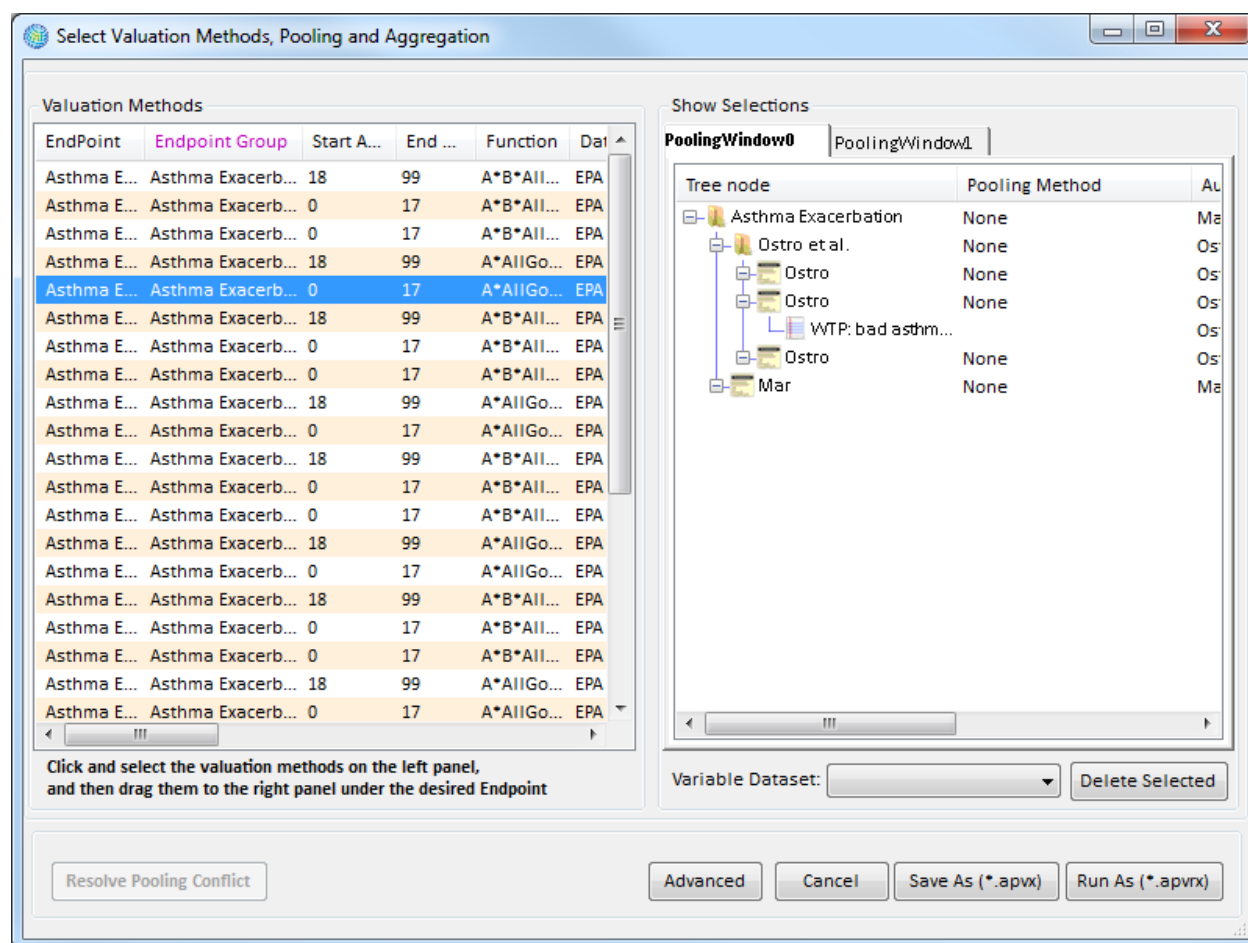
When BenMAP-CE runs the APV Configuration, it will generate a valuation result for each **Valuation Method** you select by running the method's **Function** on the selected incidence results. You do not need to select a **Valuation Method** for every incidence result—incidence results without any **Valuation Methods** will simply be ignored when valuation results are generated, aggregated, and pooled.

Because valuation functions include an uncertainty distribution around them, generating valuation results is fairly complicated. The procedure depends on whether the incidence results being used were generated in Point Mode or with the default

Monte Carlo method. See Chapter 6: Incidence Estimation (Section 6.2.3) for details on these advanced configuration settings.

In Point Mode, BenMAP-CE simply runs the valuation functions once using the point estimate of the incidence result and the mean of the valuation function as inputs.

With the Monte Carlo feature, on the other hand, BenMAP-CE generates one hundred percentile points (from the 0.5th percentile to the 99.5th percentile) to represent the distribution of the inputs to the valuation function. To get the value of the health incidence, BenMAP-CE multiplies each combination of values from the incidence result with each of the hundred valuation points, and puts the results into a holding container. (For example, if the incidence result has 10 percentile points and there are 100 valuation points, then the holding container will have 1,000 values.) Finally, the holding container is sorted low to high and binned back down to 100 percentile points (representing the 0.5th percentile to the 99.5th percentile of the economic value of the incidence).

Step 2. Sort results

Depending on how your incidence results were pooled, the columns in the valuation pooling windows can be resorted in the same way as the **Select and Pool Incidence Estimates** window columns. This resorting will have the same sort of impact on the tree structure of valuation results that it had on the tree structure of incidence results. (See Step 2 in the section on Pooling Incidence Results.)

Step 3. Select pooling methods

The same pooling methods are available for valuation results which were available for incidence results. (See Step 2 in the section on Pooling Incidence Results.) You should note that when more than one valuation method is selected for a particular pooled incidence result, it is possible to pool the generated valuation results.
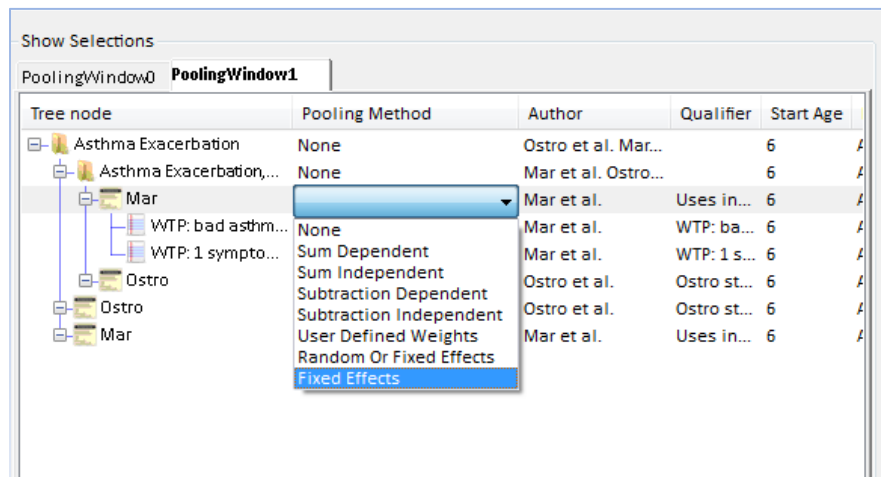
Step 4. Select Variable Dataset

In order to proceed to the next step, you must select a **Variable Dataset** from the drop-down menu beneath the pooling window. The **Variable Dataset** can include a variety of data, such as income and poverty data that might be used in health or valuation functions. For the default EPA health and valuation functions, you just need to select the *EPA Standard Variables*.

If you have developed your own setup, then you need to make sure that you also load a **Variable Dataset**. This is necessary even if you do not need the extra variables that can be included in this dataset.
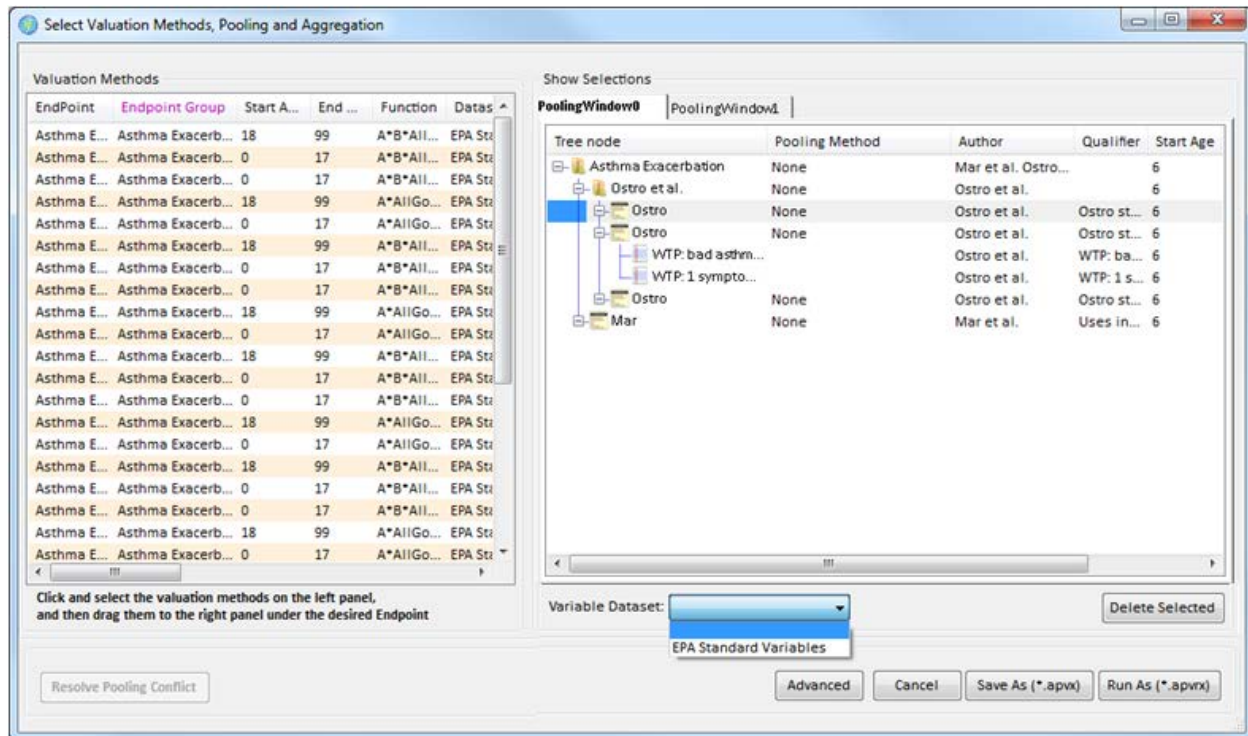
7.2.4 APV Configuration Advanced Settings

When specifying the incidence pooling options, you may click on the **Advanced** button on the bottom of the **Select and Pool Incidence Estimates window.** This button will open the **Advanced Pooling Settings** window. The features of this window are described in the next section.
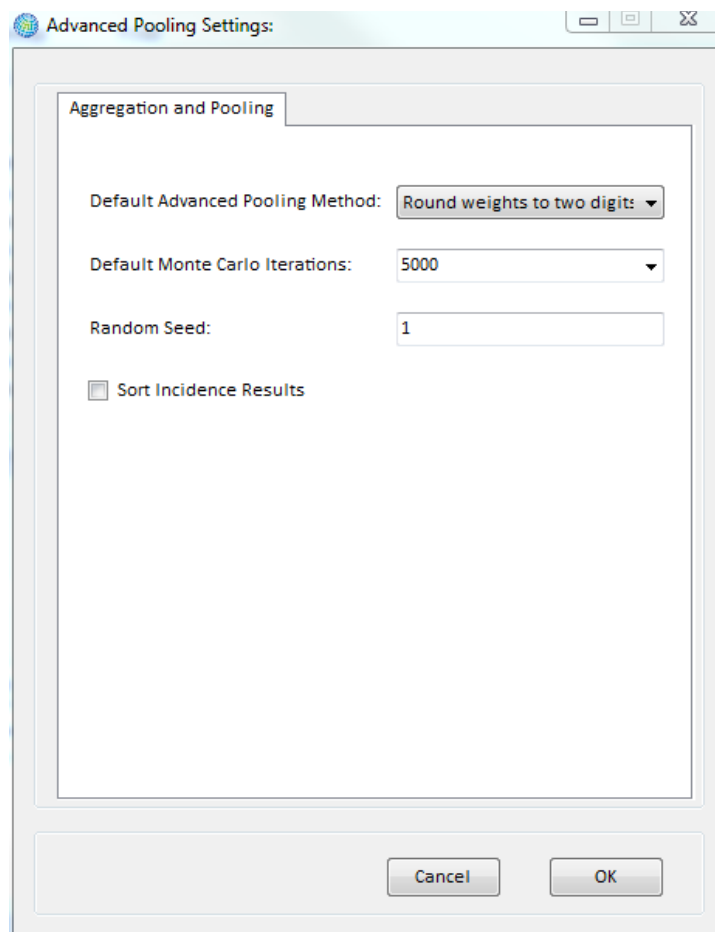
7.2.4.1   Aggregation and Pooling

Default Advanced Pooling Method

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. BenMAP-CE lets users assign "subjective" weights and it assigns weights using a fixed effects or a random effects approach. There are three options for using weights available in the **Default Advanced Pooling Method** drop-down list:

Round weights to two digits. BenMAP-CE rounds each weight to two digits (e.g. 0.73), and then multiplies these weights by 100 to get two-digit integers. Each entire distribution (set of percentile points) is then put into a holding container an integral number of times, according to its integral weight. This holding container is then sorted low to high and binned down to the appropriate number of percentile points.

Round weights to three digits. BenMAP-CE rounds each weight to three digits (e.g. 0.732), and then multiplies these weights by 1,000 to get three-digit integers. Each distribution (set of percentile points) is then put into a holding container for an integral number of times, according to its integral weight. This holding container is

then sorted low to high and binned down to the appropriate number of percentile points.

Use exact weights for Monte Carlo. BenMAP-CE uses exact weights and a Monte Carlo simulation. During each iteration of the procedure, a particular result is selected with a probability equal to its weight. Once a result is selected, one of its percentile points is chosen at random and put into a holding container. This is done some number of times (see *Monte Carlo Iterations*, below), and the holding container is then sorted low to high and binned down to the appropriate number of percentile points.

Default Monte Carlo Iterations

This drop-down list is only enabled when *Use exact weights for Monte Carlo* is selected as the **Default Advanced Pooling Method**. It specifies the number of iterations the Monte Carlo simulation should be run (see above). Its initial value is set by the **Default Monte Carlo Iterations** value from the **Advanced Pooling Settings** window (see Step 1, above).

Random Seed

The **Advanced Pooling Settings** window allows the specification of a **Random Seed**. Many of the pooling methods require the generation of sequences of random numbers, e.g. choosing a random percentile point during a Monte Carlo simulation. Providing a specific **Random Seed** value allows you to ensure that the same sequence of random numbers is generated as in a previous analysis, thus allowing exact results to be reproduced.

If you do not set the **Random Seed** for a particular run, one will be generated automatically from the system clock (the number generated will depend on the date and time, and should change every minute). Normally, you should not set the **Random Seed** value. If you need to reproduce a specific set of results, however, the random seed used to generate previous APV Configuration Results can be determined from an APV Configuration Result file (*.apvrx*) Audit Trail Report.

Sort Incidence Results

The *Sort Incidence Results* should generally be always checked. This setting ensures that the incidence percentile-point results are sorted low to high.

7.2.4.2   Currency Year and Income Growth

The **Currency Year and Income Growth** window, accessed via the **Advanced** button on the **Valuation Method** window allows you to specify an **Inflation Dataset** and a **Currency Year**, which in combination allow you to change the currency year to account

for inflation. The **Income Growth Adjustment** panel allows you to adjust the valuation estimates to account for the growth in income over time for WTP functions.

Inflation Adjustment

The **Inflation Adjustment** needs to be carefully considered in relation to the valuation dataset that you are using. (This is discussed in detail in the section on loading inflation data in Chapter 4.) The default valuation database in the *United States* setup has a currency year of *2000*, so the inflation dataset has a value of *1* for the year *2000*.

Income Growth Adjustment

Willingness to pay (WTP) estimates are believed to be tied to the income of individuals. As income rises over time, WTP estimates are likely to increase as well. The **Income Growth Adjustment** is designed to take this phenomenon into account, allowing you to account for income growth between the time when WTP estimates were calculated and the year of your analysis.

As with the **Inflation Adjustment**, the **Income Growth Adjustment** has a close connection to the valuation estimates. For example, the valuation estimates in the *United States* setup are assumed to be based on income levels from 1990, so the income growth adjustment database has a value of 1 for the year 1990. (This is discussed in detail in the section on loading income growth data in Chapter 4.)

To use the **Income Growth Adjustment**, you need to choose a dataset and then choose the income year that you want to use. It is common to set the **Year** variable to the year of the population forecast in your analysis. Of course, you can only choose from the available data. If the income growth adjustment data only goes to 2024 and the population data in your analysis are for 2030, then there will be some, unavoidable mismatch.

7.3   Open & Modify Existing APV Configuration

If you have an existing APV configuration (*.apvx) file or APV result file (*.apvrx), you can open, and edit it. Double click **Aggregate, Pool & Value** from the main tree menu to load the APV Configuration or APV Result file.
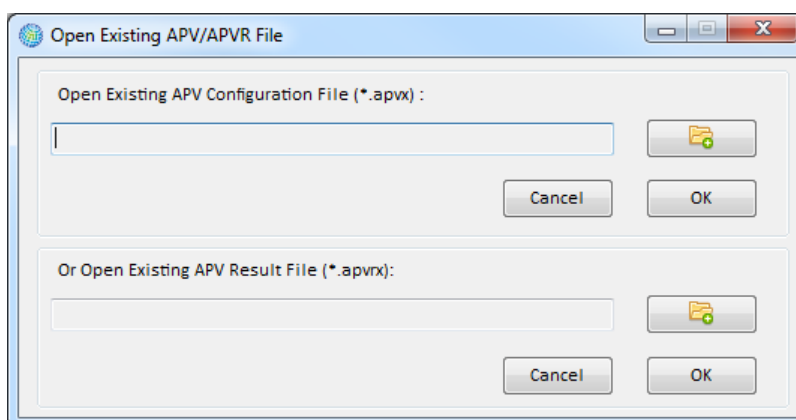
If you have only a few changes to make to an existing configuration, it is typically much quicker to open the previous configuration, rather than entering all of your choices again. Note that the various parts of an APV Configuration are quite interdependent, so modifying part of the configuration may cause other parts to be reset. For example, modifying the tree structure for incidence pooling will cause the valuation method selection and valuation pooling tree structure to be cleared and reset. Changing the **Configuration Results Filename** in the **Select and Pool Incidence Estimates** window will not reset the incidence or valuation pooling trees as long as the new file contains incidence results generated from the same health impact functions as the old file. This can be quite helpful for generating new APV Configuration Results from several different Configuration Results files which were generated from different baseline/control scenarios, but with the same set of health impact functions.

7.4   Run APV Configuration

After having specified the various aggregation, pooling, and valuation options, you can save your APV Configuration for future use. The file that you save has an ".apvx" extension. The configuration that you have specified for APV is similar to the configuration that you developed for choosing health impact functions. (That configuration has a ".cfgx" extension.) Both files allow you to save choices that you have made, and re-run them at a later time.

To save your APV configuration with your valuation pooling choices, click the **Save As (*.apvx)** button, and name your configuration file. We suggest that you save this in the Configurations folder. When ready to generate APV Configuration results, click the **Run As (*.apvrx)** button. BenMAP-CE then requires that you specify a file in which to save the results, with an ".apvrx" extension.

7.5   View and Export Pooled Incidence and Valuation Results

Using the results from the APV Results file (".apvrx" extension), you can create, view and export reports that reflect the choices you made about how to aggregate, pool and value your results.

7.5.1   Pooled Incidence Results

The **Pooled Incidence Results** report provides results aggregated and pooled to the level that you previously specified in the Aggregation, Pooling, and Valuation Configuration file. This report has fewer Health Impact Function fields than the **Aggregated Incidence Results Report**, and values for others will be blank. This is because after pooling, only enough fields are retained to uniquely identify individual results.

To generate pooled incidence results, click the **Pooled Incidence** tab in the upper portion of the main window. Double-click to select the study results you would like to view. The selected results should show up in the **Data** tab below, if not, click the **Show Results** button. Notice that you cannot re-aggregate the results in this stage, because you have already defined how to aggregate the results. You may also view the results on a map using the **GIS Map** tab and on a bar chart using the **Chart** tab.

7.5.2   Pooled Valuation Results

The **Pooled Valuation Results** report presents valuation results aggregated and pooled to the level you specified using the **Advanced** button when creating the APV configuration file. As with the **Pooled Incidence Results Report**, fewer Pooled Valuation Method fields are available, because only enough fields are retained to uniquely identify individual results.

Click the **Pooled Valuation Results** tab to begin viewing these results. Similar to the Pooled Incidence Results, you can double-click the study of interest and view the results

table in the **Data** tab. Similarly, you may view the results on a map using the **GIS Map** tab and on a bar chart using the **Chart** tab.

7.6   Frequently Asked Questions

I am at the BenMAP-CE valuation window and cannot proceed. What should I do?

In order to proceed to the next step, you must select a Variable Dataset from the drop-down menu in the Select Valuation Methods, Pooling, and Aggregation window. The files in the Variable Dataset can include a variety of data, such as income and poverty data that might be used in health or valuation functions. For the default EPA health and valuation functions, you just need to select the EPA Standard Variables.

If you have developed your own setup, then you need to make sure that you also load a Variable Dataset. This is necessary even if you do not need the extra variables that can be included in this dataset.

How do I edit or add other valuation functions?

To edit or add valuation functions you need to go to Modify Setup option in the Tools drop- down menu available in the upper left-hand corner of the main BenMAP-CE window. See the valuation function section in the chapter on Loading Data for details on how to do this.

How do I know what year dollars (currency year) were used?

You can find the answer in the Audit trail for the APVR file that you generated.

Do the currency year and year of the population data have to match?

No. The currency year and the year of the population data do not need to match. Currency years are always historical because we do not forecast inflation.

Chapter 8
GIS/Mapping

In this chapter…

- **Learn about BenMAP-CE's mapping functions.**
- **Map different variables and modify the map display.**

Chapter 8 Table of Contents

The BenMAP-CE Geographic Information System (GIS) will display maps of air quality, health and economic data. These maps can help answer a number of questions:

- *Quality assurance*: Do your air quality changes seem to be distributed correctly? Are your air quality changes and health impacts occurring in approximately the same location?

- *Presentations*: In what states or provinces are most of the benefits/disbenefits of your policy scenario concentrated?

- *Analysis*: Which air quality grid cells contain the highest ozone values?

The main **GIS Map** tool will be available once you have successfully completed the first stage of the BenMAP-CE analysis (**Air Quality Surfaces** in the main BenMAP-CE tree menu).

8.1   Overview of Mapping

The GIS will map three categories of data:

1. *Air quality (.aqgx)*. Air quality grid maps represent summary air quality metrics (e.g., daily average, daily maximum, or other metric where available) within each grid cell. Air Quality Grids can be added to the **GIS Map** by following the steps in Chapter 5: Creating Air Quality Grids.

2. *Incidence (.cfgrx)*. A configuration results file contains the results of your analysis reported at each air quality grid cell. These results have not been aggregated, pooled or valued. For more information on choosing configuration settings, see Chapter 6: Estimating Incidence.

3. *Aggregated, pooled and valued results (.apvrx).* These are results that have been aggregated to a coarser spatial scale, see Chapter 7: Aggregation, Pooling, and Valuation.

When mapping APV Configuration results, you can generate six different types of maps: (1) *Incidence*, (2) *Aggregated Incidence*, (3) *Pooled Incidence*, (4) *Valuation*, (5) *Aggregated Valuation*, and (6) *Pooled Valuation.* For more information on how to load these files into BenMAP-CE, see Chapters 6 and 7 (Estimating Incidence, and Aggregation, Pooling, and Valuation).

8.2   Results Panel

Once you have gone through all the steps outlined in the previous chapters to import data and files, you are ready to begin exploring the visual results. Each of your selected endpoint groups will be displayed in the results panel (upper right portion of the main BenMAP-CE window).
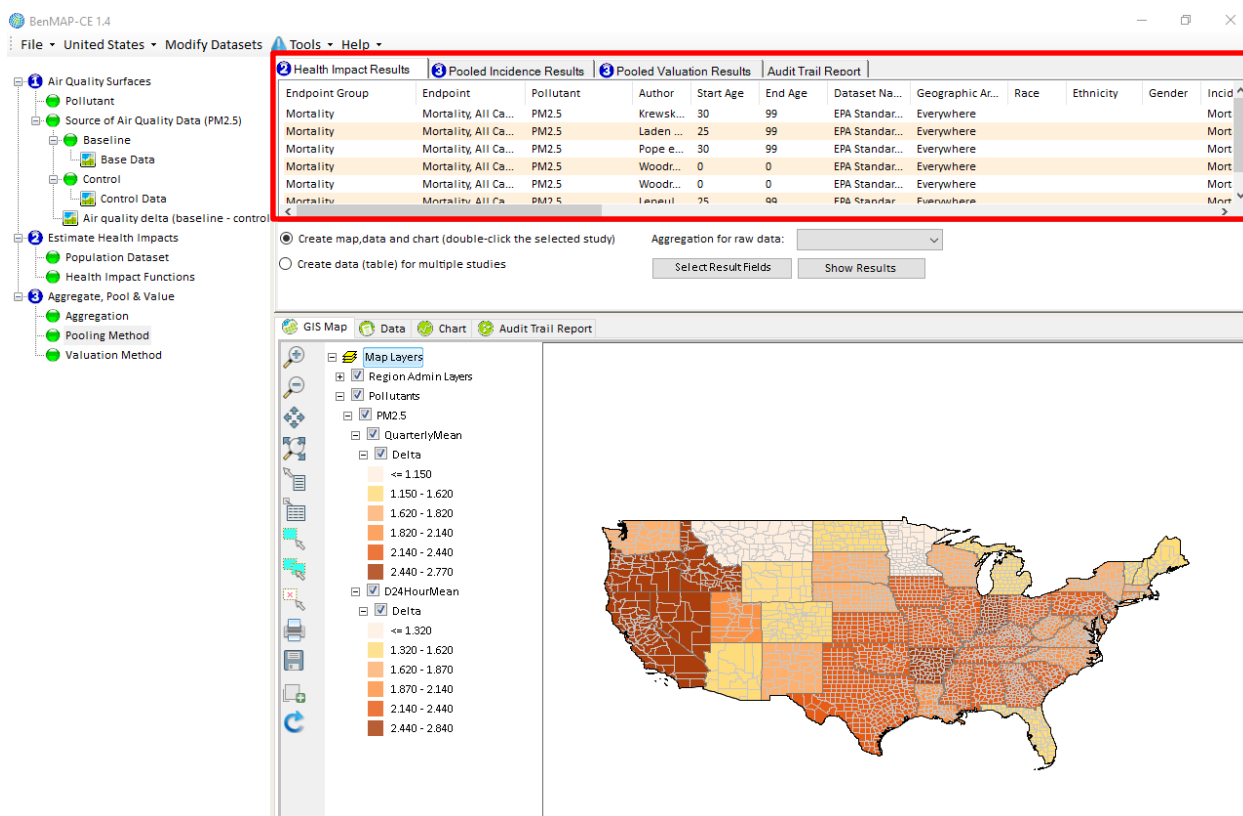
There are four tabs within the results panel, **Health Impact Results**, **Pooled Incidence Results**, **Pooled Valuation Results**, and **Audit Trail Report**. To create a results map layer for an endpoint group, select an entry, or entries, from the list and drag it down onto the GIS Mapping panel. BenMAP-CE will create the map and place an entry into the GIS table of contents below the newly created results group. The layer will most likely be listed under the author's name as subgroup of the main results group. This process of dragging-and-dropping can be done with as many entries as you like, on any of the first three tabs.

Audit Trail Reports facilitate transparency and reproducibility by reporting a summary of your assumptions underlying each step of the analysis. This is described in more detail below.

8.3   GIS Mapping Panel

The GIS Mapping panel is the centerpiece of this tool. Here is where you can view, edit, add, and remove layers from the table of contents and GIS map viewer.
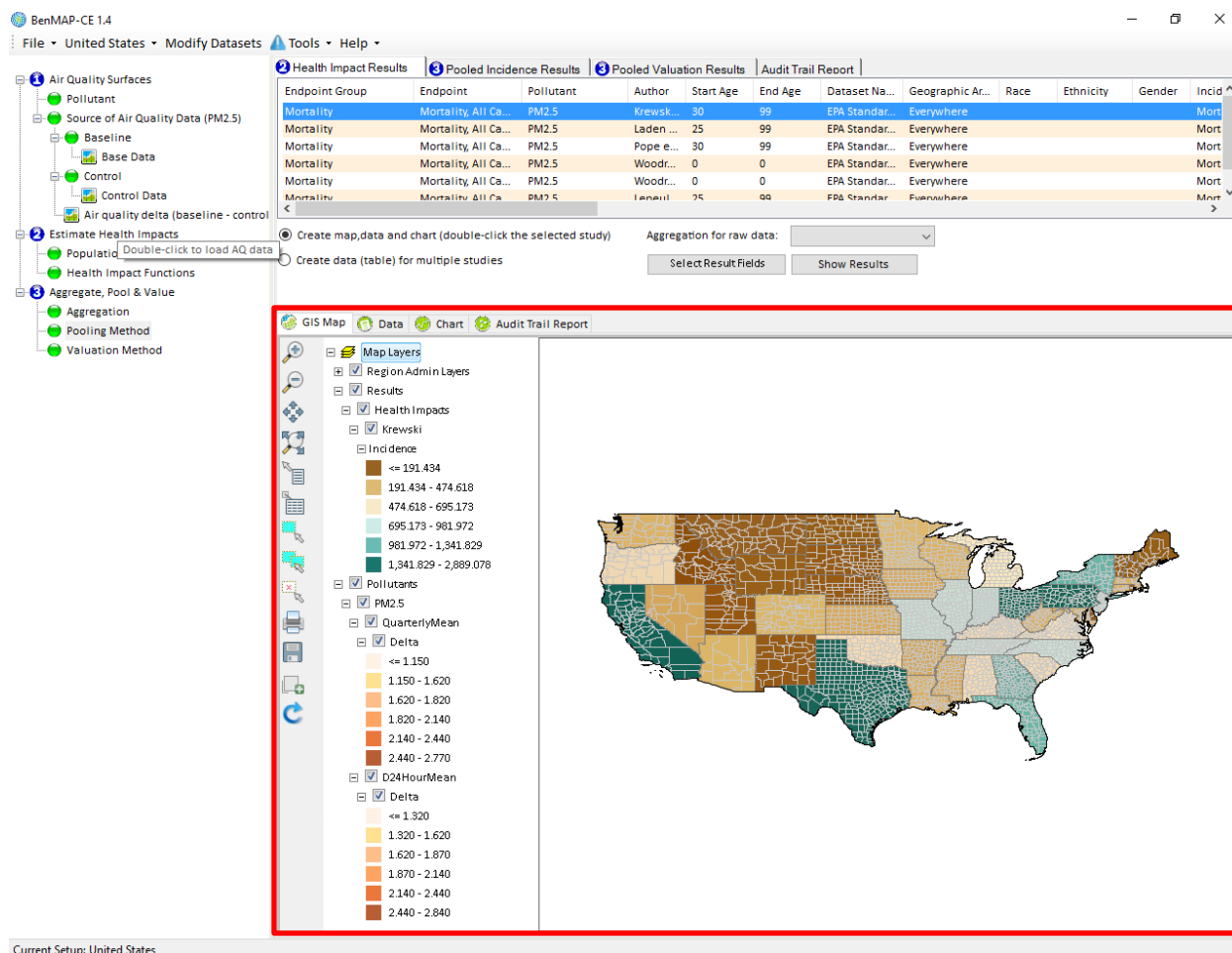
8.3.1   GIS Map Tab

After creating an air quality surface, the **GIS Map** tab is used to view the air quality data (double-click on an available air quality surface to display it). Here you will find a GIS table of contents, toolbar, and interactive map.

8.3.1.1   GIS Table of Contents

The GIS table of contents is where you will find all your loaded map layers. The layers are sorted in groups, with subgroups below them. The general setup will include *Region Admin Layers* group, *Pollutants* group, and *Results* group. You can select or deselect as many layers or groups as you like, for viewing on the map to the right.
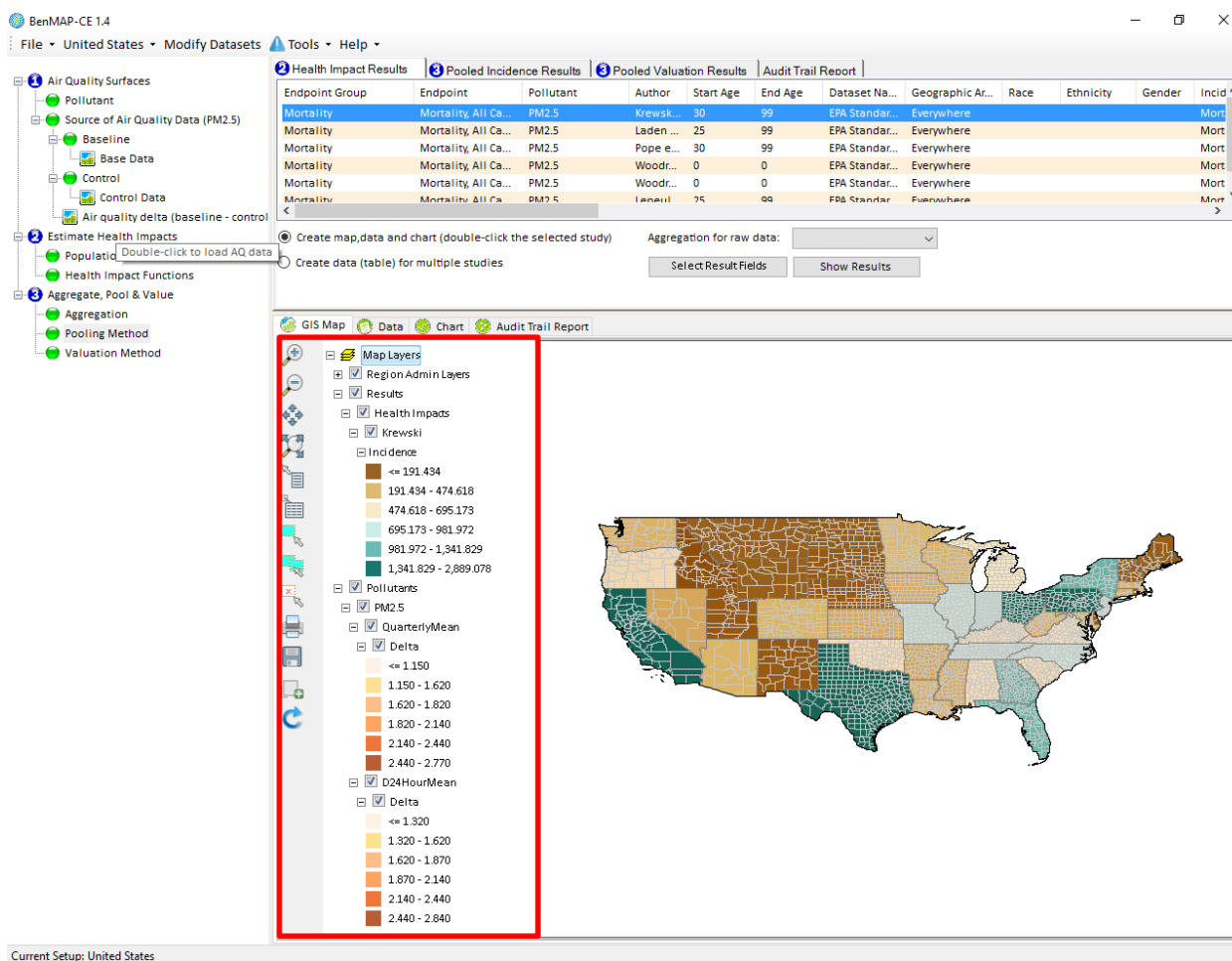
- **Region Admin Layers**: This is where the administrative grids can be found for country, state, and county boundaries. For preloaded data, regional administrative layers have been defined. For example, in the *United States* setup, the *Nation*, *State*, and *County* layer will be automatically selected. As discussed in Chapter 4, you can preselect default administrative layers for new setups you create when you import air quality grids. You can turn administrative layers on or off by expanding the group.

- **Pollutants**: This is where all the available pollutant data (from the tree menu) will be visible. Under this main group, there will be a group for each pollutant that was selected (e.g., *PM$_{2.5}$*). Below the individual pollutant, there will be metrics that were defined earlier during import and loading (e.g., *Quarterly Mean*, *D24HourMean*). Below each metric, there will be entries for each Air Quality Grid that was loaded (e.g., *Baseline*, *Control*, and *Delta*).

- **Results**: This is where the layers for Health Impact Results, Pooled Incidence Results, and Pooled Validation Results will be listed. Under the results group, there will be a subgroup labeled for each set of results. Usually, these entries will be labeled by the study author's name.

Color Ramps

The color ramps are standard for the imported files. The default color ramp for the **Baseline** and **Control** Air Quality Grids goes from light green to dark blue. The default color ramp for the **Delta** Air Quality Grid goes from light yellow to dark red. Finally, the default color ramp for the **Results** group entries will be different from each other and any other color ramps that are already in the table of contents.

These color ramps can be changed by the user, if desired. There are two possible ways to change the colors:[1]

1) To change the whole color scheme: Right click on a layer and select the **Properties** option. This will open a **Layer Properties** window, where you can change the color ramps and other properties.

---

[1] BenMAP-CE uses DotSpatial to incorporate, analyze, and map spatial data. The editing tools for layer properties are those included in the DotSpatial libraries and have not been customized for BenMAP-CE. For more information, see: http://dotspatial.codeplex.com/.
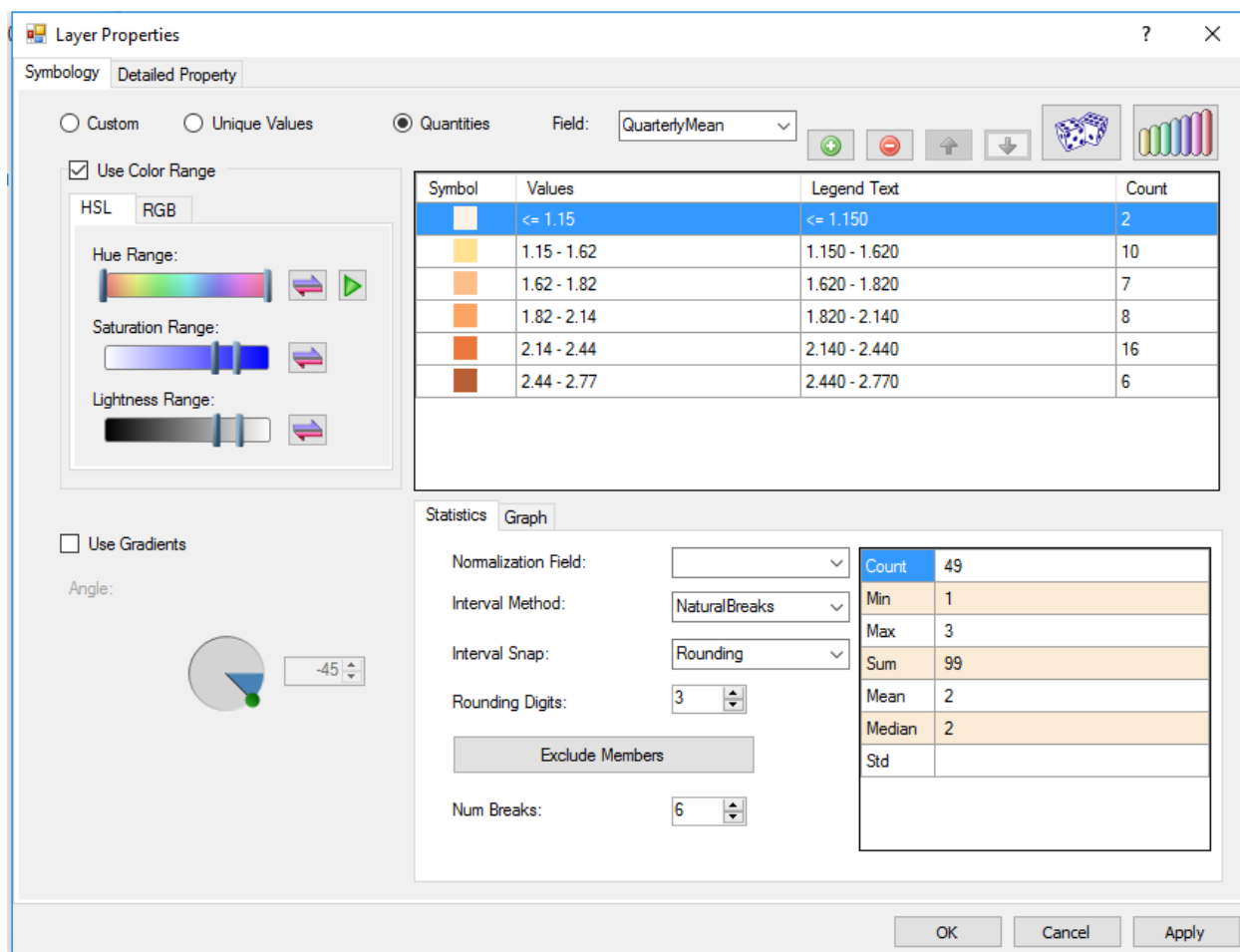
2) To change one index within a color ramp: Click on the color box or number range that you wish to change. This opens the **Polygon Symbolizer Properties** window.
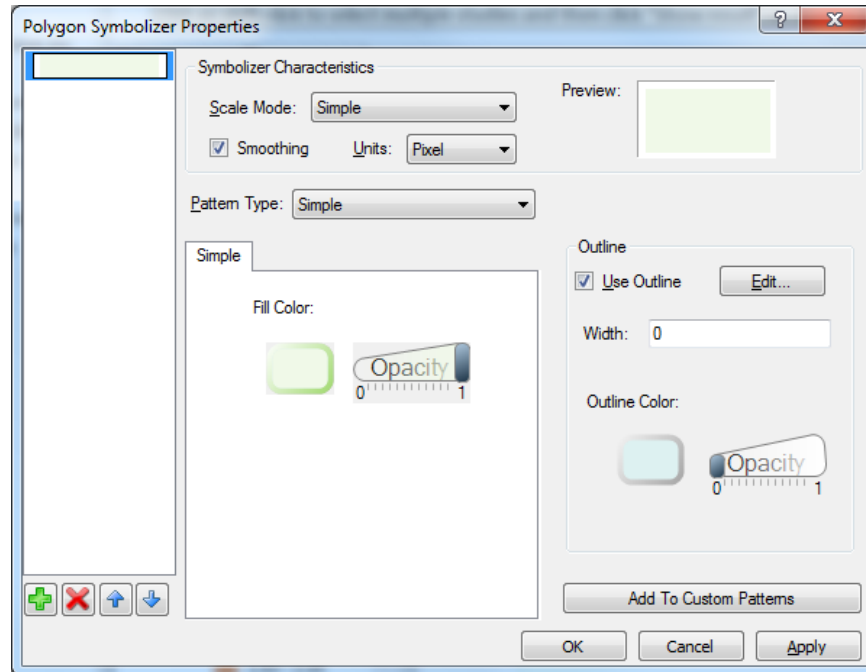
Add/Remove a Group

To add a group to the GIS table of contents, right-click on a group, and click **Create New Group** from the pop-up menu. This will add the new group below the group that was right-clicked on.

To remove a group from the GIS table of contents, right-click on the group you wish to remove, and then click **Remove Group** from the pop-up menu. This will remove the entire group from the table of contents.

Remove a Layer

To remove a spatial layer from the GIS table of contents, right-click on the layer you wish to remove, and then choose **Remove Layer** from the pop-up menu. This will remove the layer from the table of contents.

Note: Adding a layer will be discussed in the GIS Toolbar section (Section 8.3.1.2).

8.3.1.2   GIS Toolbar

There are a number of standard buttons used in most map viewing programs which you can use to navigate and customize the map view. To see the name of each button in the toolbar (to the left of the GIS table of contents), simply hold the cursor over it.

Increase zoom. Allows you to zoom in.

Decrease zoom. Allows you to zoom out.

Drag mode. Allows you to manually move the map by clicking and dragging.

Zoom to full extent. Allows you to view the whole map that you are viewing.

Click to display info for the cell in popup window. Allows you to display information (all the variable values) for individual cells or points by clicking on them.

Attribute Table. Opens an Attribute Table Editor Window where the user can view, edit, or filter the data for a specific layer.

Select Features. Allows you to select features from the map. Press and hold Ctrl to select multiple features at one time.

Select By Location. Allows you to select features from a layer based on its spatial relationship with another layer.

Clear Selection. Clear selections from all layers.

Add Layer. Allows you to add a new layer to the table of contents and map. Typically the new layer is added near the top of the table of contents.

Export Map Image. Prepare print layout of GIS map and legend.

Save Shapefile. Save and export results from a GIS layer on current map.

Change Projection to *projection type*. Allows you to change/toggle the type of map projections between the following types: GCS/NAD 83 (standard), Albers.

Creating a Query

You can create a query on any map layer that shows up in the GIS table of contents. After importing a file (such as **Baseline** or **Control**) you can begin a query by selecting a layer. The selected layer will become highlighted with a light blue bubble around it. Next, click the **Attribute Table button** in the GIS toolbar. This will open the **Attribute Table Editor** window, shown below:
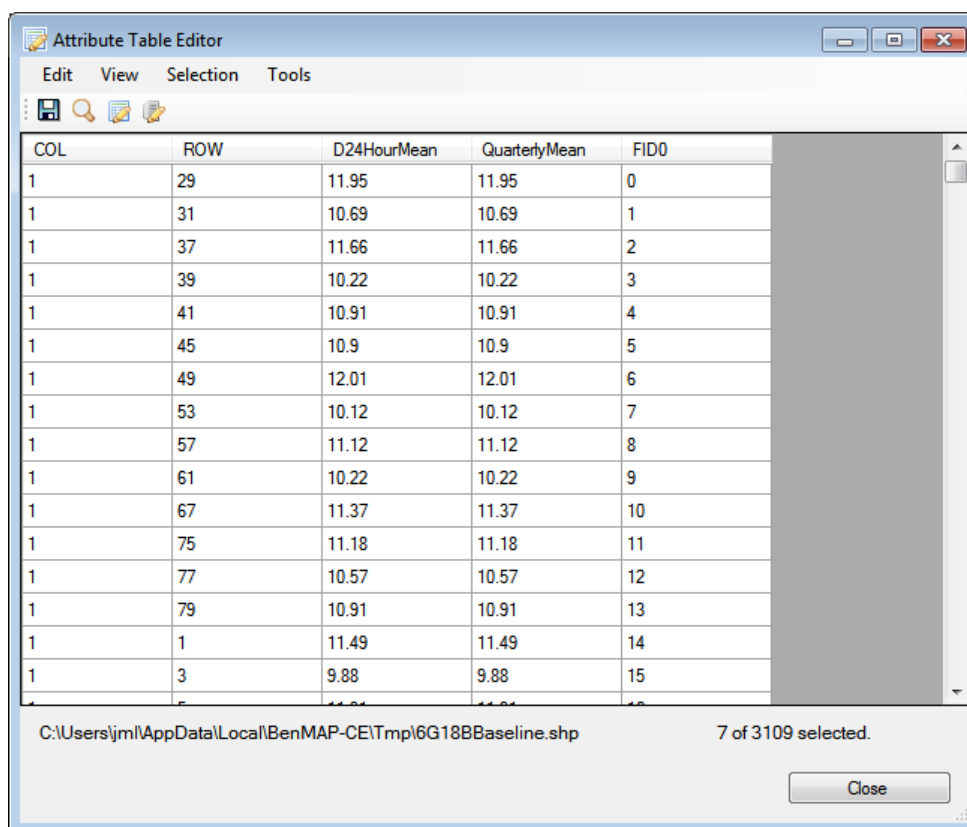
From the window above, click **Selection** from the top menu bar. This will open a drop-down menu where you can select the **Query** option.
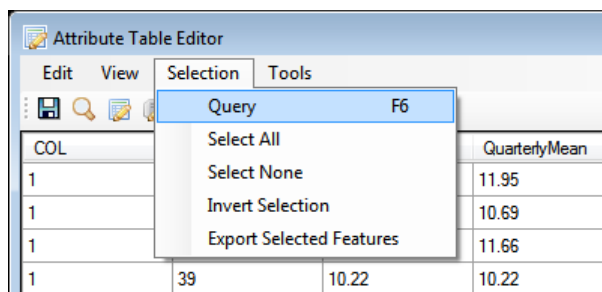
This will open **Expression Editor** window (shown below):
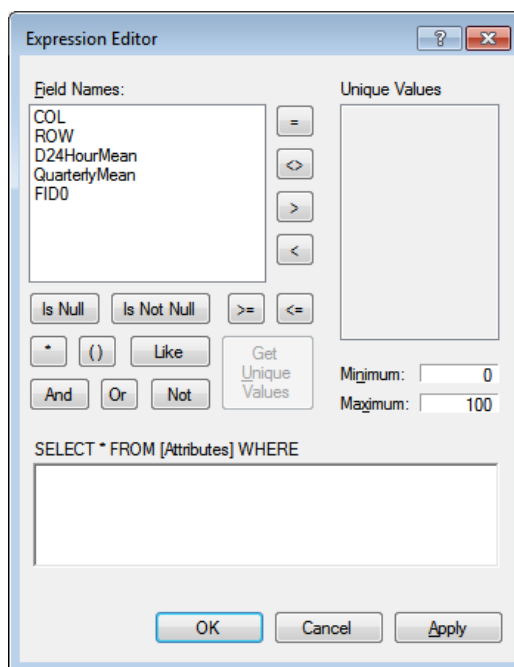
The **Expression Editor** is used to query your data. First, select a **Field Name** to query, by double-clicking on a desired entry. The **Field Name** should show up in brackets in the bottom text box. Next, select an operator from the panel of buttons. The selected operator symbol should show up in the bottom text box next to the field name. Finally, to complete the first query entry, click in the bottom text box (after the operator symbol) and enter a value that you would like to compare against (e.g., *[D24HourMean] >= 15)*. More attributes can be added to the query by clicking the **And**, **Or**, or **Not** buttons. Once you are satisfied with the query statement, click **OK** on the **Expression Editor** window. It may take a few minutes to find all the results. The **Attributes Table Editor** should appear, with the requested selections highlighted. The map on the GIS portion of the main BenMAP-CE home screen should show the query results.

To save these results, click the **Selection** button from the top menu bar of the **Attribute Table Editor** window. From the dropdown menu, select the **Export Selected Features** entry. This will open a **Save As** window, where you can save the query into a shapefile (*.shp*). Click the **Close** button on the **Attribute Table Editor**.
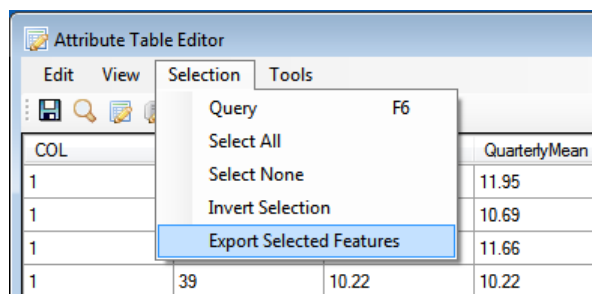
To add the query as a new layer, click the '**+**' button located in the GIS toolbar (located next to the GIS table of contents). An **Open** window will be displayed, where you can select your recently saved shapefile. Select the file and click **Open**. The new layer should appear near the top of the GIS table of contents. The layer can be dragged-and-dropped into any map group you would like within the GIS table of contents.

8.3.2 Data Tab

The **Data** tab allows you to view all the data that is being presented in the map.

At the bottom of this tab, there are a few options:

▪ The left-most set of buttons allows you to toggle between pages of data.

▪ The middle option allows you to change the number of digits that appear after the decimal point.

▪ The far right entry allows you to export the data table. Clicking the **Export** button opens a **Save As** window, allowing you to save the data as a .csv file.

8.3.3 Chart Tab

The **Chart** tab allows you to select certain regions of data within the main layer to compare localized results more easily.
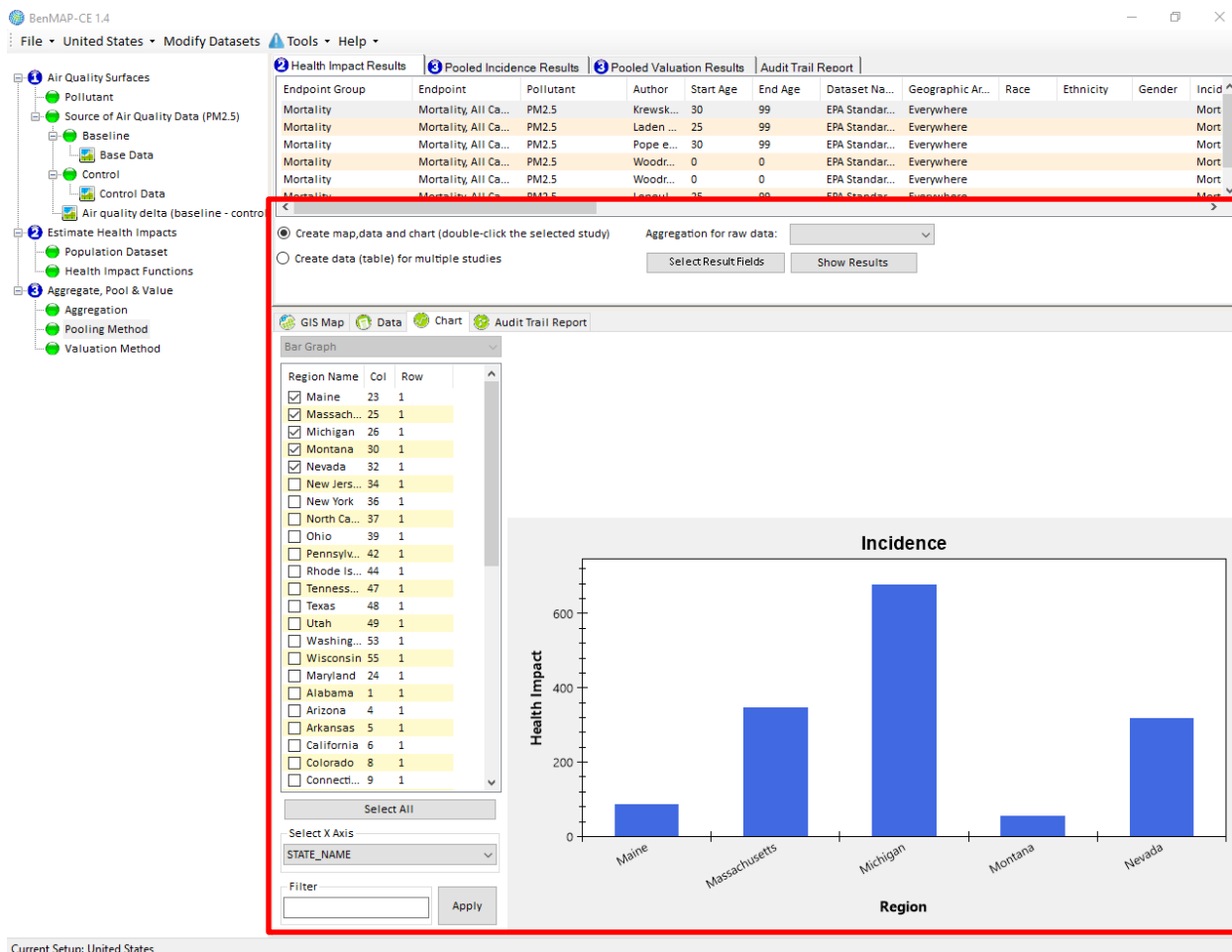
You have the option of selecting the regions that you would like to compare using the list to the left of the chart by checking and unchecking certain regions. The chart automatically updates with each new selection.[2]

8.3.4 Audit Trail Report

Audit Trail Reports facilitate transparency and reproducibility by reporting a summary of your assumptions underlying each of five types of files generated by BenMAP-CE: Air Quality Grids (with the *.aqgx* extension), Incidence Configurations (with the *.cfgx* extension),

---

[2] The data charts were originally developed using ZedGraph software, which is no longer supported. The BenMAP-CE development team is considering options to update and improve this feature.

Configuration Results (with the *.cfgrx* extension), Aggregation, Pooling, and Valuation Configurations (with the *.apvx* extension), and Aggregation, Pooling, and Valuation Results (with the *.apvrx* extension).
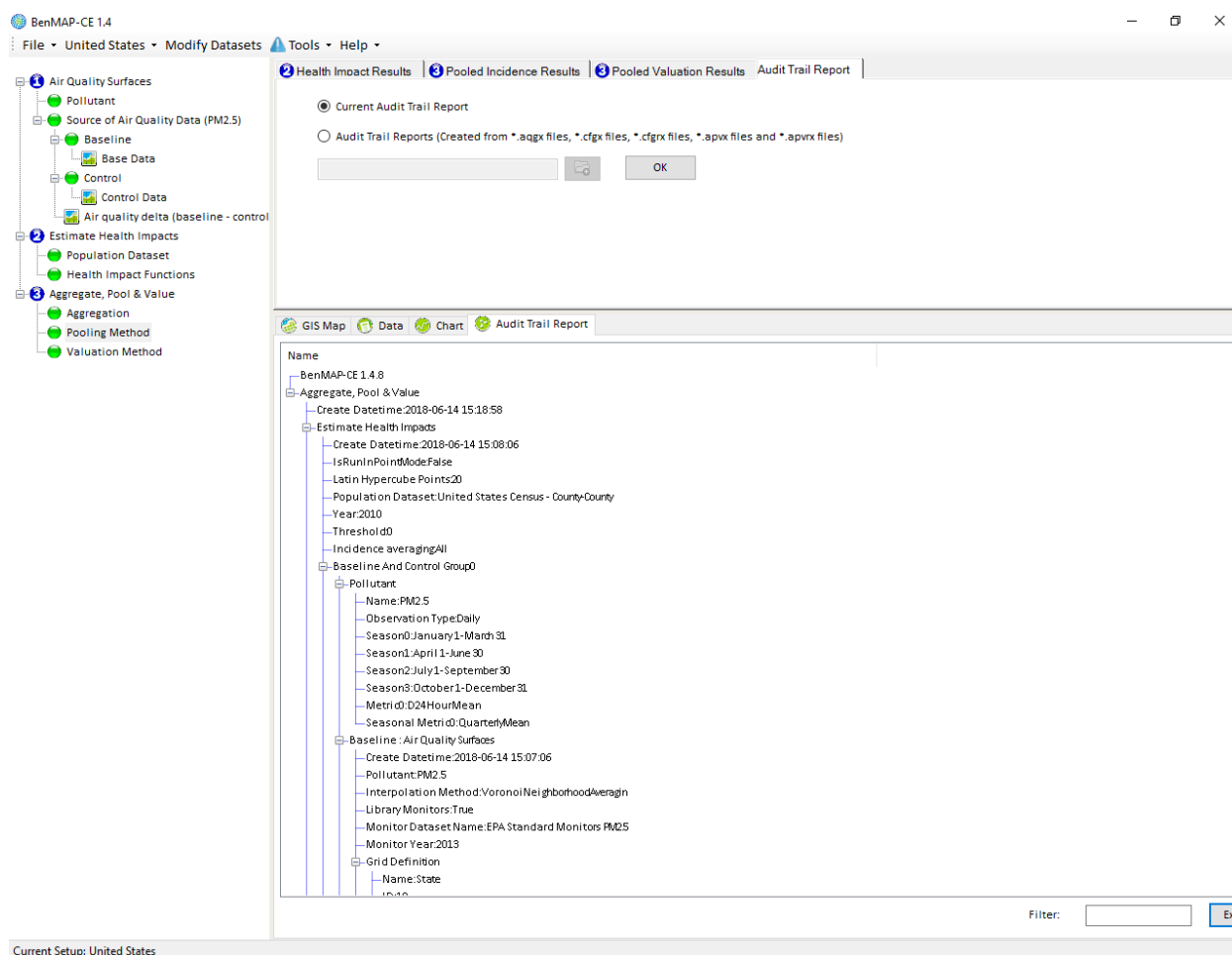
Note that each successive step in an analysis contains a summary of its inputs and attributes, and those of each previous step in the analysis. For example, in the above report the attributes of the *Health Impact Function* file used to generate the APV Results are present in the *Estimate Health Impacts* node. Similarly, the metadata for both the baseline and control air quality grids are present under the *Estimate Health Impacts* node.

The process of creating an Audit Trail is described below:

- Click the **Audit Trail Report** tab in the results window. Select *Current Audit Trail Report* (this is the default setting). Click **OK**.

- Carefully review the report, ensuring that the air quality grids, population data, health incidence data, health impact functions and economic value estimates appear as you expected.

▪ Click the **Export** button to save the audit trail report. The default location for saving audit trail export (*.txt*, *.ctlx*, or *.xml*) files will be the location you chose earlier for saving your shape files.

Audit Trail Reports have three export options: *.txt*, *.ctlx*, and *.xml*. These file types can all be viewed using a standard text editor. The *.txt* and *.xml* files will contain the same information displayed in the **Audit Trail Report** window (however, the *.xml* file contains tags to retain the tree structure). If you are familiar with the command line feature, you may use the audit trail report to produce a control (*.ctlx*) file using an existing analysis, rather than creating one from scratch. The control file documents variables and configurations (file paths) associated with an analysis. For more information about the command line tool see Appendix L.

8.4   Frequently Asked Questions

Can I reorder the data layers?

Yes, data layers may be dragged within the GIS table of contents to reorder them within a group.

How do I export shapefiles?

When viewing any of the displayed maps n the **GIS Map** tab (lower right frame of the main window), click on the GIS toolbar icon for *Save Shapefile* (looks like a 3.5-inch diskette). Follow the prompts to provide a name and location for the file. BenMAP-CE will export a set of files (*.dbf*, *.prj*, *.shp*, *.shx*) associated with the shapefile that you can use with any GIS viewer.

How do I save a GIS map as an image?

To export the map as an image, click the *Export map* image icon (immediately below the *Save Shapefile* icon). This will use built-in DotSpatial GIS tools to allow you to save the map as a formatted image (*.png*) file. Alternately, use the *Print Screen* (PrtScn) button on your keyboard to create an image (saved in memory) which you can then paste into a graphics editor or document.

Can I display my map using a projection other than GCS-NAD 83 or Albers?

Yes, click the GIS toolbar icon for "**change projection to…**"; this feature will allow you to alternate between GCS-NAD83 and Albers projections. For more options, right-click on the **Map Layers** feature in the GIS table of contents and select **Projection** from the pop-up menu. This will display information about the current projection. Click the **Change Projection** button to view and apply other available projections.

Chapter 9
Tools Menu

In this chapter…

- **Learn about the options in the Tools menu.**

Chapter 9 Table of Contents

The **Tools** menu, available on the main BenMAP-CE screen, provides access to six special add-on tools and an **Options** menu. Below we summarize the purpose of each tool. Note that other sections of the manual have already covered several of these items, so we merely list them here and point you to the appropriate section.
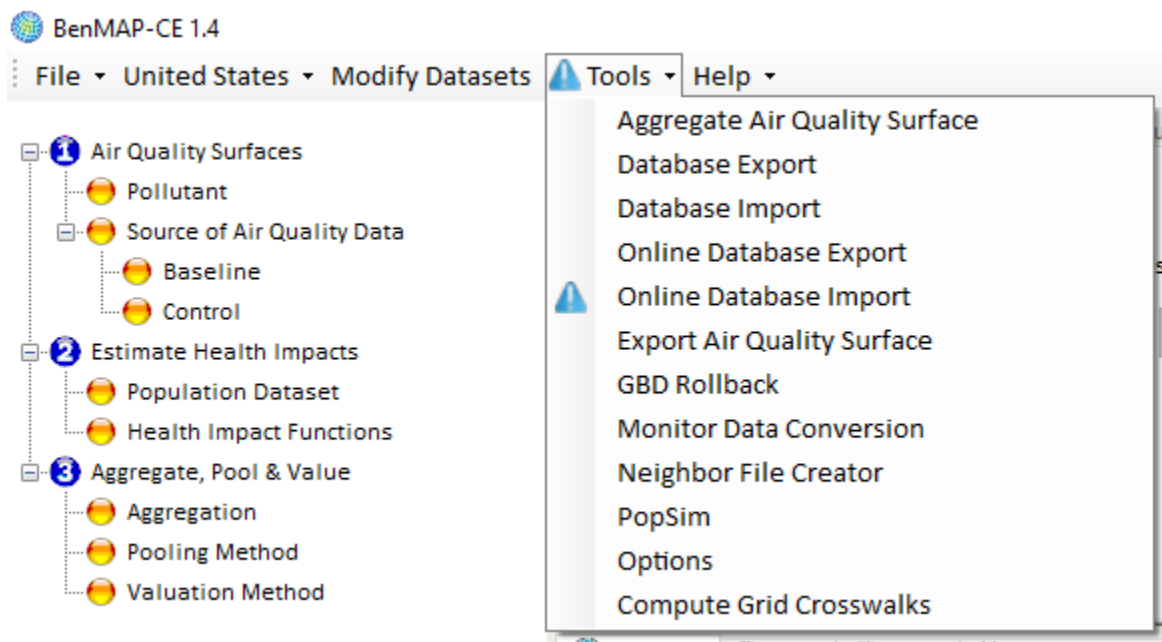
- **Aggregate Air Quality Surface.** Use this tool to change an air quality grid based on one grid definition to a coarser grid definition, using a simple spatially weighted average approach. For example—you could aggregate your air quality surface from 12km × 12km grids to U.S. Counties.

- **Database Export.** Export your entire database (every item in your setup) or parts of your setup (e.g. all GridDefinitions, or individual Health Impact Function Datasets) to a specified file location.

- **Database Import.** Import an entire setup or parts of an individual setup from a specified file location.

- **Online Database Export.** Export your BenMAP-CE dataset(s) and upload to a cloud-based data archive to share with the BenMAP-CE community.

- **Online Database Import.** Import dataset(s) provided by the BenMAP-CE community from a shared cloud-based data archive.

- **Export Air Quality Surface.** Generate a data file that contains all of the air quality values saved within an air quality surface (.aqgx) file.

- **GBD Rollback.** This application estimates the air pollution-attributable health burden, and the benefits of improved air quality, in each country using data from the Global Burden of Disease (GBD) study.

- **Monitor Data Conversion.** Convert monitor data files into BenMAP-CE format.

- **Neighbor File Creator.** Identify the monitors and weights used in the interpolation process when creating air quality grids.

- **PopSim.** Simulate the cumulative effects of air pollution on different age groups over time (U.S. data only).

- **Options.** View and edit the general options for BenMAP-CE.

- **Compute Grid Crosswalks.** Remove or re-create crosswalks between grid definitions.

9.1 Air Quality Surface Aggregation

Using the **Aggregate Air Quality Surface** tool, you can change an air quality grid based on one grid definition to another (coarser) grid definition, using a simple spatially weighted average approach.

To start, choose *Aggregate Air Quality Surface* from the **Tools** drop-down menu. This will bring up the **Aggregate Air Quality Surface** window. Click the **Browse** button to find the air quality grid (.aqgx file) that you want to change and then use the **Aggregation Surface** drop-down list to select the new grid definition that you want to use. For example, you might want to aggregate a 12km model grid to the county-level. Click **OK** when done.
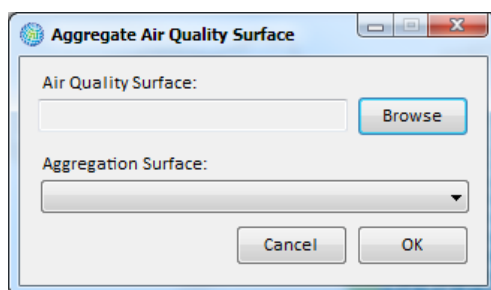
This will bring up the **Save Aggregated Air Quality Grid** window, where you specify the name of the newly aggregated air quality surface (.aqgx file) you are creating and its location. After the file is created, BenMAP-CE will return you to the main BenMAP-CE screen. You can then use the new file just as you would any other air quality surface.

Below is an example of an air quality surface for $PM_{2.5}$ created by the CMAQ model using a 12km grid. (The following images are cropped from the main BenMAP-CE window to focus on the GIS panel.)
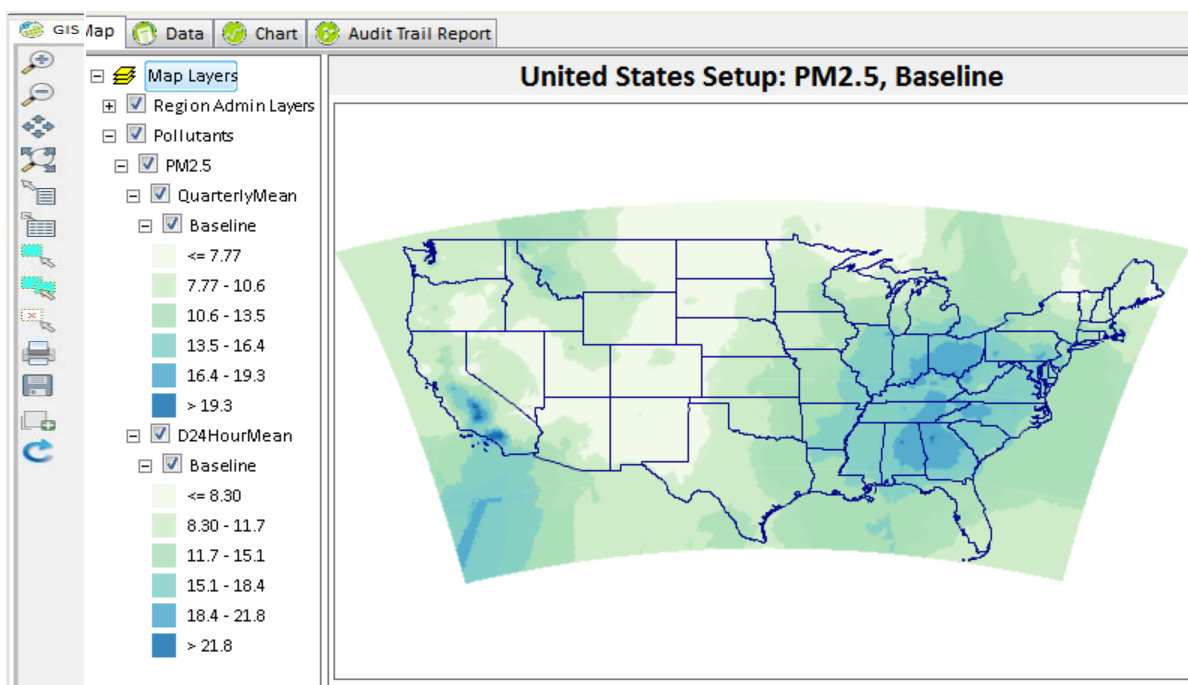
Here is an example of the above 12km air quality surface aggregated to the county level.
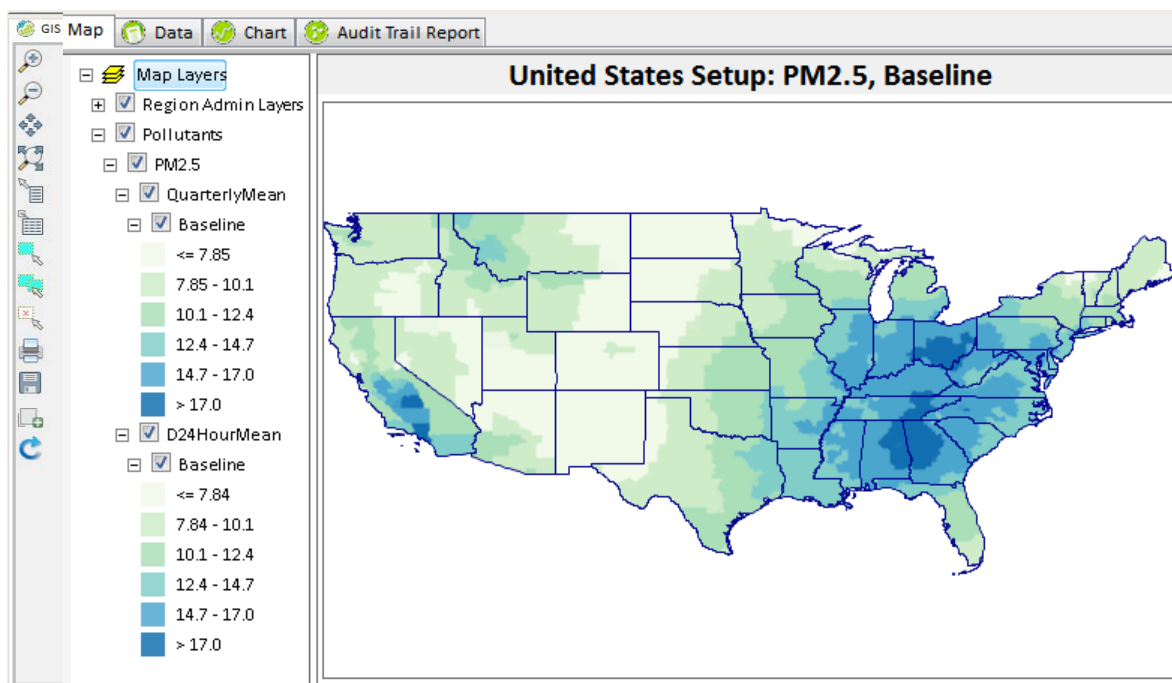
And, here is an example of the 12km air quality surface aggregated to the state level.
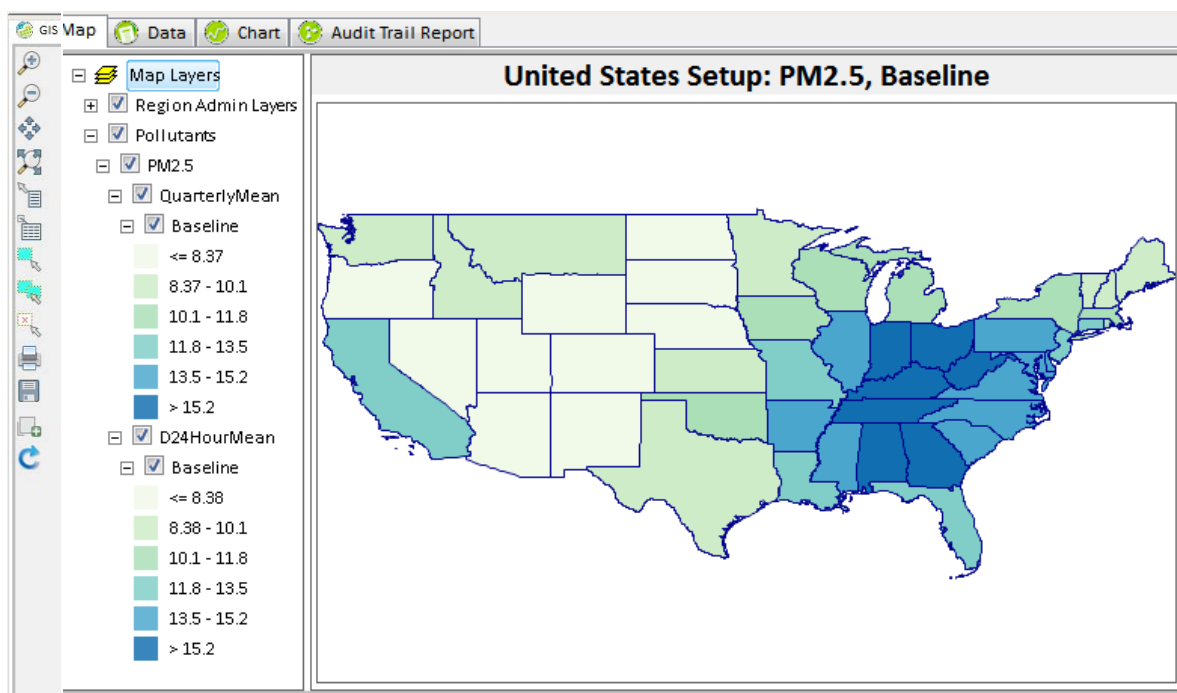
Observe that there is more variation in the county-level file than the state-level file. This is expected, because BenMAP-CE is just using a simple spatially weighted average of the data.

9.2   Database Export

BenMAP-CE allows you to export and import entire databases (all **Available Setups**), individual setups (e.g., *United States, China*), and parts of individual setups (e.g. all **Grid Definitions**, or individual **Health Impact Function** datasets). This functionality can be used to archive data, share data with other BenMAP-CE users, move databases between computers, as well as to view data in other applications like Excel or ArcMap. In particular, all of the steps involved in creating a setup can be done just once, after which the data can be exported and then imported on other computers. Version 1.4 also allows users to import databases exported from Version 1.3. This can be useful for transferring existing setups to the new version of BenMAP-CE without needing to re-recreate them.

To export part or all of an existing setup, go to the **Tools** menu, and choose the *Database Export* option.

This will bring up the **Database Export** window. Initially, all of the setups are listed in a tree menu, which is initially in a collapsed view. To expand any of the menu items, click on the '+' sign to the left of menu item. This will expand the tree menu to show additional listings for the expanded item. To collapse the tree menu, simply click on the '-' sign.

Choose a dataset to export by selecting it from the tree menu. The default Type of Export is BenMAP CE database. Press **OK**. In the screenshot below, we have chosen to export *EPA Standard Monitors O3* to a BenMap-CE database file.
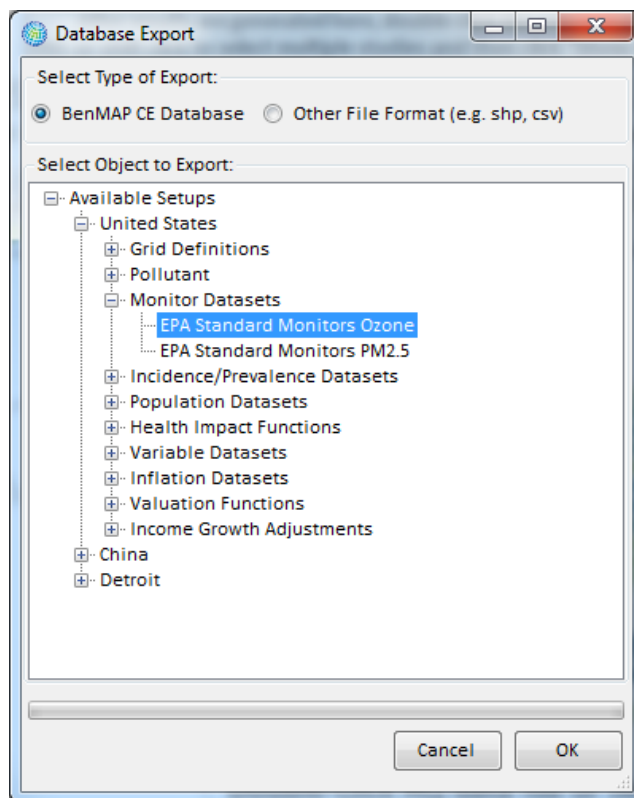
This will bring up the **Save As** window. From here you may name the export file, select the export file type and choose the save location.
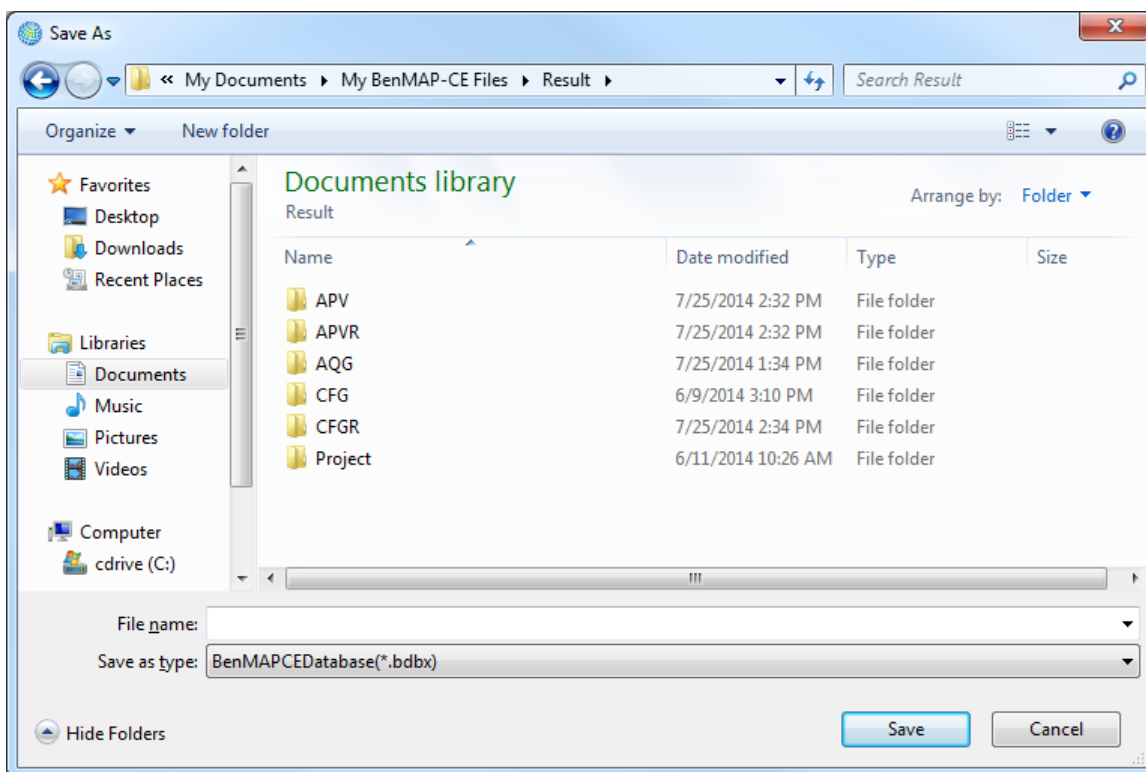
NOTE: Exported BenMAP-CE database files have a *.bdbx* extension, and are a binary format not suitable for viewing in external applications.

To view exported data in external application, choose "Other file Format". The tool will export Grid definitions in shapefile (*.shp*) format and the rest of the data in CSV (*.csv*) format. If the dataset contains multiple tables, it will be exported into multiple CSV files. In the screen below, we have chosen to export *EPA Standard Variables* to a CSV file.
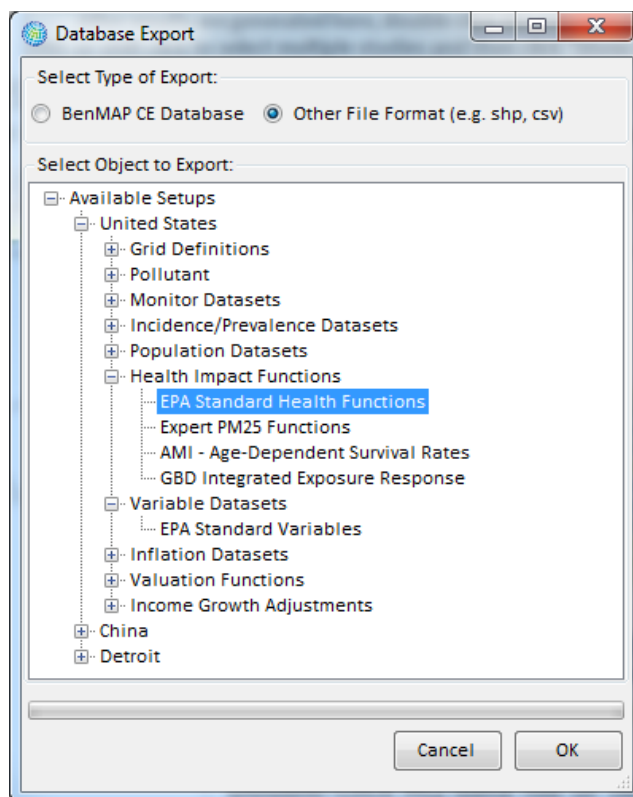
In the **Save As** window, the File name has been pre-populated with dataset name. You may leave it as is or change to something you want.
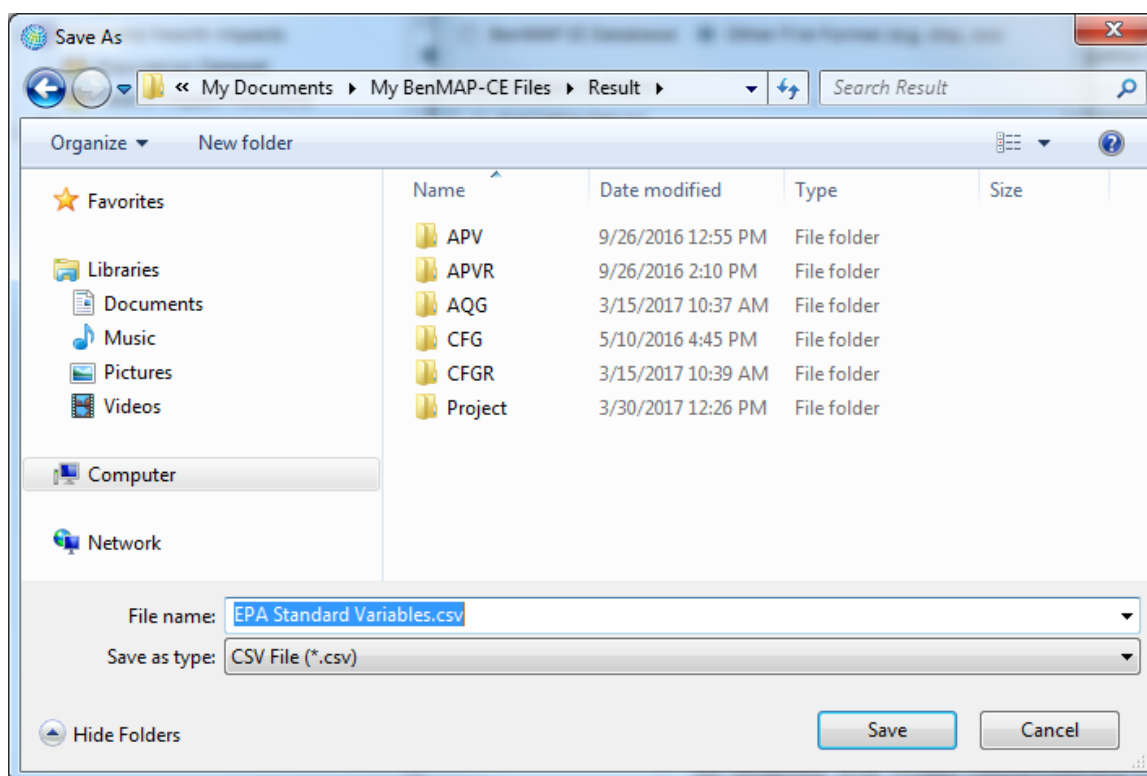
In this case, each variable of *EPA Standard Variables* dataset is saved as one CSV file with its variable name appended to the end of the dataset name.
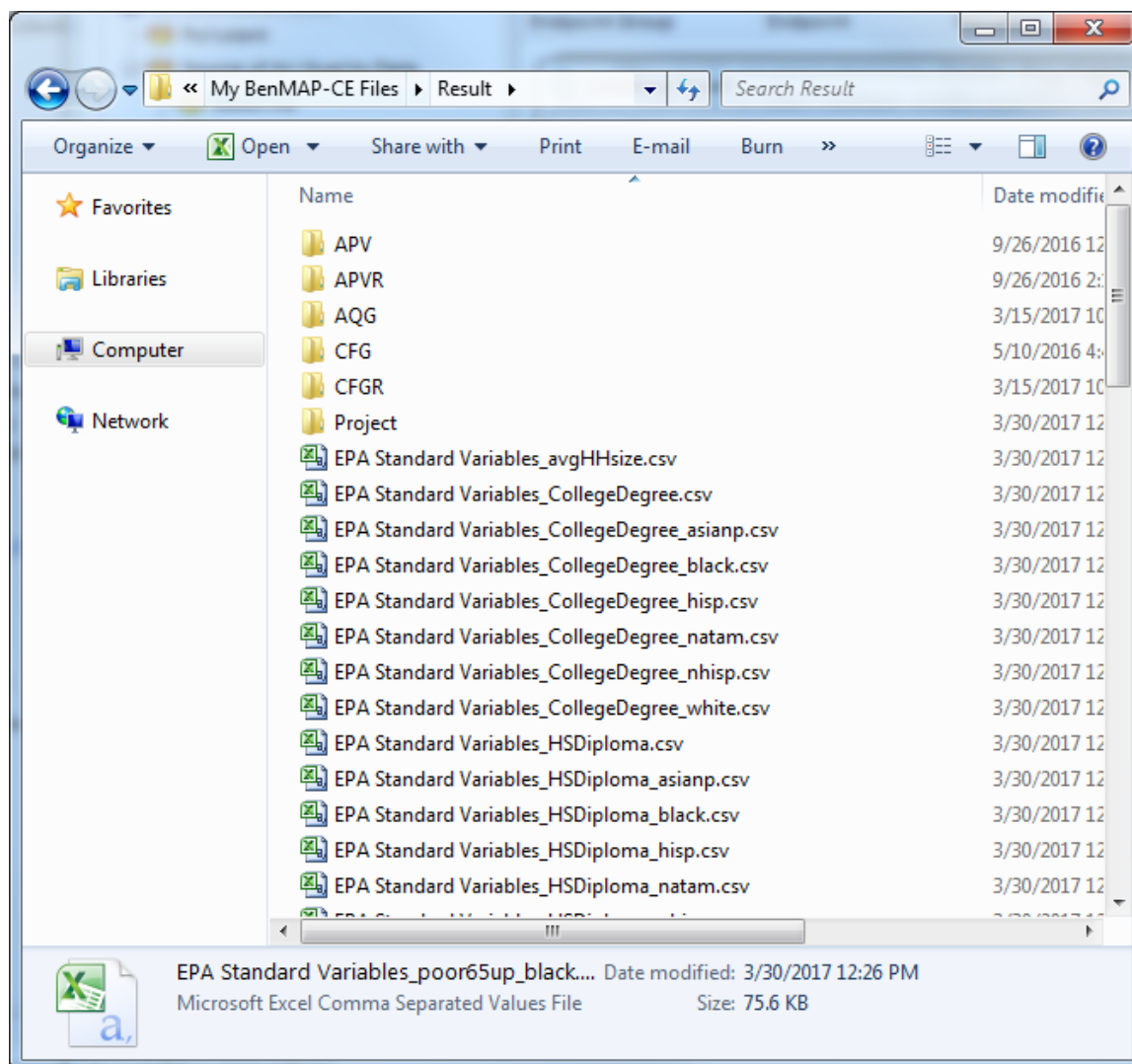
NOTE: When exporting datasets in Other File Format, only one dataset can be selected at one time.

9.3   Database Import

Import entire setups or parts of individual setups. This option is described in the Import Setups section of Chapter 4: Loading Data.

To import part or all of an existing setup, go to the **Tools** menu, and choose the *Database Import* option. This will bring up the **Database Import** window.
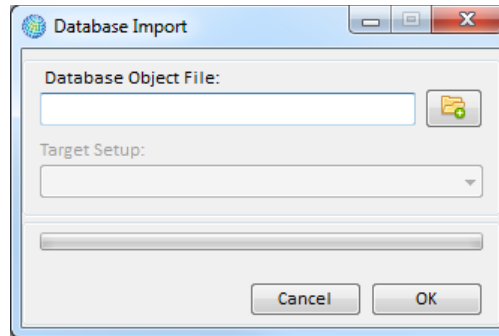
The **Database Object File** identifies the file that you want to import. Click on the **Browse** icon to locate the file. This will display the **Open** window.

Find and select the *.bdbx* file that you want to import, and then click **Open**. This will return you to the **Database Import** window. Click **OK** to finish the import process.

If the import file contains a subset of a setup, such as a collection (e.g., a set of Grid Definitions) or an individual dataset (e.g., a single grid definition from among many available), select the **Setup** into which it should be imported from the **Target Setup** drop-down list. Click **OK** to finish.

NOTE: Duplicates of datasets (typically identified by their names, e.g., '*Detroit Population*') will default to the existing dataset in the Setup. New datasets (i.e., non-duplicated) will be added to the setup.

9.4   Online Database Export

The **Online Database Export** feature was designed to facilitate sharing of BenMAP-CE datasets among the user community. The data is stored in an online database. Sharing and use of this data is at the discretion of the user community.

To share data in the online repository, select *Online Database Export* from the BenMAP-CE **Tools** menu. In the **EPRI Online Database Export** form, provide your name (required field), organization, and a description of the data you are sharing. Then select the object you wish to export from the tree menu. Similar to the database export feature (see Section 9.2), you can select an entire setup, dataset type, or specific data element. Once you make your selection and click "**OK**", you will be prompted to confirm the upload.
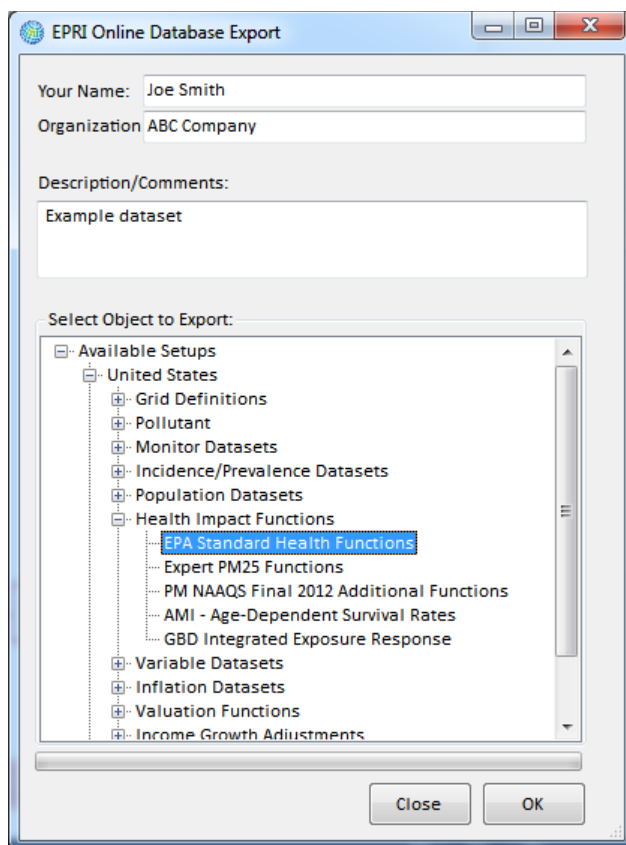
9.5 Online Database Import

The **Online Database Import** feature was designed to provide user access to an online repository of BenMAP-CE datasets shared by the user community. To import data from the online repository, select *Online Database Import* from the BenMAP-CE **Tools** menu. Select the data you wish to import by clicking on the record selector on the left side of the data grid. The record will be highlighted. Next choose your **Target Setup** from the dropdown list. Then, click the "**OK**" button. The progress bar near the bottom of the screen will update to show progress of the import. When complete, the system will notify you that the "*The database file was imported successfully."*
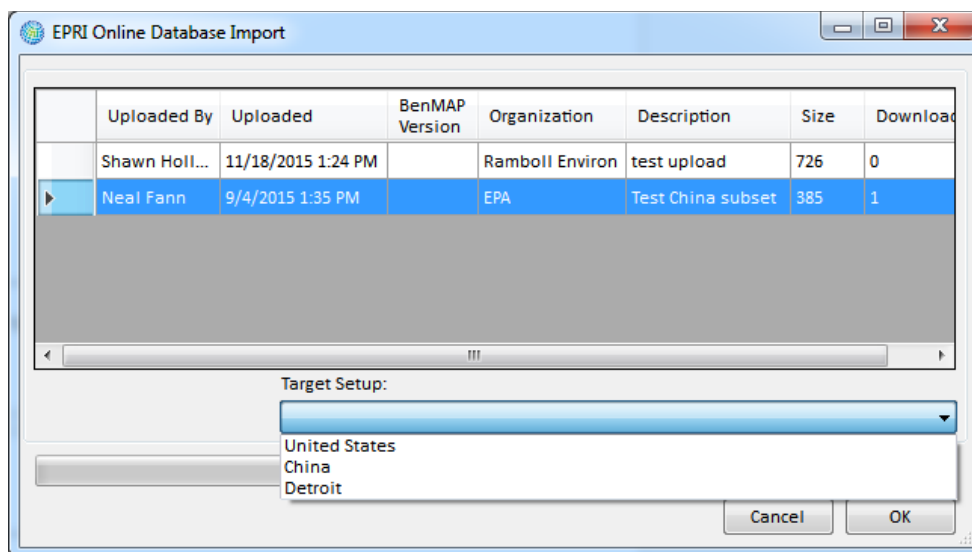
9.6   Export Air Quality Surface

The **Export Air Quality Surface** tool generates a data file (*.csv*) reporting all of the data contained in the air quality grid. After choosing *Export Air Quality Surface* from the **Tools** menu, the **Export Air Quality Surface** window will appear. Click the **Browse** button to choose the air quality grid that you want to examine.
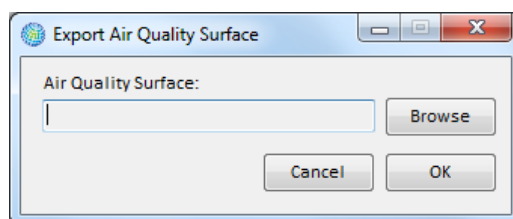
Click **OK** after you have selected your file. Use the **Save As** window to choose the directory where you want to save your file. And in the **File name** box, type in the name of the file.

To help keep track of what you are doing, you might want to use the same file name as your air quality grid, or something very similar. (If you use the same name, you can always distinguish the two files by the extension. An air quality grid has an *.aqgx* extension and the file you are generating here has a *.csv* extension.)

When done, click the **Save** button. You can view the files you have created with any database viewer. For each **Metric** and **Seasonal Metric**, you can see the actual values. In addition, you can see the **Statistics** calculated for each. In the example below for a 12km PM$_{2.5}$ air quality surface, you can see in the first grid cell (Column = 1, Row = 246) that the mean of the *D24HourMean* values is 9.15 and the mean of the *QuarterlyMean* values is 9.2.

Col	Row	D24HourMean	QuarterlyMean
1	246	9.15	9.2
2	246	9.16	9.2
3	246	9.16	9.21
4	246	9.17	9.21
5	246	9.17	9.22
6	246	9.54	9.65

Note that the exported files may be very large (tens to hundreds of megabytes in file size and with row counts exceeding typical spreadsheet applications). With large files, you might need to use a database program to work with the files. Alternatively, these files can also be read by simple text editors.

9.7   GBD Rollback

The World Health Organization global burden of disease (GBD) study measures burden of disease using the disability-adjusted-life-year (DALY). This time-based measure combines years of life lost due to premature mortality and years of life lost due to time lived in states of less than full health. The DALY metric was developed in the original GBD 1990 study to assess the burden of disease consistently across diseases, risk factors and regions.[1]

The **GBD Rollback** tool uses data from the 2013 GBD study to allow users to estimate the human health burden of $PM_{2.5}$ levels in each country as well as the benefits of reducing these air pollution levels. Users can "roll back," or adjust ambient $PM_{2.5}$ levels in one or more countries or regions and calculate the total burden, or avoided deaths, in that region. The tool also estimates avoided life years lost, changes in life expectancy, and the economic benefits associated with avoided deaths.

This feature is analogous to the monitor roll-back tool already available in core BenMAP-CE (discussed in Chapter 5); that tool allows users to adjust downward (or upward) air quality monitoring data in specified locations according to various algorithms (including proportional, quadratic and incremental rollbacks). The tool uses a grid with 0.1 degree resolution (approximately 10km) grid cells and includes:

- 2015 $PM_{2.5}$ pollution concentrations[2]. Negative concentrations were adjusted to zero.
- 2015 global population[3] data stratified by age and gender. Elder populations were combined into an "80UP" age group to align with the incidence dataset.

---

[1] For more information on the GBD, see: http://www.who.int/topics/global_burden_of_disease.

[2] Cohen, A. J.; Brauer, M.; Burnett, R. T. (2017). Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet, 389(10082), 1907-18.

- 2013 mortality incidence[4] for six health endpoints (COPD, cerebrovascular disease, ischemic heart disease, lung cancer, acute lower respiratory infection, and non-accidental) stratified by age and gender. Neonatal ("0 to 0") and "1 to 4" age groups were combined into a "0 to 4" age group to align with population data.

- The 2013 Integrated Exposure Response (IER) function[5] employed by the 2013 GBD study to estimate premature mortality associated with ambient air pollution. The 2013 IER function estimates premature mortality from COPD, cerebrovascular disease, ischemic heart disease, and lung cancer.

- The Shape-Constrained Health Impact Function (SCHIF) developed by Burnett et al. (in preparation)[6]. Like the IER, the SCHIF is a meta-analytic concentration-response function developed using data from many $PM_{2.5}$-related epidemiological studies. Depending on the country being evaluated, the SCHIF estimates draw from mortality incidence from either (a) all non-accidental causes, or (b) a "re-attributed" incidence rate representing deaths due to COPD, cerebrovascular disease, ischemic heart disease, lung cancer, and acute lower respiratory infection).

---

[3] Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Administrative Unit Center Points with Population Estimates. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4F47M2C.

[4] Global Burden of Disease Study 2013. Global Burden of Disease Study 2013 (GBD 2013) Incidence, Prevalence, and Years Lived with Disability 1990-2013. Seattle, United States: Institute for Health Metrics and Evaluation (IHME), 2015. http://ghdx.healthdata.org/record/global-burden-disease-study-2013-gbd-2013-incidence-prevalence-and-years-lived-disability

[5] Burnett RT, Pope CA, Ezzati M, Olives C, Lim SS, Mehta S, Shin HH, Singh G, Hubbell B, Brauer M, Anderson HR, Smith KR, Balmes JR, Bruce NG, Kan H, Laden F, Pruss-Ustun A, Turner MC, Gapstur SM, Diver WR, Cohen A. (2014). An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environ Health Perspect 122:397-403; http://dx.doi.org/10.1289/ehp.1307049

[6] Burnett RT, Chen H, Szyszkowicz M, Fann N, Hubbell B, Pope CA, Apte JS, Brauer M, Cohen A, Weichenthal S, Coggins J, Di Q, Brunekreef B, Frostad J, Lim SS, Kan H, Pruss-Ustun AM, AARP collaborators, ACS collaborators, CTS collaborators, Canadian Breast Screening collaborators, CanCHEC/CCHS collaborators, Chinese Male Cohort Collaborators, CUELS collaborators, English Cohort collaborators, Hong Kong collaborators, NHIS collaborators, NHS collaborators, Rome Census Cohort collaborators, VHM&PP collaborators. 2017. A new approach to estimating global mortality burden from outdoor fine particle exposure. Manuscript in preparation.

To begin, select *GBD Rollback* from the **Tools** menu. The **GBD Rollback Tool** window will be displayed.
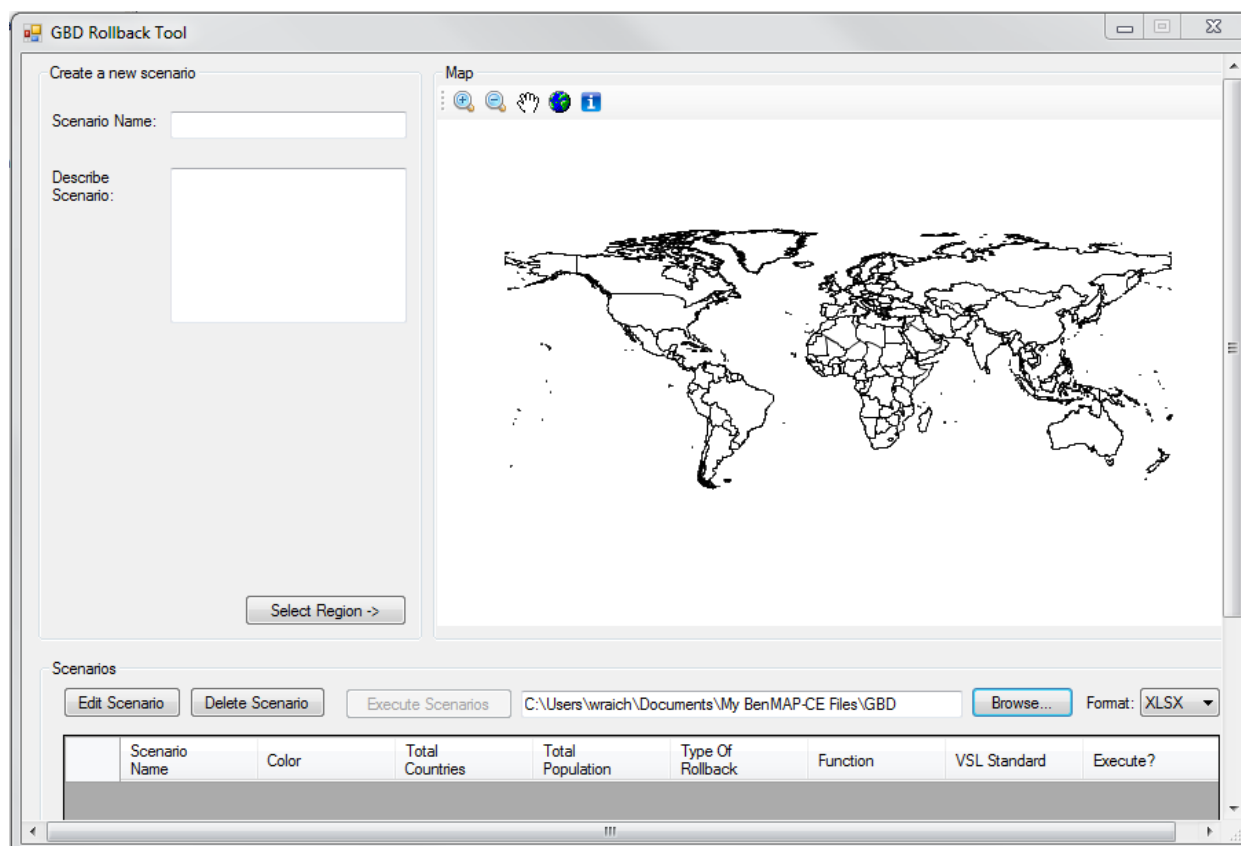
9.7.1 Create New Scenario

To create a new scenario, first provide a **Scenario Name** (required) and description for the scenario (optional). Then, click the **Select Region** button.

9.7.2 Select Regions or Countries

The **Region Selection** box provides options to view *Regions* (as defined in the GBD study) with the associated countries listed beneath, or to view *Countries*. Multiple selections are possible with either selection method and you can toggle between the two.

If you view *Regions*, you can select an entire region (all countries within this region) by checking the box next to the desired region. To view the individual countries for a region, click on the '+' sign to the left of the region. This will expand the tree menu. To collapse the tree menu, simply click on the '-' sign.

If you do not know the name of the region which contains the country you want to select, click the option to view *Countries*. You can search the alphabetically-sorted list to find countries.

As you make selections, the map will highlight them in a bright blue color. Note that the map is not interactive (i.e., you cannot click on the map to make selections). The toolbar located above the map allows you to zoom in and out, pan, view full extent, and identify countries.
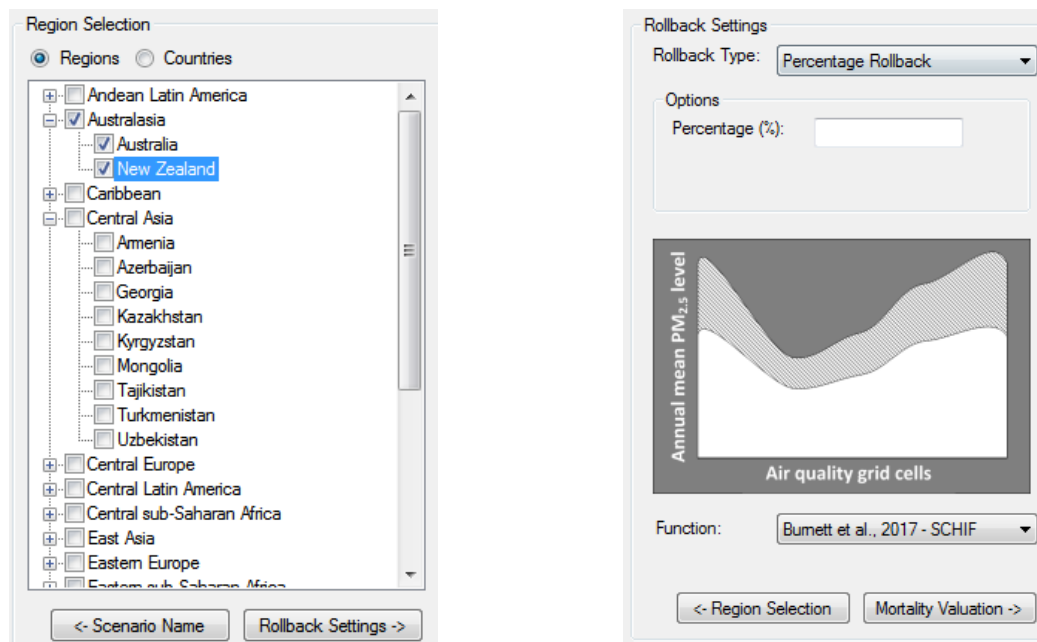
9.7.3   Choose Rollback Settings

After selecting the countries or regions to analyze, click the **Rollback Settings** button. The **Rollback Settings** pane allows you to choose the **Rollback Type** (*Percentage Rollback*, *Incremental Rollback*, or *Rollback to a Standard*) and configuration options.[7] (An illustrative example for the selected **Rollback Type** will be displayed below the *Options* box.) The tool allows for a negative rollback, indicating an increase in pollution concentration.

For *Percentage* and *Incremental Rollback* types, you must enter a percent value or whole number by which to reduce the pollutant concentration in all grid cells in the selected countries. If you select *Rollback to a Standard*, you may select from a list of national pollutant standards enforced by various countries in different years[8].

Standard Group	Concentration Limit (µg/m³)	Exposure Duration
WHO	10	Annual
US Primary	12	Annual
US Secondary	15	Annual

---

[7] The tool uses a background PM$_{2.5}$ concentration = 5.8 µg/m³ (lowest measured level in epidemiological literature).

[8] Air Quality Standards / Guidelines / Objectives For Different Countries (updated April 2014). Table developed by Scott Voorhees, US EPA.

EU	25	Annual
Japan	15	Annual
China Class I	15	Annual
China Class II	35	Annual
Hong Kong	35	Annual
India	40	Annual
Singapore	12	Annual
WHO	25	24-hour
US Primary	35	24-hour
Japan	35	24-hour
China Class I	35	24-hour
China Class II	75	24-hour
Hong Kong	75	24-hour
India	60	24-hour
Singapore	37.5	24-hour

The **Function** dropdown allows users to select either the SCHIF (Burnett et al., 2017) or the IER function used in the 2013 GBD studies (Burnett et al., 2014).

9.7.4 Choose Mortality Valuation Settings

Once you have completed the Rollback Settings window, click the **Mortality Valuation** button. The **Mortality Valuation Settings** pane allows you to specify how you wish to calculate the economic benefits associated with avoided premature deaths.

Estimated reductions in premature mortality are valued using country-specific estimates of the value per statistical life (VSL). This value is not the monetary value of individual lives. Rather, it reflects the amount individuals are willing to pay to incrementally reduce their risks of death from adverse health conditions that may be caused by environmental pollution. For example, if each individual in a population is willing to pay \$20 to reduce his or her risk of death by 10 in 100,000, the VSL for that population would be \$200,000 (= \$20 ÷ (10/100,000)). For 500 avoided deaths in that population, the economic benefit would be \$100,000,000 (= 500 * \$200,000).

The **VSL Standard** dropdown menu allows users to select the VSL estimate applied to mortality risk reductions. Because VSL estimates from primary research are not available in most countries, it is necessary to transfer international estimates across countries, adjusting for differences in income levels.[9] Users may select one of two base

---

[9] Mortality risk reductions have been shown to be sensitive to income; individuals are willing to pay more for risk reductions as their income increases. Thus, the VSL is commonly adjusted to reflect differences in the average income levels both across countries and within a country over time.

VSL estimates generously provided by the authors of World Bank (2016).[10] VSL estimates were derived by the authors using a benefit-transfer approach, with a base VSL of $3.83 million (2011 USD) as recommended by a meta-analysis of willingness-to-pay (WTP) studies by the Organisation for Economic Co-operation and Development (OECD). The base VSL is transferred to other countries and years using the following formula:

$$VSL_{c,n} = VSL_{OECD} * \left(\frac{Y_{c,n}}{Y_{OECD}}\right)^e$$

where $VSL_{c,n}$ is the VSL for country c in year n; VSL_{OECD} is the base OECD VSL; $Y_{c,n}$ is the GDP per capita for country c in year n; Y_{OECD} is the average GDP per capita for the sample of OECD countries (roughly $37,000); and e is the income elasticity of the VSL. In all instances, monetary values are adjusted for price inflation and converted to 2011 U.S. dollars at purchasing power parity (PPP) rates. While the authors provide VSL estimates for a range of income elasticity (e) values, we present their central results, which assume an elasticity of 1.2 for low- and middle-income countries, and 0.8 for high-income countries. This benefits transfer approach is also used with a base VSL estimate from USEPA of $8.7 million (2011 USD).

Additionally, we use country-level VSL estimates to derive population-weighted average VSLs by GBD regions. These VSLs allow GBD Tool users to value reduced mortality at the regional level. The regional VSLs are also used to impute missing country-level VSLs for 29 of the 210 countries in the GBD tool. The "Metadata" tab in the GBD Tool Output spreadsheet displays VSL estimates by country and region.

The left pane of the **Mortality Valuation Settings** window provides a list of countries with their respective VSL estimates. Once you have selected the base VSL estimate (OECD or USEPA), click the **Save Scenario** button to save your rollback configuration. The map will update the color of the selected countries so that each saved scenario is uniquely different.

---

[10] World Bank and IHME (2016). *The Cost of Air Pollution: Strengthening the Economic Case for Action.* Washington, DC: 2016 International Bank for Reconstruction and Development/The World Bank.
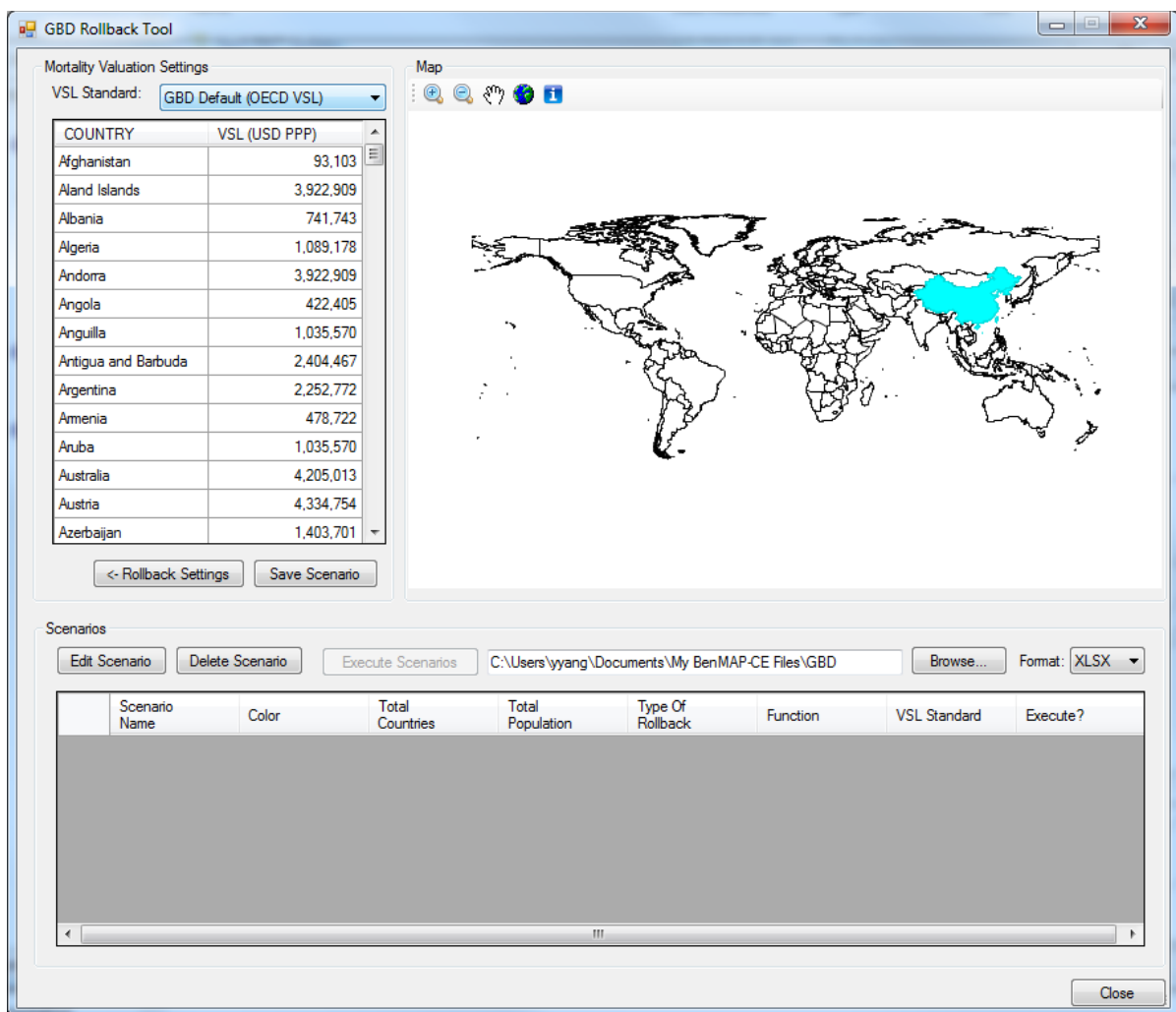
The table at the bottom of the window will update to display the attributes for the saved scenario (*Scenario Name, Color, Total Countries, Total Population, Type of Rollback, etc.*), along with an *Execute?* option for you to indicate whether the tool should execute or ignore the scenario. If you double-click in the Total Countries or Total Population fields, the program will display a table of the individual countries and populations included in the scenario.
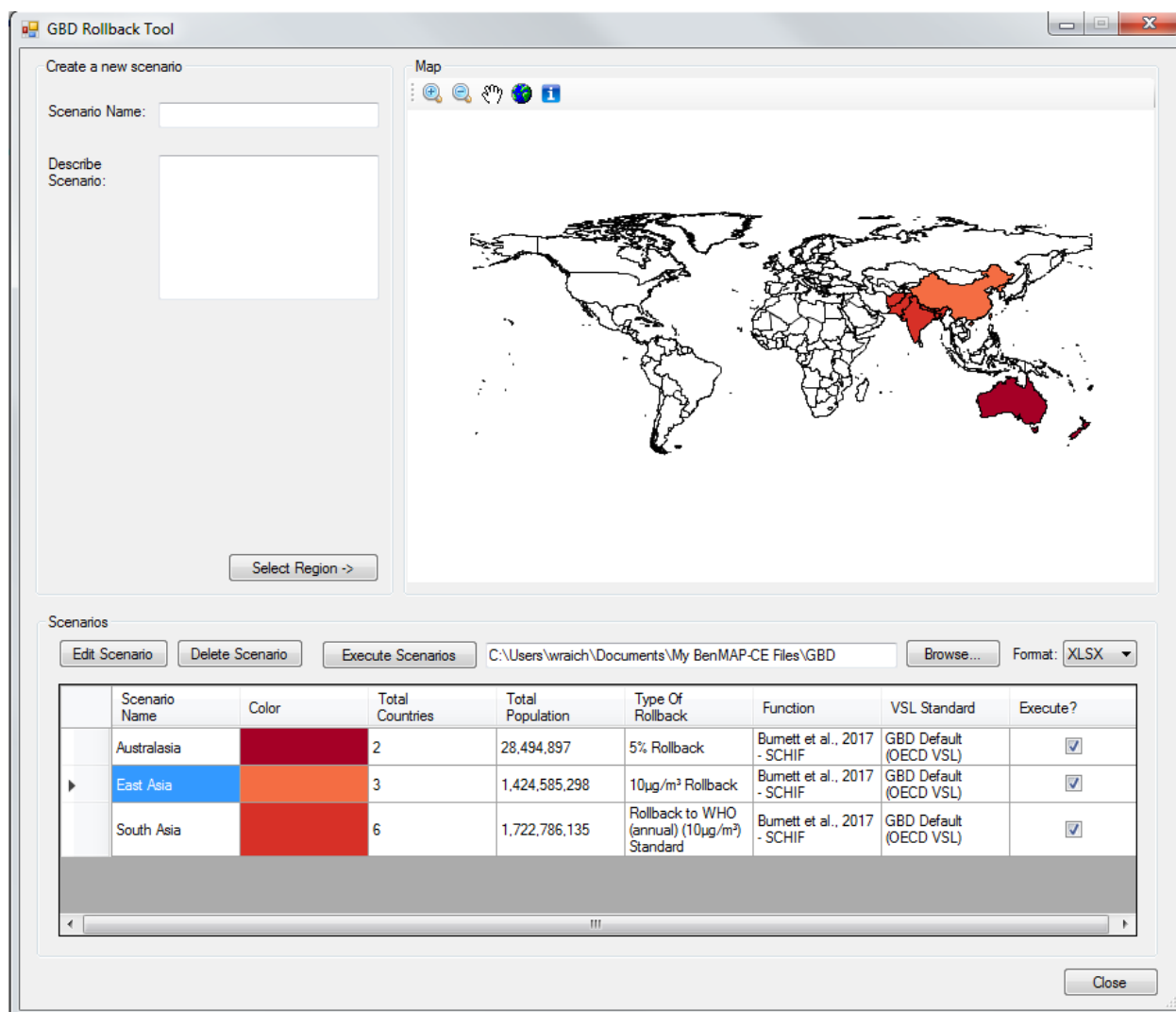
9.7.5 Execute Scenarios and Save Results

Click the **Execute Scenarios** button to generate results for saved scenarios. If a selected country does not have sufficient data (i.e., population, pollution concentration, or incidence rates) to perform the analysis, it will be bypassed and a message (including the country name) will be provided.

If you are adding to a list of saved scenarios and do not want to re-run the previously executed ones, uncheck *Execute?* in the summary table so that only the new scenario is checked.

Upon execution, the **GBD Rollback** tool will export results in (*.xlsx* format) to a default file location.[11] You can change the file location by clicking the **Browse** button. You can also change the export file format to *CSV* by clicking on the **Format** drop-down list. The results file is named using the *Scenario Name* followed by the time stamp of scenario execution.

- The *XLSX* option provides a formatted summary table, detailed results, charts, and metadata about the supporting datasets used in this analysis.

- The *CSV* option will only create two files: one containing the summary data and one with detailed results.

If you want to edit or delete saved scenarios, select one record at a time from the scenario table (use the record selector on the left side of the grid) and click the **Edit Scenario** or **Delete Scenario** button. If a scenario is edited and re-executed, the timestamp in the filename will help the user identify the new version.

Once you close the **GBD Rollback** tool, the scenarios are cleared from memory; scenarios are not saved in the BenMAP-CE database. Information about the scenario configurations are saved in the results file to help you document the analysis and re-create it if necessary.
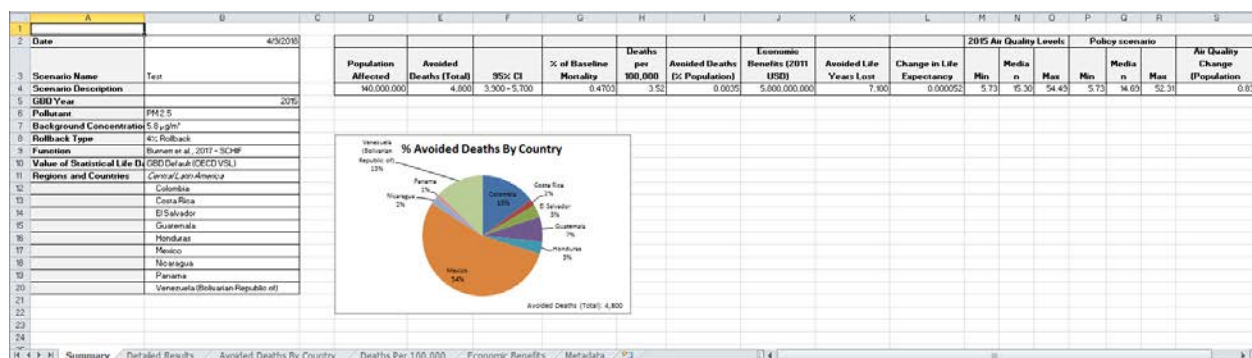
The GBD results file (*.xlsx* format) contains six worksheet tabs, described as follows:

- **Summary**: Gives a basic background on the scenario chosen, including name, description, pollutants, rollback type, health impact function, VSL selection, and countries. It also provides a quick overview of total and affected population results.

- **Detailed Results**: Gives in-depth breakdown of population and results for each selected country. Important fields on this tab include *Population Affected*, *Avoided*

---

[11] The files are stored by default at the following location: C:\Users\<User Name>\Documents\My BenMAP-CE Files\GBD.

Deaths (Total), Economic Benefits, Avoided Life Years Lost, and *Change in Life Expectancy*.

- **Avoided Deaths by Country**: Displays the number of avoided deaths for the region and by country in a bar graph. This graph was generated from the data on the **Detailed Results** tab.

- **Deaths Per 100,000**: Displays the number of deaths for each country (with a population multiplier of 100,000) in a bar graph. This graph was generated from the data on the **Detailed Results** tab.

- **Economic Benefits**: Displays the economic benefits associated with avoided premature mortality for the region and each country in a bar graph. This graph was generated from the data on the **Detailed Results** tab.

- **Metadata:** Provides supplementary information about the underlying data and functions used by the **GBD Rollback** tool.

9.8 Monitor Data Conversion

The **Monitor Data Conversion Tool** provides the user with a mechanism to transform their data from commonly available formats (e.g., one monitoring result per row) to BenMAP-ready format.
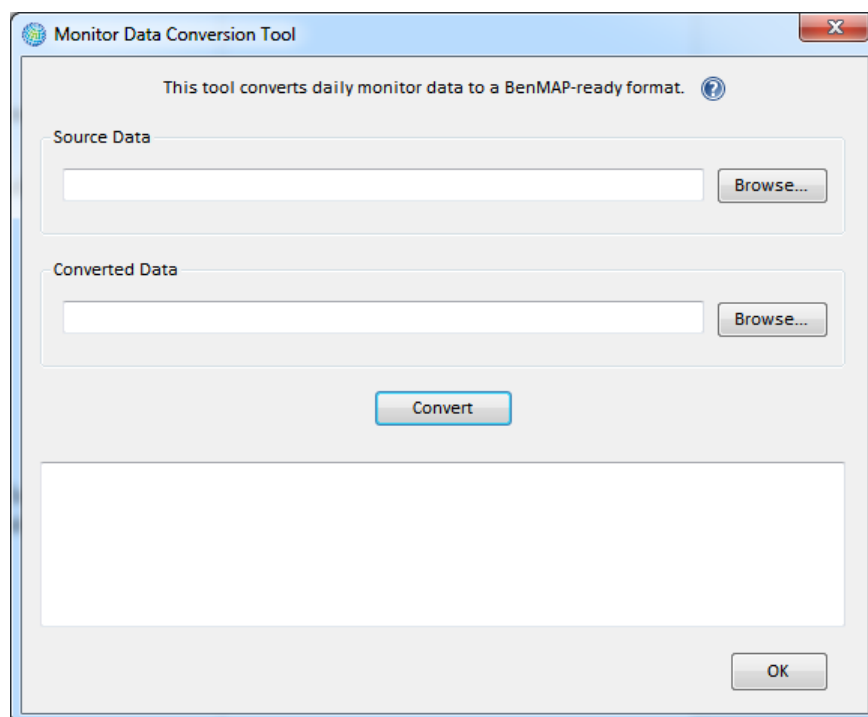
To view an example, click the help (?) button near the top of the form.
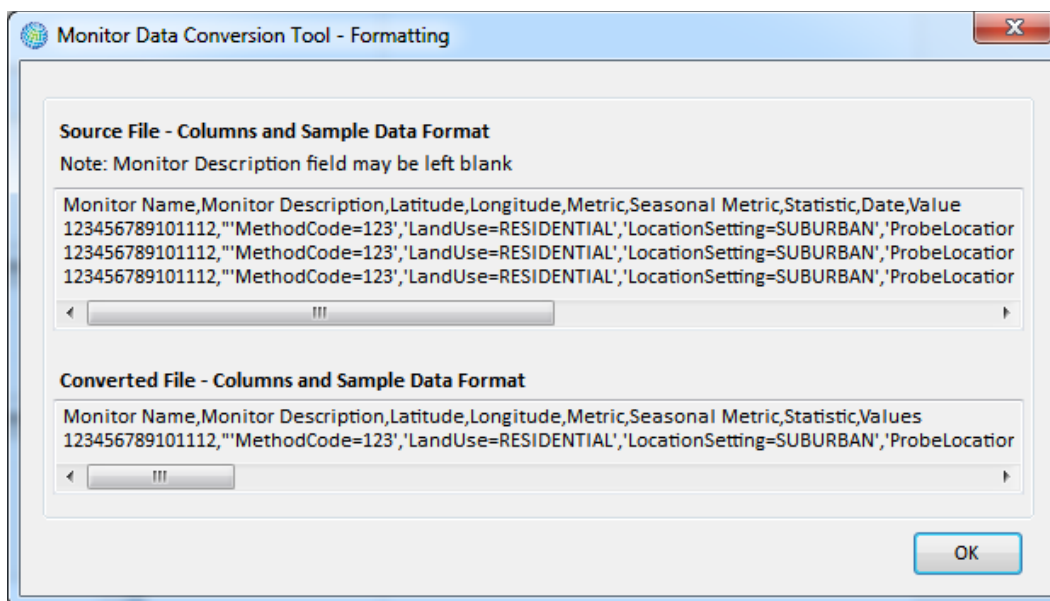
To use this tool, first click "**Browse…**" to locate your source data file. Then click "**Browse…**" to specify the file name and location for the converted file. Click the "**Convert**" button to begin the conversion process. Once the file is converted, click "**OK**" to close the form.

9.9   Neighbor File Creator

The **Neighbor File Creator** tool generates a file containing gridded monitor data created by BenMAP-CE for *Monitor Data* or *Monitor Rollback* air quality surfaces.

To start, choose *Neighbor File Creator* from the **Tools** drop-down menu. This will bring up the **Create Neighbors File** window. Click on the **Browse** button, and find the air quality surface you want to analyze.
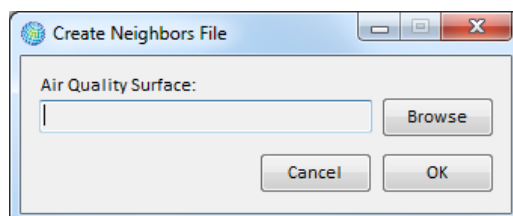
After locating the file, click the **OK** button. The file path for the selected file will be displayed in the **Air Quality Surface** box. Click **OK**. A **Save As** window will open. Provide a file name for the **Neighbors file (.csv)** you want to create and click **Save**. An example Neighbors File is shown below.

Col	Row	MonitorName	Weight	Distance
1	246	410030013881011	0.215167	627.5258
1	246	530410006881011	0.253047	533.5881
1	246	530090009881011	0.286508	471.2713
2	246	410030013881011	0.213318	620.813
2	246	530410006881011	0.252881	523.688
2	246	530090009881011	0.287978	459.8632
3	246	410030013881011	0.211387	614.2613
3	246	530410006881011	0.252682	513.8765
3	246	530090009881011	0.289525	448.4836

The first two columns specify the *Column* and *Row* variables for each grid cell. In the example above, you will see that Column = 1 and Row = 246 are repeated three times, indicating that three different monitors were used to estimate air quality at this grid cell. The *MonitorName* column provides the monitor identifier. The *Weight* column specifies the weight used in the air quality calculation (e.g., *Voronoi Neighbor Averaging*). And the *Distance* column gives the distance (in kilometers) from the monitor location to the center of the grid cell.

Note that if an air quality grid was created using the *Closest Monitor* option (see Chapter 5 under the Monitor Data section), then only a single monitor is used for any given grid cell. As a result, the neighbor file will contain the same 5 fields, but the *Weight* column will contain a value of "1" all the way down. In addition, there will only be one entry for each grid cell.

Col	Row	MonitorName	Weight	Distance
1	246	410030013881011	1	627.5258
2	246	410030013881011	1	620.813
3	246	410030013881011	1	614.2613
4	246	410030013881011	1	607.8755
5	246	410030013881011	1	601.661
6	246	410671003881011	1	548.5376
7	246	410671003881011	1	541.038
8	246	410671003881011	1	533.702

9.10 PopSim

The Population Simulation (PopSim)[12] model was designed to estimate two outputs related to premature mortality attributed to the CAAA: life-years lost, and changes in life expectancy. The population simulation approach provides some advantages over the BenMAP-CE model in terms of simulation of the dynamic effects of mortality across a population through time, but also has several significant disadvantages relative to BenMAP-CE in terms of the spatial resolution of pollutant exposure estimates. As a result, the population simulation approach operates as a supplement to the BenMAP-based primary estimates for selected measures of the impact of reducing risks of premature mortality.

> The core BenMAP-CE program estimates changes in adverse health effects based on changes in air quality for one specified analysis year, even though certain health benefits may occur after the analysis year. Conversely, the PopSim tool estimates the change in population mortality risk over a multi-year period, but it is not yet possible to estimate the economic value of these impacts in BenMAP-CE.

PopSim is designed to track the effect of alternative assumptions about the mortality effects of fine particulate matter ($PM_{2.5}$) air pollution in the U.S. population over time. The tool incorporates detailed life table data for historical years, by age, gender, and cause of death, obtained from the Census Bureau and the Centers for Disease Control (CDC). It also incorporates Census mortality and population projections for future years, again by age and gender, using the projected death and birth rates that underlie the Census Bureau's published population projections.

The PopSim model allows users to:

- Simulate population in the U.S. by single year cohorts of age and gender for years between 1980 and 2050 under alternative assumptions about the degree of hazard posed by air pollution relative to baseline historical and projected Census mortality rates;

- Estimate changes in life years relative to baseline Census mortality rates;

- Apply air pollution hazards differentially by cause of death; and

- Analyze the effect of alternative cessation lag structures on the timing of total mortality and on total life years in the U.S. population, based on differential application by cause of death or other specifications of cessation lag.

The dynamic life-table approach used in this model can theoretically provide improved estimates of the mortality impacts of air pollution in future years over the more common

---

[12] The PopSim tool in BenMAP-CE is based upon an Access-based model originally developed for EPA by Industrial Economics, Inc. This Population Simulation model is described in Chapter 5 of EPA's "Second Prospective Study – 1990 to 2020 – Benefits and Costs of the Clean Air Act. U.S. Environmental Protection Agency, Office of Air and Radiation, April 2011."

static approach because it explicitly accounts for the year-to-year cascade of impacts on mortality and population following an air pollution change.

To use the PopSim tool, first select **PopSim** from the BenMAP-CE **Tools** menu.
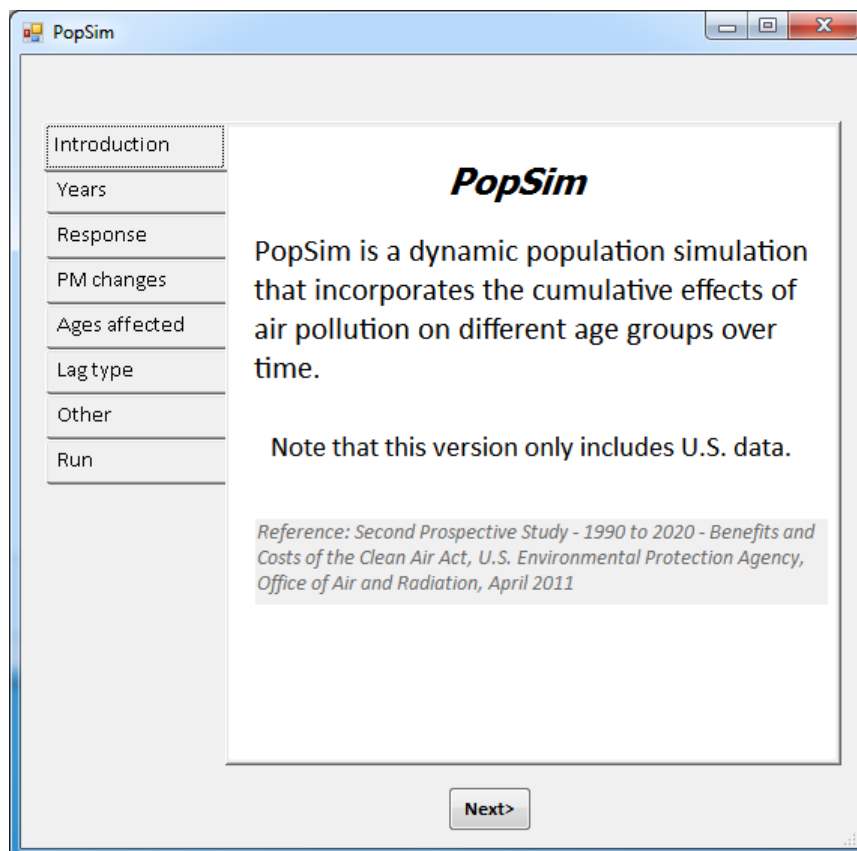
Click through each of the menu items (e.g., **Years**, **Response**, **PM changes**, **Ages affected**, **Lag type**, and **Other**) and review the default settings, or adjust the values as needed. For more information on the underlying datasets or model specifications, please refer to the "*Population Simulation Model for Air Pollution Hazards. Version 3.0. User Manual and Documentation*" (IEc, September 2015).

When you are finished with the model settings, click the **Run model** button on the **Run** screen. A progress bar will be displayed as the calculations are performed. When the model run has completed, a dialog will be displayed. Click **OK** to acknowledge the message. Then click the **Output** button. Specify a location to save the model results. By default, the tool will save to the "*My Documents\My BenMAP-CE Files\PopSim*" folder. The progress bar will once again display as the files are saved to your destination. When complete, a dialog will indicate "*File Saved*". Click **OK** to acknowledge the message. You may close the PopSim tool, or return to the beginning to initiate a new model run. The tool will remember the settings from the previous run only.

9.11 Options

You can customize some of the generic options for BenMAP-CE using the *Options* window available from the **Tools** menu.
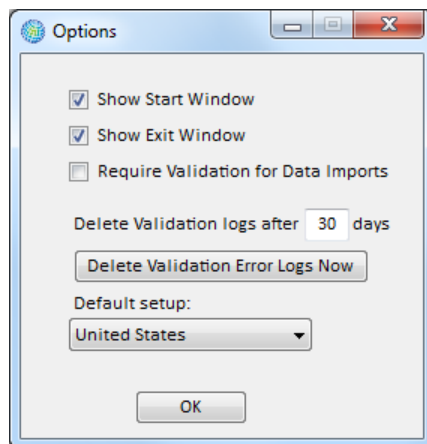
- **Show Start Window**: If unchecked, the **Welcome** window will not appear during subsequent start-ups. You can update your choice at any time.

- **Show Exit Window:** If unchecked, the window confirming you would like to exit will not appear during subsequent shut-downs. You can update your choice at any time.

- **Require Validation for Data Imports:** If checked, you will be required to validate their input files prior to importing datasets. If unchecked, validation will be an available option but it will not be required.

- **Delete Validation logs after __ days:** You can specify the number of days BenMAP-CE will retain the validation logs (the default value is 30 days). If the number of days is left blank, BenMAP-CE will not automatically delete any validation logs.[13]

- **Delete Validation Error Logs Now:** Select this button if you want to immediately clear all validation logs*. Note: There is no confirmation option here – once the button is clicked, all logs will be deleted.*

- Default setup: Select the preferred setup to appear by default in the main BenMAP-CE window.

9.12 Compute Grid Crosswalk

The Compute Grid Crosswalk tool allows users to manually remove crosswalks and re-generate selected crosswalks. A crosswalk is a file used to relate data, such as air quality, population and demographic data, at one spatial scale to another. Crosswalks

---

[13] Validation logs are saved at C:\Users\<User Name>\Documents\My BenMAP-CE Files\ValidationResults.

are created during when importing a grid definition or calculating health impacts and economic benefits. BenMap-CE does not remove or overwrite a crosswalk automatically if it already exists. In cases where a crosswalk was broken due to database error or force quitting of the application, users can use this tool to remove the crosswalk from the database. Users may then let BenMap-CE automatically re-generate the crosswalk during processing or re-create it in the tool to save time later.

To re-create crosswalks, go to the **Tools** menu, and choose the **Compute Grid Crosswalks**.

This will bring up the Crosswalk Calculator window. You will see all grid definitions associated to the active setup are listed in both windows. To switch to a different setup, choose another setup from **Available Setups**.
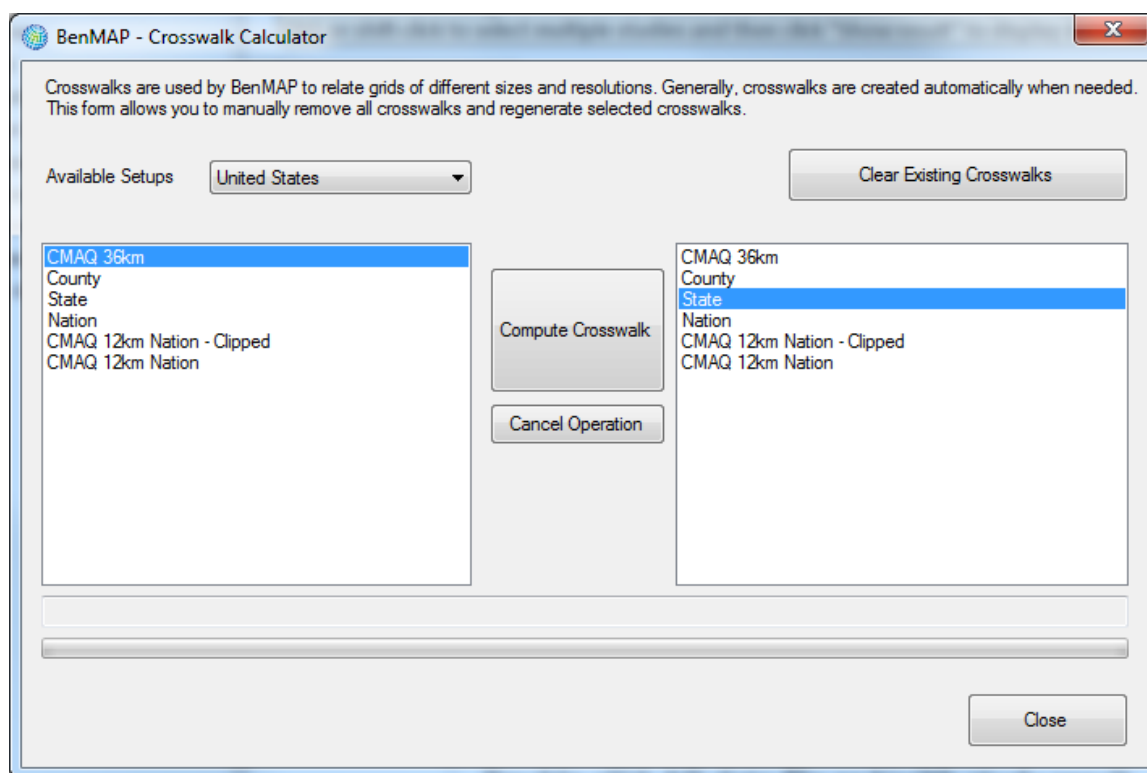
To remove all crosswalks among the grid definitions in the windows, click **Clear Existing Crosswalks**. A message box will pop up asking you to confirm this operation. Click **OK** to confirm. In the screenshot below, we are trying to remove all existing crosswalks for *Detroit* setup from the database. NOTE: crosswalks for default setups (e.g. *China* and *United States*) are locked and cannot be removed.
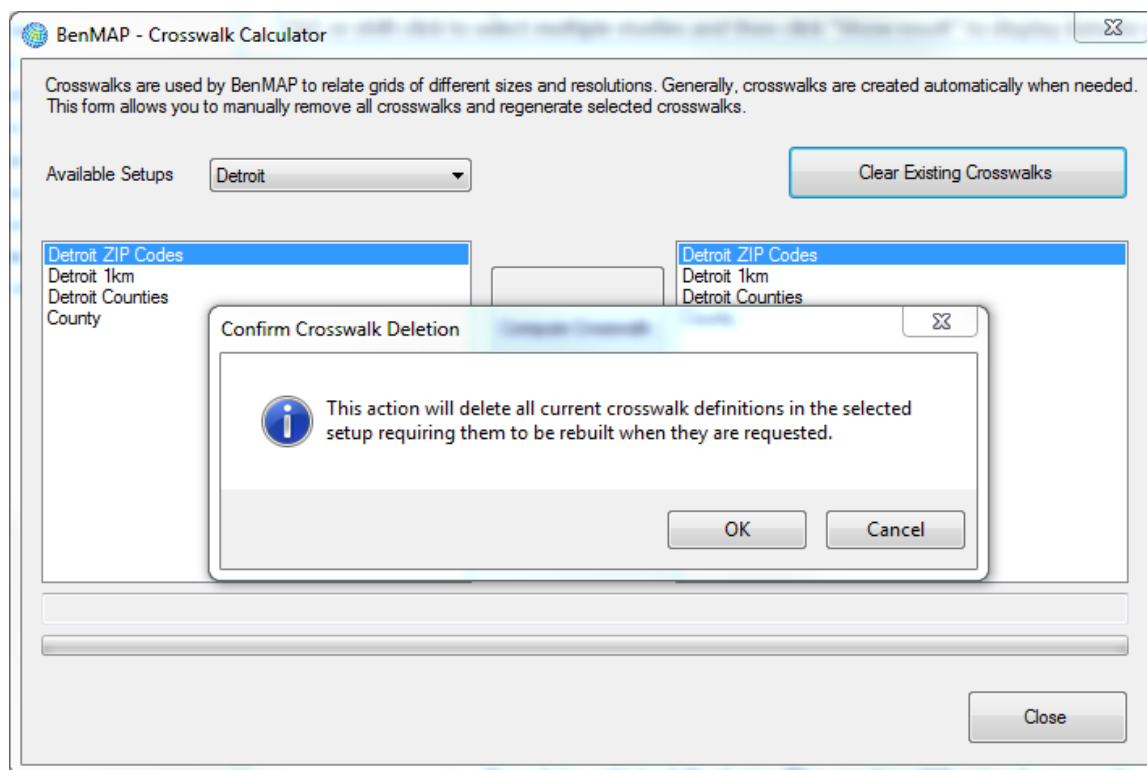
To create or overwrite a crosswalk between two grid definitions, highlight one grid definition from the left side, highlight the other grid definition from the right side, and click Compute Crosswalk button. If the crosswalk already exists, you will be asked to confirm replacement by clicking **OK** in the message box. In the screenshot below, we selected *Detroit ZIP Codes* from the left window and *Detroit Counties* from the right window to create crosswalks between the grids for both directions.
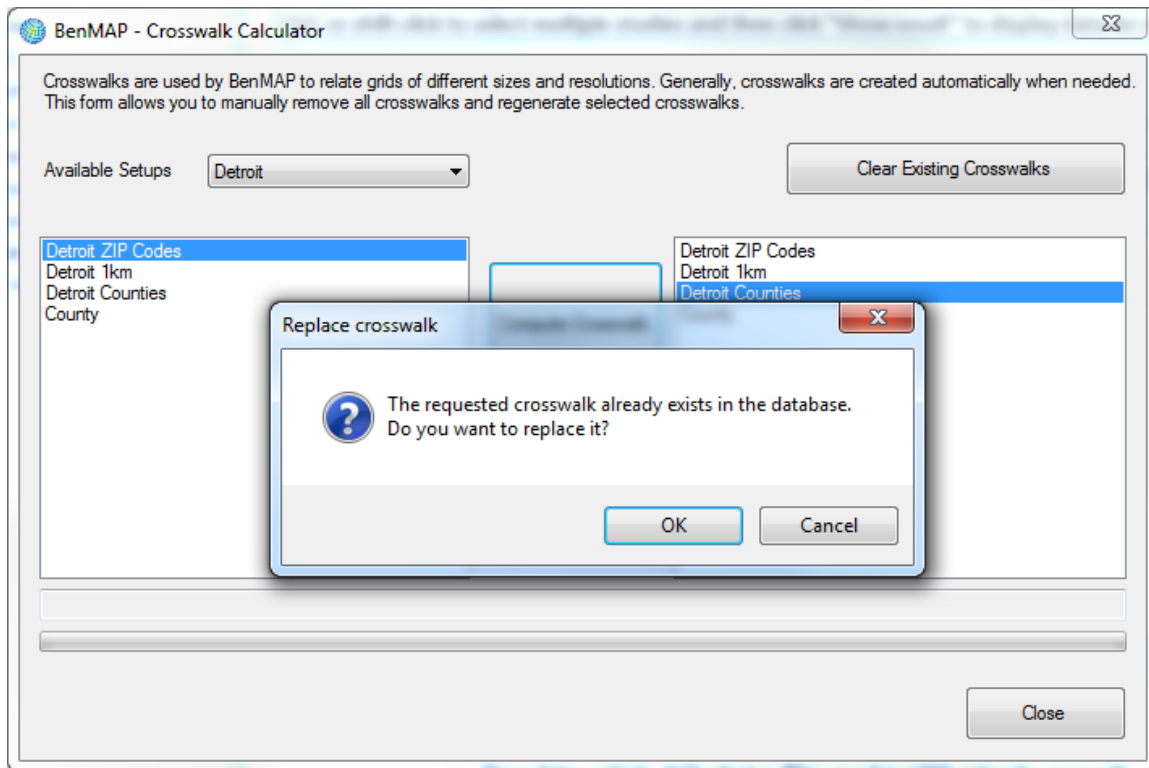
The calculation may process slowly depending on the size and resolution of your grid cells. If you have to interrupt the process, simply click Cancel Operation. It is recommended to highlight the grid definition with higher resolution from the left window and the one with lower resolution from the right window. This will speed up the calculation.

Environmental Benefits Mapping and Analysis Program – Community Edition

User's Manual

Appendices

July 2018

Appendices Table of Contents

Appendix A.   Monitor Rollback Algorithms

This Appendix details the algorithms BenMAP-CE calculates when you perform a "monitor rollback." The monitor rollback adjusts the air quality monitoring data to reflect hypothetical changes in air pollution in a given location. The program allows you to roll back the monitoring data using three approaches: 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 an area in which you perform a specified rollback. Three rollback **types** are available:

- Percentage Rollback. Monitor values are reduced the same percentage.

- Incremental Rollback. Monitor values are changed 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.

A.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

A.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, the values 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

A.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 (PM$_{2.5}$) 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 **Intraday 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).

A.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 three supported rollback methods for Interday Rollbacks: Percentage, Incremental, and Peak Shaving. 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.

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

A.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 meet 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

A.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 - 40)

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 0 30 10 0 0 18 20 0 0 0 0 20 10 15

Target Metric Values:

 30 35 40 70 50 40 4 58 60 40 40 30 40 60 50 55

A.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 |

A.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 two supported rollback methods for Intraday Rollback – Percentage and Incremental. 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.

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).*

A.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.

A.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.

A.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.

A.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.

Appendix B. Air Pollution Exposure Estimation Algorithms

BenMAP groups counts of individuals into what we refer to as "population grid cells," where the grid cells typically correspond to some type of grid used in an air quality model, such as the CMAQ model, or to a grid defined by political boundaries such as the counties of the United States. In the United States setup, the program includes population counts aggregated to each 12km by 12km grid cell. In the next step, BenMAP estimates the air pollution exposure for each grid-cell, thus assuming 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.

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 B-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 throughout the year.   BenMAP accounts for missing days from different data sources as described below.

Table B-1. Metrics Typically Used in Concentration-Response Functions for Criteria Air Pollutants

Measurement Frequency	Metric Name	Metric Description
Daily (e.g., PM$_{2.5}$)	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 throughout 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.

B.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 assumes that the missing days have the same values as the seasonal average of available data.

B.2   Closest Monitor

When using the closest 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. The annual metrics and daily metrics are then used to calculate 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).

= Center Grid-Cell "E"

* = Air Pollution Monitor
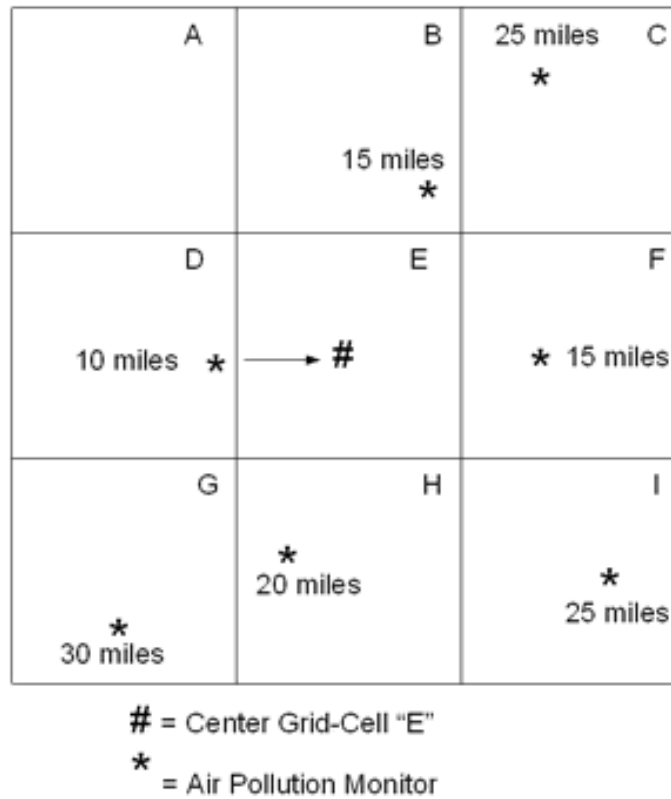
B.3 Voronoi Neighbor Averaging (VNA)

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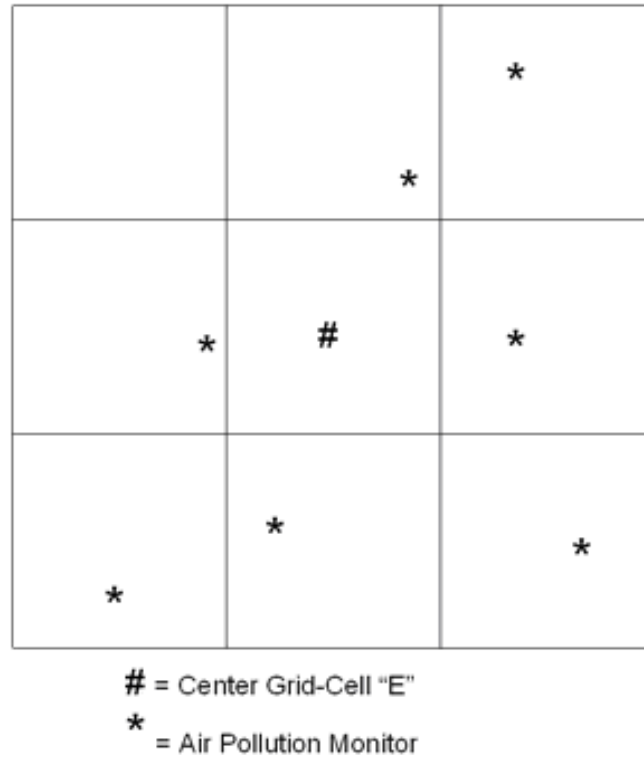
In particular, BenMAP identifies the nearest monitors, or "neighbors," by drawing a polygon, or "Voronoi" cell, around the center of each BenMAP grid cell. The polygons have the special property that the boundaries are the same distance from the two closest points.

\# = Center Grid-Cell "E"

* = Air Pollution Monitor
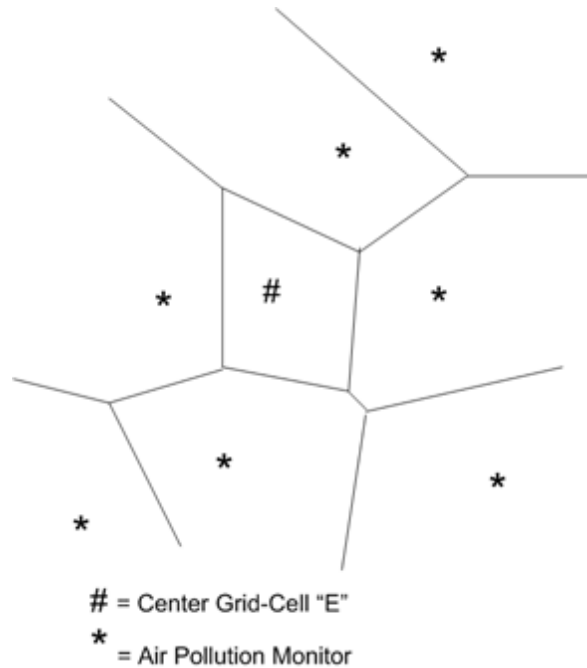
BenMAP then chooses those monitors that share a boundary with the center of grid-cell "E." These are the nearest neighbors, BenMAP uses these monitors to estimate the air pollution level for this grid-cell.

= Center Grid-Cell "E"

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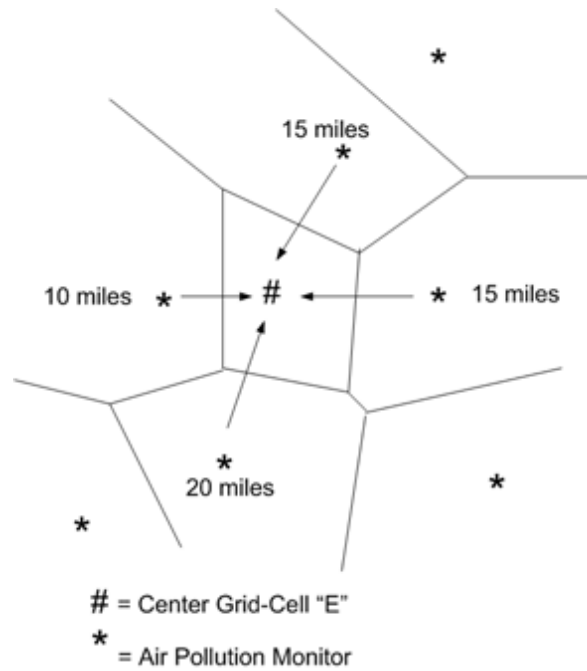
To estimate the air pollution level in each grid-cell, BenMAP calculates the metrics for each of the neighboring monitors, and then calculates an inverse-distance weighted average of the metrics. The further the monitor is from the BenMAP grid-cell, the smaller the weight.

In the figure below, the weight for the monitor 10 miles from the center of grid-cell E is calculated as follows:

$$weight_{10miles} = \frac{\frac{1}{10}}{\left(\frac{1}{10} + \frac{1}{15} + \frac{1}{15} + \frac{1}{20}\right)} = 0.35$$

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
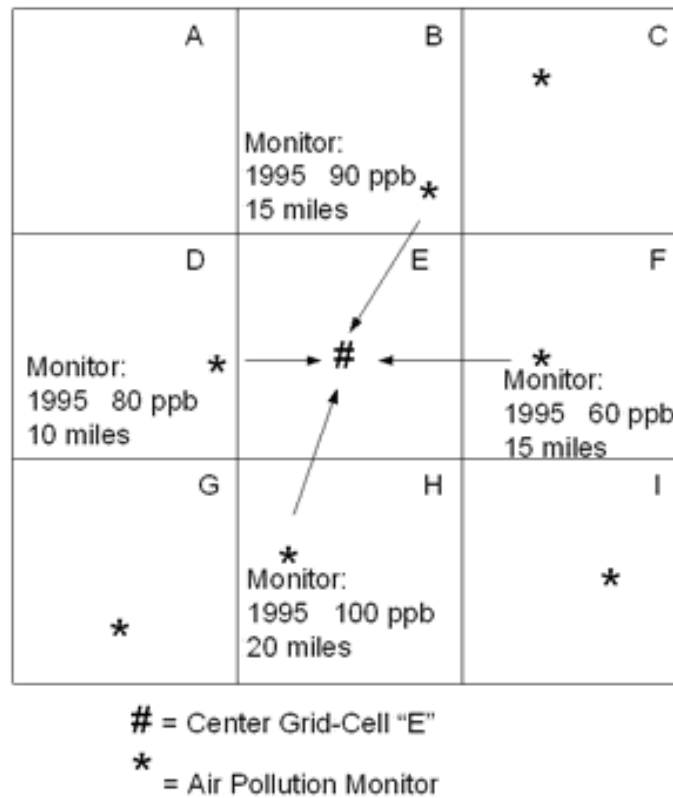
B.4 Fixed Radius

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.

B.5 Monitor or Model Data with Missing Days

When estimating air pollution exposure, it will often happen that metrics are missing for one or more days in the season or year. To remedy this, BenMAP calculates seasonal average values to substitute for missing daily values within each season. When combining air pollution metrics from multiple monitors, BenMAP first calculates the seasonal average values for the daily metrics, substitutes these for missing values, and then performs the user-specified interpolation method, such as VNA.

Appendix C.   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 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 parameters may be estimated by regression techniques but are often estimated by maximum likelihood techniques. The form of the model, however, is still log-linear.

C.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 fine particulate matter ($PM_{2.5}$) may be in $\mu g/m^3$ 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 to estimate 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.

C.2   Review Relative Risk and Odds Ratio

The terms relative risk and odds ratio are related but distinct. Table C-1 provides the basis for demonstrating their relationship.

Table C-1. Relative Risk and Odds Ratio Notation

	Fraction of Population		**Adverse Effect Measure**	
Exposure	**Affected**	**Not Affected**	**Relative Risk**	**Odds**
Baseline Pollutant Exposure	y_0	$1-y_0$	y_0/y_c	$y_0/(1-y_o)$
Control Pollutant Exposure	y_c	$1-y_c$		$y_c/(1-y_c)$

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_0, 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:

$$Odds\ Ratio = \frac{\left(\frac{y_0}{1-y_0}\right)}{\left(\frac{y_c}{1-y_c}\right)}$$

This can be rearranged as follows:

$$Odds\ Ratio = \frac{y_0}{y_c} \times \left(\frac{1-y_c}{1-y_0}\right) = RR \times \left(\frac{1-y_c}{1-y_0}\right)$$

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.

C.3   Linear Model

A linear relationship between the rate of adverse health effects (incidence rate) and various explanatory variables is of the form:

$$y = \alpha + \beta \times PM$$

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:

$$\Delta y = y_0 - y_c = \beta * (PM_0 - PM_c) = \beta * \Delta PM$$

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:

$$\beta_{lowerbound} = \beta - \left(1.96 \times \sigma_\beta\right) = 0.0006 - (1.96 \times 0.0003) = 1.2 \times 10^{-5}$$

$$\beta_{upperbound} = \beta + \left(1.96 \times \sigma_\beta\right) = 0.0006 + (1.96 \times 0.0003) = 0.00119$$

It is then straightforward to calculate lower and upper bound estimates of the change in asthma.

C.4   Log-linear Model

The log-linear relationship defines the incidence rate (y) as:

$$y = B \times e^{\beta * PM}$$

Or, equivalently,

$$ln(y) = \alpha + \beta * PM,$$

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 α = 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,

$$B = B_0 \times e^{\beta_1 x_1 + \cdots + \beta_n x_n}$$

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:

$$\Delta y = y_0 - y_c = B\left(e^{\beta PM_0} - e^{\beta PM_c}\right)$$

This may be rewritten as:

$$\Delta y = B \times e^{\beta \bullet PM_0}\left(1 - e^{-\beta(PM_0 - PM_c)}\right) = y_0\left(1 - \frac{1}{\exp(\beta \times \Delta PM)}\right)$$

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 ($\sigma\beta$) 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:

$$\beta_{lowerbound} = \beta - \left(1.96 \times \sigma_\beta\right)$$

$$\beta_{upperbound} = \beta + \left(1.96 \times \sigma_\beta\right),$$

Where the adjustment on the mean of ±1.96 times the standard error produces the 2.5[th] and 97.5[th] percentiles of the normal distribution, which are used to approximate a 95% confidence interval. These values can be changed to capture different lower and upper bounds.

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:

$$RR = \frac{y_0}{y_c} = e^{\beta \bullet \Delta PM}$$

Taking the natural log of both sides, the coefficient in the C-R function underlying the relative risk can be derived as:

$$\beta = \frac{In(RR)}{\Delta PM}$$

The coefficients associated with the lower and upper bounds (e.g., the 2.5th and 97.5th 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:

$$\sigma_{\beta, 2.5\,percentile} = \frac{\beta - \beta_{2.5\,percentile}}{1.96}$$

$$\sigma_{\beta, 97.5\,percentile} = \frac{\beta_{97.5\,percentile} - \beta}{1.96}$$

$$\sigma_\beta \cong \frac{\sigma_{\beta, 2.5\,percentile} + \sigma_{\beta, 97.5\,percentile}}{2}$$

C.5   Logistical 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:

$$y = prob(occurrence \mid X \times \beta) = \left(\frac{e^{X \bullet \beta}}{1 + e^{X \bullet \beta}} \right),$$

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:

$$y = \frac{e^{X \bullet \beta}}{1 + e^{X \bullet \beta}} \times \frac{e^{-X \bullet \beta}}{e^{-X \bullet \beta}} = \frac{1}{1 + e^{-X \bullet \beta}}$$

The odds of an occurrence is:

$$odds = \frac{y}{1 - y} = \frac{\left(\dfrac{1}{1 + e^{-X \bullet \beta}}\right)}{1 - \dfrac{1}{1 + e^{-X \bullet \beta}}}$$

$$\Rightarrow \quad odds = \frac{\left(\dfrac{1}{1 + e^{-X \bullet \beta}}\right)}{1 - \dfrac{1}{1 + e^{-X \bullet \beta}}} = \frac{\left(\dfrac{1}{1 + e^{-X \bullet \beta}}\right)}{\left(\dfrac{e^{-X \bullet \beta}}{1 + e^{-X \bullet \beta}}\right)} = \frac{1}{e^{-X \bullet \beta}} = e^{X \bullet \beta}$$

$$\Rightarrow ln(odds) = X \times \beta$$

The odds ratio for the control scenario (odds$_c$) versus the baseline (odds$_0$) is then:

$$odds\ ratio = \frac{odds_c}{odds_0} = \frac{\left(\dfrac{y_c}{1 - y_c}\right)}{\left(\dfrac{y_0}{1 - y_0}\right)} = \frac{\left(\dfrac{1}{e^{-X_c \bullet \beta}}\right)}{\left(\dfrac{1}{e^{-X_0 \bullet \beta}}\right)} = \frac{e^{X_c \bullet \beta}}{e^{X_0 \bullet \beta}}$$

The *change* in the probability of an occurrence from the baseline to the control (Δy), assuming that all of the other covariates remain constant, may be derived from this odds ratio:

$$odds\ ratio = \frac{\dfrac{y_0}{1 - y_0}}{\dfrac{y_c}{1 - yc}} = \frac{e^{X_0 \bullet \beta}}{e^{X_i \bullet \beta}} = e^{\beta \bullet \Delta X} = e^{\beta \bullet \Delta PM}$$

$$\frac{y_c}{1 - y_c} = \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM}$$

$$y_c = (1 - y_c) \times \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM}$$

$$y_c + y_c \times \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM} = \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM}$$

$$y_c \left(1 + \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM} \right) = \frac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM}$$

$$y_c = \frac{\dfrac{y_0}{1 - y_0} \times e^{-\beta \bullet PM}}{1 + \dfrac{y_0}{1 - y_0} \times e^{-\beta \bullet \Delta PM}} = \frac{y_0 \times e^{-\beta \bullet \Delta PM}}{1 - y_0 + y_0 \times e^{-\beta \bullet \Delta PM}} = \frac{y_0}{(1 - y_0) \times e^{\beta \bullet \Delta PM} + y_0}$$

$$\Delta y = y_0 - y_c = y_0 - \frac{y_0}{(1 - y_0) \times e^{\beta \bullet \Delta PM} + y_0}$$

$$\Delta Incidence = \Delta y \times pop = y_0 \times \left(1 - \frac{1}{(1 - y_0) \times e^{\beta \bullet \Delta PM} + y_0} \right) \times pop$$

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:

$$\beta_{lowerbound} = \beta - \left(1.96 \times \sigma_\beta \right)$$

$$\beta_{upperbound} = \beta + \left(1.96 \times \sigma_\beta \right),$$

where the adjustments to the mean of plus or minus 1.96 times the standard error represent the 2.5th and 97.5th percentiles of the normal distribution, and are used to approximate a 95% confidence interval. These values can be changed to capture different lower and upper bounds.

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:

$$ln(odds\ ratio) = \beta \times \Delta PM$$

$$\Rightarrow \beta = \frac{ln(odds\ ratio)}{\Delta PM}$$

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 y_c 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:

$$\beta = \frac{In(odds\ ratio)}{\Delta PM} ,$$

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/m^3 of PM$_{2.5}$ exposure.

Assuming that the baseline rate for chronic bronchitis (y0) should be 7.17 percent, the question becomes whether the "increment" of 45 µg/m^3 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:

$$odds\ ratio_1 = \frac{\left(\dfrac{y_0}{1-y_0}\right)}{\left(\dfrac{y_c}{1-y_c}\right)} = \frac{\left(\dfrac{1.81 \times 0.0717}{1-(1.81 \times 0.0717)}\right)}{\left(\dfrac{0.0717}{1-0.0717}\right)} = 1.931$$

$$odds\ ratio_2 = \frac{\left(\dfrac{y_0}{1-y_0}\right)}{\left(\dfrac{y_c}{1-y_c}\right)} = \frac{\left(\dfrac{0.0717}{1-0.0717}\right)}{\left(\dfrac{\dfrac{0.0717}{181}}{1-\dfrac{0.0717}{1.81}}\right)} = 1.873$$

$$\Rightarrow \beta_1 = \frac{In(1.931)}{45} = 0.01462$$

$$\Rightarrow \beta_2 = \frac{In(1.873)}{45} = 0.01394$$

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:

$$RR = \frac{y_0}{y_c} = (1-y_0) \times e^{-\Delta PM \bullet \beta} + y_0$$

Assuming that:

$$e^{-\Delta PM \bullet \beta} \cong \left(1 - y_0\right) \times e^{-\Delta PM \bullet \beta} + y_0$$

$$\Rightarrow RR \cong e^{-\Delta PM \bullet \beta}$$

It is then possible to calculate the underlying coefficient:

$$\frac{In(RR)}{-\Delta PM} \cong \beta$$

$$\Rightarrow \beta_3 = \frac{In(1.81)}{45} = 0.01319$$

Since this coefficient estimate is based on the assumption that

$$e^{-\Delta PM \bullet \beta} \cong \left(1 - y_0\right) \times e^{-\Delta PM \bullet \beta} + y_0 \quad ,$$

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:

$$\Delta y = \left[y_0 \times \left(e^{\beta \bullet \Delta PM} - 1 \right) \right]$$

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:

$$\Delta y_1 = \frac{.0717}{\left(1 - 0.0717\right) \times e^{10 \times 0.01462} + 0.0717} - 0.0717 = -0.00914$$

$$\Delta y_2 = \frac{.0717}{\left(1 - 0.0717\right) \times e^{10 \times 0.01394} + 0.0717} - 0.0717 = -0.00874$$

$$\Delta y_3 = 0.0717 \times \left(e^{-10 \times 0.01319} - 1 \right) = -0.00886$$

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 is a reasonable approach.

C.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:

$$h(X,t) = h_0(t)e^{X \bullet \beta} \; ,$$

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:

$$RR = \frac{h(X_0, t)}{h(X_c, t)} = \frac{h_0(t)e^{X_0 \bullet \beta}}{h_0(t)e^{X_c \bullet \beta}} = e^{\Delta PM \bullet \beta},$$

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 the functional relationship between ΔPM and Δy, as described previously for log-linear C-R functions.

Appendix D.   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.

D.1   Mortality

This section describes how we developed county mortality rates for the years 2015 through 2050 to use in BenMAP. First, we describe the source of 2012-2014 baseline mortality data and how we calculated county-level mortality rates. We then describe how we used national-level Census mortality rate projections to develop county-level mortality rate projections for years 2015-2060.

D.1.1 Mortality Data for 2012-2014

We obtained county-level mortality and population data from 2012-2014 for seven causes for the contiguous United States by downloading the data from the Centers for Disease Control (CDC) WONDER database (http://wonder.cdc.gov).

Since the detailed mortality data obtained from CDC do not include population, we combined them with U.S. Census Bureau population estimates exported from BenMAP. We then generated age-, cause-, and county-specific mortality rates using the following formula:

$$R_{i,j,k} = \frac{D_{i,j,k}(2012) + D_{i,j,k}(2013) + D_{i,j,k}(2014)}{P_{i,k}(2012) + P_{i,k}(2013) + P_{i,k}(2014)}$$

where $R_{i,j,k}$ is the mortality rate for age group i, cause j, and county k; D is the death count; and P is the population.

For county-age group cells with fewer than 10 deaths, CDC WONDER suppresses the exact death count. For these observations, a mortality rate cannot be calculated. For each combination of age group and mortality cause, we used the following procedure to deal with suppressed counts.

For each combination of state, age group and mortality cause, we grouped counties with unsuppressed mortality figures and summed their reported death counts. We then subtracted these unsuppressed deaths from the state-level age- and cause-specific death count, which includes suppressed deaths. We divided the resulting state-wide

death count in suppressed counties by the age-specific populations in those counties. This calculation resulted in an age- and cause- specific average mortality rate for suppressed counties;

$$R_{s,i,j} = \frac{D_{T,i,j} - D_{u,i,j}}{P_{s,i,j}}$$

Where $R_{s,i,j}$ is the state average suppressed mortality rate for age group i and cause j; $D_{T,i,j}$ is the total state death count for age group i and cause j; $D_{u,i,j}$ is the aggregated state-level unsuppressed death count for age group i and cause j; and $P_{s,i,j}$ is the aggregated population for age group i and cause j in suppressed counties.

In some instances, age- and cause-specific death counts were suppressed at both the county and state level. In these cases, we substituted national-level age- and cause-specific mortality rates for the respective missing county mortality rates.

Following CDC WONDER (http://wonder.cdc.gov), we treated mortality rates as "unreliable" when the death count is less than 20. For each combination of age group and mortality cause, we used the following procedure to deal with the problem of "unreliable" rates:

For a given state, we grouped the counties where the death count was less than 20 and summed those death counts across those counties. If the sum of deaths was greater than or equal to 20, we then summed the populations in those counties, and calculated a single rate for the "state collection of counties" by dividing the sum of deaths by the sum of populations in those counties. This rate was then applied to each of those "unreliable" counties.

If the sum of deaths calculated in the above step was still less than 20, the counties in the "state collection of counties" were not assigned the single rate from the above step. Instead, we proceeded to the regional level, according to the regional definitions shown below in Table D-1. In each region, we identified all counties whose death counts were less than 20 (excluding any such counties that were assigned a rate in the previous step). We summed the death counts in those counties. If the sum of deaths was greater than or equal to 20, we then summed the populations in those counties, and calculated a single rate for the "regional collection of counties" by dividing the sum of deaths by the sum of populations in those counties. This rate was then applied to each of those counties in the "regional collection of counties."

Table D-1. Regional Definitions from U.S. Census

Region	States Included
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

If the sum of deaths calculated in the previous (regional) step was still less than 20, the counties in the "regional collection of counties" were not assigned the single rate from the above step. Instead, we proceeded to the national level, identifying all counties in the nation whose death counts were less than 20 (excluding any such counties that were assigned a rate in the previous steps). We summed the death counts in those counties and divided by the sum of the populations in those counties to derive a single rate for the "national collection of counties." This rate was then applied to each of those counties in the "national collection of counties." In these cases where national adjustment still did not yield a death count greater than 20, we simply calculated a single rate for the "national collection of counties, even though it was "unreliable," and assigned it to those counties in the "national collection of counties."

Table D-2. National Mortality Rates (per 100 people per year) by Health Endpoint and Age Group, 2012-2014

Mortality Category	ICD-10 codes	Infant*	1-17	18-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
Mortality, All Cause	All	0.59396	0.01951	0.07804	0.10665	0.17264	0.40542	0.86162	1.79670	4.62837	13.58034
Mortality, Non-Accidental	A00-R99	0.55495	0.00949	0.01874	0.04112	0.10876	0.33084	0.79395	1.73208	4.49595	13.20867
Mortality, Respiratory	J00-J98	0.01297	0.00102	0.00127	0.00253	0.00570	0.02013	0.06560	0.20585	0.57827	1.42362
Mortality, Chronic Lung	J40-J47, J67	0.00053	0.00032	0.00040	0.00074	0.00186	0.01033	0.04045	0.13873	0.36008	0.68593
Mortality, Lung Cancer	C34	0.00002	0.00001	0.00007	0.00033	0.00282	0.02378	0.07992	0.19701	0.32952	0.31820
Mortality, Ischemic Heart Disease	I20-I25	0.00033	0.00004	0.00039	0.00234	0.01242	0.04854	0.12174	0.25698	0.68000	2.27271
Mortality, Cardio-Pulmonary	I00-I78, J10-J18, J40-J47, J67	0.00539	0.00069	0.00099	0.00214	0.00502	0.01794	0.05877	0.18453	0.51055	1.26213

*We estimate post-neonatal mortality (deaths after the first month) for infants because the health impact function (see Appendix E) estimates post-neonatal mortality.

D.1.2 Mortality Rate Projections 2015-2060

To estimate age- and county-specific mortality rates in years 2015 through 2060, we calculated annual adjustment factors, based on a series of Census Bureau projected national mortality rates (for all- cause mortality), to adjust the age- and county-specific mortality rates calculated using 2012-2014 data as described above. We used the following procedure:

For each age group, we obtained the series of projected national mortality rates from 2013 to 2050 (see the 2013 rate in Table D-3) based on Census Bureau projected life tables.

We then calculated, separately for each age group, the ratio of Census Bureau national mortality rate in year Y (Y = 2014, 2015, ..., 2060) to the 2013 rate. These ratios are shown for selected years in Table D-4.

Finally, to estimate mortality rates in year Y (Y = 2015, 2020, ..., 2060) that are both age-group-specific and county-specific, we multiplied the county- and age-group-specific mortality rates for 2012-2014 by the appropriate ratio calculated in the previous step. For example, to estimate the projected mortality rate in 2015 among ages 18-24 in Wayne County, MI, we multiplied the mortality rate for ages 18-24 in Wayne County in 2012-2014 by the ratio of Census Bureau projected national mortality rate in 2015 for ages 18-24 to Census Bureau national mortality rate in 2013 for ages 18-24.

**Table D-3. 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+
Calculated CDC 2012-2014	0.594*	0.020	0.078	0.107	0.173	0.405	0.862	1.797	4.628	13.580
Census Bureau 2013**	0.654	0.029	0.088	0.102	0.183	0.387	0.930	2.292	5.409	13.091

* The Census Bureau estimate is for all deaths in the first year of life. BenMAP uses post-neonatal mortality (deaths after the first month, i.e., 0.23 per 100 people) because the health impact function (see Appendix E) estimates post- neonatal mortality. For comparison purpose, we also calculated the rate for all deaths in the first year, which is 0.684 per 100 people).

**For a detailed description of the model, the assumptions, and the data used to create Census Bureau projections, see the working paper, "Methodology and Assumptions for the 2012 National Projections," which is available on http://www.census.gov/population/projections/files/methodology/methodstatement12.pdf

**Table D-4. Ratio of Future Year All-Cause Mortality Rate to 2013 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+
2015	0.93	0.93	0.96	1.02	0.96	0.96	1.01	1.02	1.03	1.00
2020	0.94	0.94	0.98	1.04	0.97	0.98	1.02	1.03	1.03	1.00
2025	0.85	0.81	0.74	0.80	0.75	0.77	0.85	0.91	0.93	0.97
2030	0.81	0.75	0.66	0.70	0.67	0.69	0.78	0.86	0.89	0.92
2035	0.76	0.70	0.58	0.62	0.60	0.62	0.71	0.81	0.87	0.87
2040	0.73	0.65	0.51	0.53	0.53	0.56	0.64	0.76	0.84	0.86
2045	0.70	0.60	0.45	0.46	0.46	0.50	0.58	0.71	0.80	0.86
2050	0.67	0.56	0.39	0.40	0.40	0.44	0.53	0.66	0.77	0.87
2055	0.64	0.52	0.34	0.35	0.35	0.39	0.48	0.62	0.73	0.88
2060	0.61	0.48	0.30	0.30	0.31	0.34	0.43	0.58	0.70	0.87

D.1.3 Race-Stratified Incidence rates

To estimate race-stratified and age-stratified incidence rates at the county level, we downloaded all-cause mortality data from 2007 to 2016 from the CDC WONDER mortality database (https://wonder.cdc.gov/). Race-stratified incidence rates were calculated for the following age groups: < 1 year, 1- 4 years, 5-14 years, 15-24 years, 25-34 years, 35-44 years, 45-54 years, 55-64 years, 65-74 years, 75-84 years, and 85+ years. We stratified the data into two race categories, White and Non-White, and follow all methods outlined in section D.1.1. To properly impute incidence rates for suppressed and unreliable counties, we downloaded data at the state, regional, and national scales.

D.2   Hospitalizations

Hospitalization rates were calculated using data from the Healthcare Cost and Utilization Project (HCUP). HCUP is a family of health care databases developed through a Federal-State-Industry partnership and sponsored by the Agency for Healthcare Research and Quality (AHRQ). HCUP products include the State Inpatient Databases (SID), the State Emergency Department Databases (SEDD), the Nationwide Inpatient Sample (NIS), and the Nationwide Emergency Department Sample (NEDS). HCUP databases can be obtained from the following data services:

HCUP Central Distributor: Many of the HCUP databases are available for purchase through the HCUP Central Distributor. The databases include detailed information for individual discharges, such as primary diagnosis (in ICD-9 codes), patient's age and residence county.

HCUP State Partners: Some HCUP participating states do not release their data to the Central Distributor; however, the data may be obtained through contacting the State Partners. South Carolina provided county-level data.

HCUPnet: This is a free, on-line query system based on data from HCUP. It provides access to summary statistics at the state, regional and national levels.

Figure D-1 shows the level of hospitalization data (e.g, discharge-level or state-level) for each state. Note that for some states neither discharge-level, county-level nor state-level data were available. In such cases we used regional statistics from HCUPnet to estimate hospitalization rates for those states. The data year for states using HCUPnet data is 2014. For discharge-level data, the data year for most states is 2014; however, some states provided data for 2011 (CA, MS); 2012 (ME); and 2013 (AR, MA, MD, NV, SD, UT). We assume hospitalization rates are reasonably constant from 2011-2014 and consider all as 2014 rates.

Figure D-1. Hospitalization Data from HCUP
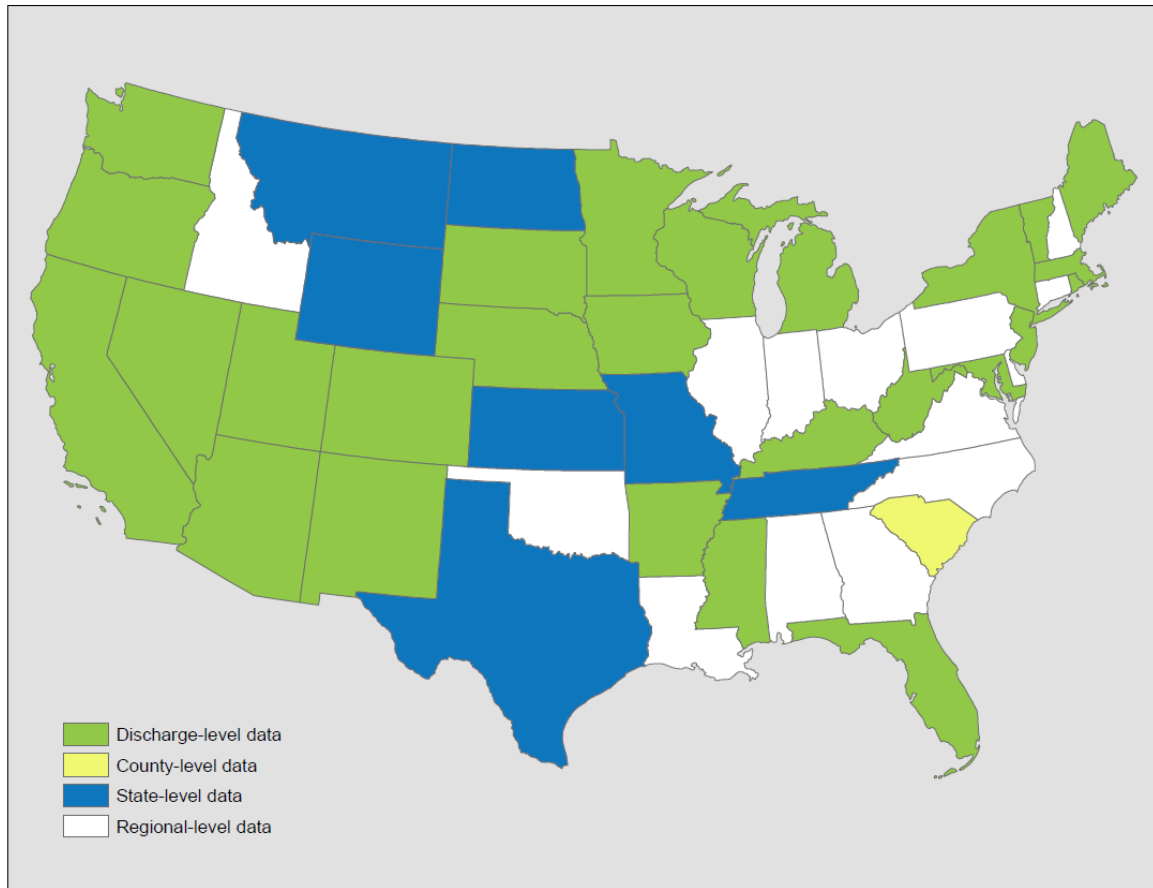
More information about HCUP can be found at http://www.hcup-us.abrq.gov/

The procedures for calculating hospitalization rates are summarized as follows:

For states with discharge-level data:

o We calculated age-, health endpoint-, and county-specific hospitalization counts. South Carolina was the only state that, while not providing discharge-level data, did provide county-level data for each age group-endpoint combination.

o The above calculation excluded hospitalizations with missing patient age or county FIPS, which may lead to underestimation of rates. Therefore we scaled up the previously calculated age-, endpoint-, and county-specific counts using an adjustment factor obtained as follows:

▪ We first counted the number of discharges for a specific endpoint in the state including those discharges with missing age or county FIPS.

- We then counted the number of discharges for the endpoint in the state excluding those records with missing age or county FIPS.

- The adjustment factor is the ratio of the two counts.

o For California and West Virginia, patient county was unavailable for all observations. For these two states, we used hospital county in place of patient county.

o For health outcomes deemed acute (acute myocardial infarction; cerebrovascular events; stroke; pneumonia; lower respiratory infection; acute cases of asthma), we distributed patients within the hospital state in cases where the patient resided out of state. We assume that everyone admitted to the hospital in a given state developed that acute condition while in that state.

o We calculated hospitalization rates for each county by dividing the adjusted county-level hospitalization counts by the Census estimated county-level population for the corresponding year (2011 - 2014). Following CDC Wonder, we treated rates as "unreliable" when the hospitalization count was less than 20, using the same procedure we used for mortality rates (see Section D.1.1).

For states with summarized state statistics (from HCUPnet) we calculated the state-, age-, endpoint- specific hospitalization rates and applied them to each county in the state. We used the previously described procedure to adjust the "unreliable" rates.

For states without discharge-level or state-level data:

o We obtained the endpoint-specific hospitalization counts in each region from HCUPnet/NIS (we refer to this count for the ith endpoint in the jth region as "TOTALij")

o For those states in the jth region that do have discharge-level or state-level data, we summed the hospital admissions by endpoint (we refer to this count for the ith endpoint in the jth region as "SUB ij").

o We then estimated the hospitalization count for states without discharge or state data for the ith endpoint in the jth region as TOTALij - SUB ij. Note that while this count is endpoint- and region- specific, it is not age-specific. We obtained the distribution of hospital admission counts across age groups based on the Central Distributor data and assumed the same distribution for the HCUPnet hospitalizations. We then applied this distribution to the estimated hospital counts (i.e., TOTALij - SUB ij) to obtain endpoint-, region-, and age-specific counts.

      o Using the corresponding age- and region-specific populations in BenMAP-CE from Woods and Poole (2015), we calculated age-specific hospitalization rates for the ith endpoint in the jth region and applied them to those counties in the region that didn't have discharge-level or state-level data.

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:

- congestive heart failure (ICD-9 428)

- dysrhythmia (ICD-9 427)

- heart rhythm disturbances (ICD-9 426-427)

- acute myocardial infarction (ICD-9 410)

- ischemic heart disease - 1 (ICD-9 410-414)

- ischemic heart disease - 2 (ICD-9 410-414, 429)

- ischemic heart disease (less myocardial infarction) (ICD-9 411-414)

- all cardiovascular (ICD-9 390-429)

- all cardiovascular (less myocardial infarctions) (ICD-9 390-409, 411-429)

- cardiovascular, cerebrovascular and peripheral vascular diseases (ICD-9 410-414, 429, 426- 427, 428, 430-438, 440-449)

- all cardiac outcomes (ICD-9 390-459)

- cerebrovascular events (ICD-9 430-438)

- stroke (ICD-9 431-437)

- peripheral vascular disease -1 (ICD-9 440-448)

- peripheral vascular disease -2 (ICD-9 440-449)

- all respiratory (ICD-9 460-519)

- respiratory illness -1 (ICD-9 466, 480-486, 490-493)

- respiratory illness -2 (ICD-9 464-466, 480-487, 490-492)

- chronic lung disease (ICD-9 490-496)

- chronic lung disease (less asthma) (ICD-9 490-492, 494-496)

- chronic lung disease (less asthma) -2 (ICD-9 490-492, 494, 496)

- chronic lung disease (less asthma) -3 (ICD-9 490-492)

- chronic lung disease (less asthma) -4 (ICD-9 491,492, 494, 496)

- pneumonia (ICD-9 480-486)

- asthma (ICD-9 493)

- lower respiratory infection (ICD-9 466.1, 466.0, 480-487, 490, 510-511)

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 D-5 presents a summary of the national hospitalization rates for 2014 from HCUP.

Table D-5. Hospitalization Rates (per 100 people per year), by Health Endpoint and Age

Hospitalization Category	ICD-9 Code	Age 0-1	2-17	18-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
Respiratory											
All Respiratory	460-519	2.387	0.363	0.166	0.212	0.340	0.737	1.297	2.292	4.151	6.343
Pneumonia	480-486	0.477	0.101	0.039	0.063	0.103	0.196	0.336	0.640	1.426	2.660
Chronic Lung Disease	490-496	0.226	0.151	0.041	0.056	0.105	0.281	0.496	0.837	1.276	1.306
Asthma	493	0.217	0.147	0.036	0.048	0.076	0.123	0.136	0.157	0.218	0.243
Cardiovascular											
All Cardiovascular	390-429	0.044	0.017	0.061	0.138	0.377	0.914	1.747	3.131	5.886	8.832
Acute Myocardial Infarction, Nonfatal	410	0.000	0.000	0.002	0.010	0.068	0.202	0.380	0.575	0.921	1.332
Ischemic Heart Disease	410-414	0.000	0.000	0.002	0.014	0.105	0.350	0.689	1.090	1.570	1.734
Dysrhythmia	427	0.016	0.005	0.014	0.025	0.057	0.145	0.319	0.684	1.357	1.917
Congestive Heart Failure	428	0.010	0.001	0.005	0.021	0.061	0.165	0.344	0.700	1.727	3.513
Stroke	431-437	0.009	0.003	0.007	0.021	0.070	0.199	0.417	0.816	1.639	2.488

D.3 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 785,000 new heart attacks per year (Roger et al., 2012). Exclusion of heart attack deaths reported by CDC Wonder yields approximately 575,000 nonfatal cases per year.

An alternative approach to the estimation of heart attack rates is to use data from the Healthcare Cost and Utilization Project (HCUP), assuming that all heart attacks that are not instantly fatal will result in a hospitalization. Details about HCUP data are described in Section D.2. According to the 2014 HCUP data there were approximately 608,795 hospitalizations due to heart attacks (acute myocardial infarction: ICD-9 410, primary diagnosis). We used rates based on HCUP data over estimates extrapolated from cohort studies because the former is a national database with a larger sample size, which is intended to provide reliable national estimates. The incidence rate calculation is also described in Section D.2 and the incidence rates for AMI hospitalization are presented in Table D-5.

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 estimated number of PM-related acute myocardial infarctions to exclude the number of cases that result in death within the first month. Note that we did not adjust for fatal AMIs in the incidence rate estimation, due to the way that the epidemiological studies are designed. Those studies consider total admissions for AMIs, which includes individuals living at the time the studies were conducted. Therefore, we use the definition of AMI that matches the definition in the epidemiological studies.

D.4   Emergency Department Visits

The data source for emergency department/room (ED or ER) visits is also HCUP, i.e., SID, SEDD, and NEDS. And the types of data providers are also the same as those described in Section D.2. Figure D-2 shows the emergency department data in each state.

Figure D-2. Emergency Department Data from HCUP
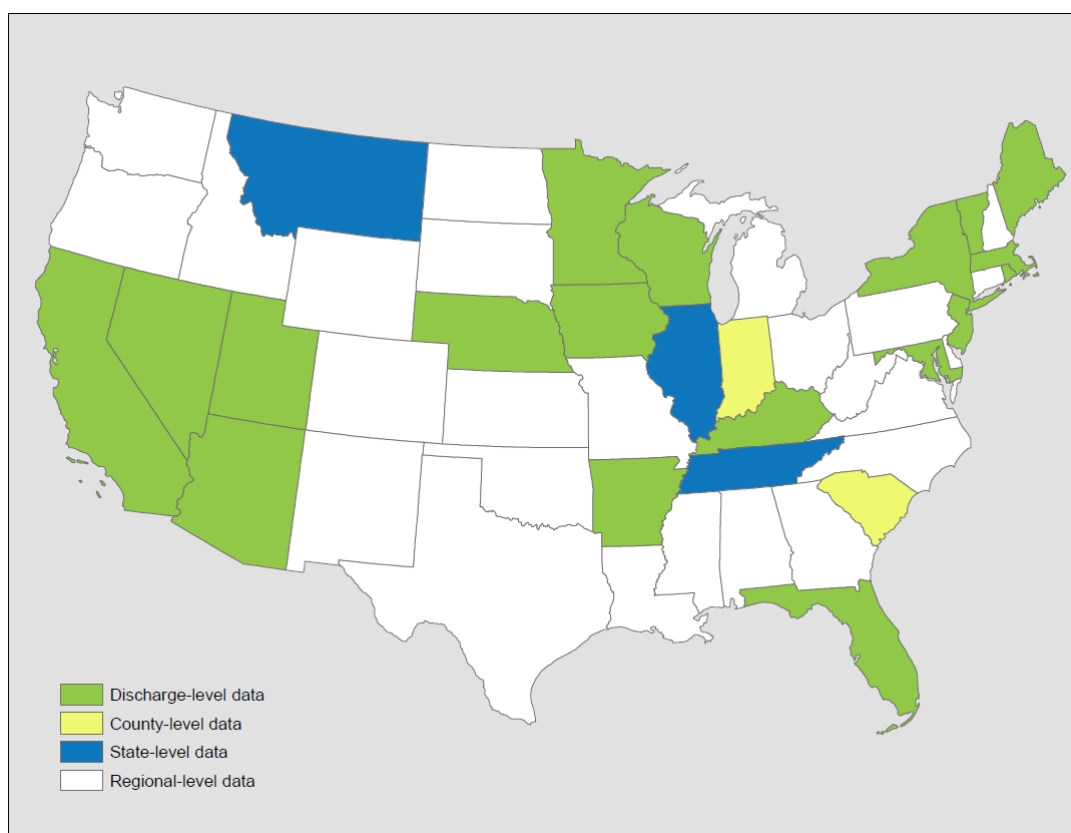
The calculation of ER visit rates is also similar to the calculation of hospitalization rates, except for the following differences:

The SEDD databases include only those ER visits that ended with discharge. To identify the ER visits that ended in hospitalization, we used a variable called "admission source" in the SID databases. Admission source identified as "emergency room" indicates that the hospital admission came from the ER - i.e., the ER visit ended in hospitalization. For each combination of age group, endpoint and county, we summed the ER visits that ended with discharge and those that resulted in hospitalization.

The data year varies across the states from 2011 to 2014; we assumed that ER visit rates are reasonably constant across these three years and consider them as 2014 rates.

Instead of using HCUPnet/NIS in the last step as described in Section D.2., we used HCUPnet/NEDS to calculate ER visit rates for states without discharge level or state level data. Table D-6 presents the estimated asthma emergency room rates by health endpoint and age group.

Table D-6. Emergency Department Visit Rates (per 100 people per year) by Health Endpoint and Age Group

Emergency Department Category	ICD-9 Codes	Age 0-17	18-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
Asthma	493	0.959	0.601	0.556	0.538	0.552	0.408	0.331	0.368	0.350
Respiratory	491-493, 460-466, 477.0-477.9, 480-486, 496, 786.07, 786.09	6.069	3.214	2.837	2.332	2.447	2.418	2.908	4.382	5.651
Cardiovascular	410-414, 427-428, 433-437, 440.0-440.9, 443-445, 451-453	0.030	0.107	0.212	0.496	1.151	2.023	3.451	6.726	11.028

D.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 D-7 presents the estimated school loss day rates. Further detail is provided below on these rates.

Table D-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.

D.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 D-8 presents a summary of these baseline rates.

Table D-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	4.37%	American Lung Association (2010a, Table 4). The rate numbers may be slightly different from those in Table 4 because we received more current estimates form ALA.
	18-44		3.15%	
	45-64		5.49%	
	65+		5.63%	
Lower Respiratory Symptoms (LRS)	7-14	Incidence	0.483	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	Adams et al. (1999, Table) U.S.

Endpoint	Age	Parameter	Rate	Source
	18-24		1.971	Bureau of the Census (1997, No.22)
	25-44		2.475	
	45-64		1.796	

NOTE: The incidence rate is the number of cases per person per year. Prevalence refers to the fraction of people that have a particular illness during a particular time period.

D.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.

D.6.2 Chronic Bronchitis Incidence Rate

The annual incidence rate for chronic bronchitis[1] 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.

D.6.3 Chronic Bronchitis Prevalence Rate

We obtained the annual prevalence rate for chronic bronchitis from the American Lung Association (2010a, Table 4). Based on an analysis of 2008 National Health Interview Survey data, they estimated a rate of 0.0437 for persons 18 and older; they also reported the following prevalence rates for people in the age groups 18-44, 45-64, and 65+: 0.0315, 0.0549, and 0.0563, respectively.

D.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 10th = 0 percent, 25th = 0 percent, 50th = 0 percent, 75th = 0.29 percent, and 90th =

---

[1] Please note that this endpoint is not regularly considered in U.S. EPA analyses (July 2018).

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 100th percentiles. Alternatively, assuming a linear slope between the 50th and 75th, 75th and 90th, 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.

D.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.

D.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.

D.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 health impact functions are based on study-specific estimates. The baseline rates for the various endpoints are described below and summarized in Table D-9. The prevalence of asthma is summarized in Table D-10.

Table D-9. Asthma-Related Health Effects Incidence Rates

Endpoint	Age	Parameter	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%	
Upper Respiratory Symptoms (URS)	9-11	Incidence	124.79	Pope et al. (1991, Table 2)

NOTE: The incidence rate is the number of asthma attacks per person per year. Prevalence refers to the fraction of people that have a particular illness during a particular time period.

D.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).

D.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.

D.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.

D.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.

D.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 from an American Lung Association (2010b) report summarizing data from NHIS. Table D-10 presents asthma prevalence rates used to define asthmatic populations in the health impact functions.

Table D-10. Asthma Prevalence Rates Used to Estimate Asthmatic Populations

Population Group	Asthma Prevalence	Source
All Ages	7.80%	American Lung Association (2010b, Table 7)
<5	6.14%	
<18	9.41%	
5-17	10.70%	
18-44	7.19%	
45-64	7.45%	
65+	7.16%	
African-American, <5	9.98%	American Lung Association (2010b, Table 9)
African-American, 5 to 17	17.76%	
African-American, <18	15.53%	American Lung Association*

* Calculated by ALA for U.S. EPA, based on NHIS data (CDC, 2008).

Appendix E.   Core Particulate Matter Health Impact Functions in U.S. Setup

In this Appendix, we present the core PM-related health impact functions in BenMAP, i.e., the functions that, as of the current release, U.S. EPA routinely uses in its regulatory analyses. 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 C 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 D presents a description of the sources for the incidence and prevalence data used in the health impact functions.

E.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 E-1 lists the long-term mortality health impact functions.

Table E-1. Core 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 µg. Unconditional dist. 2% no causality included.
Mortality, All Cause	Expert B	2006		30-99	Annual	0.011950		Log-linear	Range 4 to 10 µg. 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 µg

Effect	Author	Year	Location	Age	Metric	Beta	Std Err	Form	Notes
Mortality, All Cause	Expert F	2006		30-99	Annual	0.009370		Log-linear	Range 4 to 7 µg
Mortality, All Cause	Expert G	2006		30-99	Annual	0.006970		Log-linear	Unconditional dist. 30% no causality included
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 µg. No threshold. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0.003940		Log-linear	Range 4 to 16 µg. Threshold 0 to 5 µg. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0.003940		Log-linear	Range 4 to 16 ug. Threshold 5 to 10 µg. Conditional dist.
Mortality, All Cause	Expert K	2006		30-99	Annual	0			Range >16 to 30. Uniform distribution. No causality.
Mortality, All Cause	Expert K	2006		30-99	Annual	0			Range 4 to 16 ug. Uniform distribution. No causality.
Mortality, All Cause	Expert L	2006		30-99	Annual	0.009340		Log-linear	Range >10 to 30 µg. Unconditional dist. 1% no causality included.
Mortality, All Cause	Expert L	2006		30-99	Annual	0.007390		Log-linear	Range 4 to 10 µg. Unconditional dist. 25% no causality included
Mortality, All Cause	Krewski et al.	2009	116 U.S. cities	30-99	Annual	0.005827	0.000963	Log-linear	
Mortality, All Cause	Lepeule et al.	2012	6 Cities	25-99	Annual	0.013103	0.003347	Log-linear	
Mortality, All Cause	Woodruff et al.	2006	204 counties	Infant	Annual	0.006766	0.007339	Logistic	

E.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 PM$_{2.5}$ 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.

E.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 PM$_{2.5}$. 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 -8 to + 8. The Weibull distribution has support (l + 8), 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;å θ,**min,max**) be a truncated continuous parametric (cumulative) distribution function with (vector of) parameters θ and truncation points **min** and **max**. The nth percentile point is defined as the value xn such that F(xn; θ,**min,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:

$$F\left(xn;\theta,\min,\max\right)=n/100$$
$$F\left(xm;\theta,\min,\max\right)=m/100$$

This is a system of non-linear equations that can be solved for the unknown distribution parameters **θ**. We used the Nelder and Mead (1965) numeric optimization algorithm, available in R, to find the best-fitting estimates of parameters **θ** for the truncated distributions specified by the experts. Once estimates of **θ** were obtained, the distributions were specified fully and we had enough information to make 15,000 draws from each.

Table E-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 E-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	+8	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	+8	mean=? sd=?
	output			1.002	1.590	2.000	2.410	3.000		mean=2 sd=0.608
G	input	Normal	-8			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 µg/m³	input	Normal	-8	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 µg/m³	input	Normal	-8	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:

$$N(0.29; mean = ?, sd = ?, min = 0, max = 4) = 0.05$$
$$N(2.90; mean = ?, sd = ?, min = 0, max = 4) = 0.95$$

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/m^3 change in PM$_{2.5}$. 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/m^3 change in PM$_{2.5}$ is (1+Z/100), and the PM$_{2.5}$ 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 E-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 E-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	1.01489

Expert	Mean	Standard Deviation	Min	P25	P50	P75	Max
K2	0.00237	0.00382	-0.00402	0.00000	0.00000	0.00489	0.01488

E.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/m^3 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 E-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 E-4. Description of the Non-Parametric Expert Functions

Expert	Information	Min	P5	P10	P25	P50	P75	P95	Max
B1 4-10 μg/m^3	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 μg/m^3	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 μg/m^3	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 μg/m^3	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 μg/m^3	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	

Expert	Information	Min	P5	P10	P25	P50	P75	P95	Max
L2 >10-30 µg/m³	input	0.020	0.200		0.570	1.000	1.400	1.600	2.700
	output		0.018		0.568	1.003	1.396	1.634	

Table E-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 E-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)	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

E.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 $PM_{2.5}$); 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 $PM_{2.5}$ 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 $PM_{2.5}$ ranges - and selects from the set appropriate for a given $PM_{2.5}$ 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/m^3 to 5 μg/m^3. Thus, the set of log-linear functions in BenMAP for expert K on the range from 4 μg/m^3 to 16 μg/m^3 contains a function with the:

- PM$_{2.5}$ coefficient = 0 (no causality), which BenMAP selects with 65% probability;

- PM$_{2.5}$ coefficient expert K specified for the log-linear function on that range and no threshold, which BenMAP selects with 18% probability;

- PM$_{2.5}$ coefficient expert K specified for the log-linear function on that range and a threshold (with uniform probability) between 0 μg/m^3 to 5 μg/m^3, which BenMAP selects with 14% probability; and

- PM$_{2.5}$ coefficient expert K specified for the log-linear function on that range and a threshold (with uniform probability) between 5 μg/m^3 to 10 μg/m^3, which BenMAP selects with 4% probability.

If the PM$_{2.5}$ baseline value is greater than 16 μg/m^3, BenMAP goes through an analogous procedure to select a function from among the two functions in that set.

E.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).

E.1.1.4.1   Expert A

Figure E-1. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert A
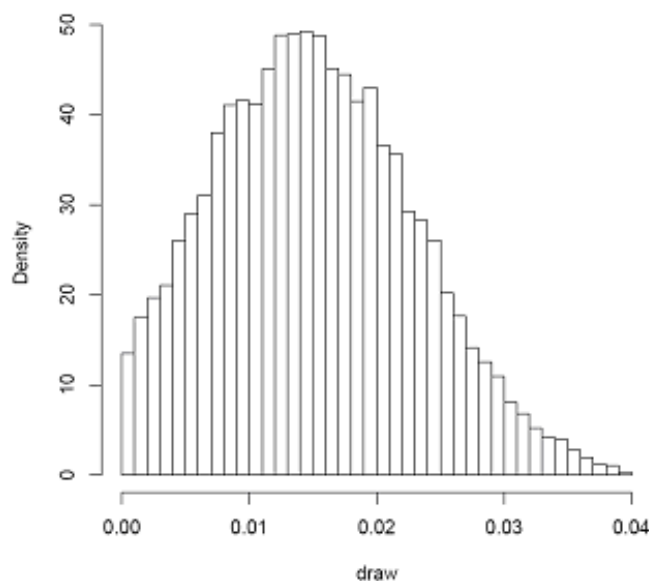
Notes:

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)).

E.1.1.4.2   Expert B

Figure E-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 µg/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 E-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 µg/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)).

E.1.1.4.3 Expert C

Figure E-3. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert C
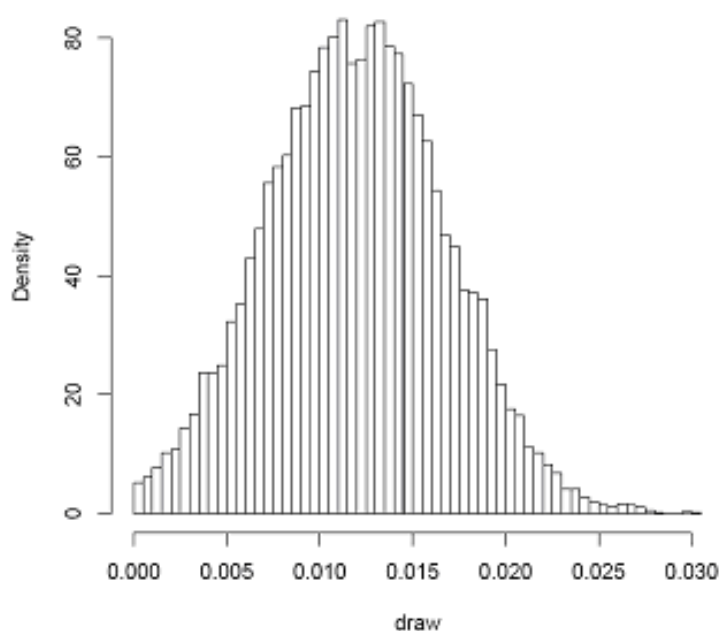
Notes:

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)).

E.1.1.4.4 Expert D

Figure E-4. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert D
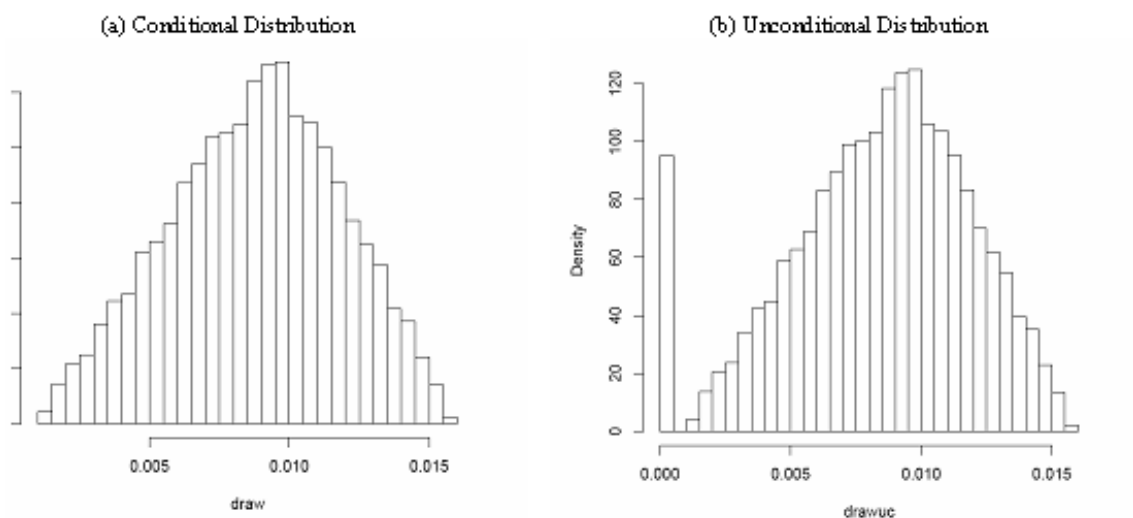
Notes:

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)).

E.1.1.4.5   Expert E

Figure E-5. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert E
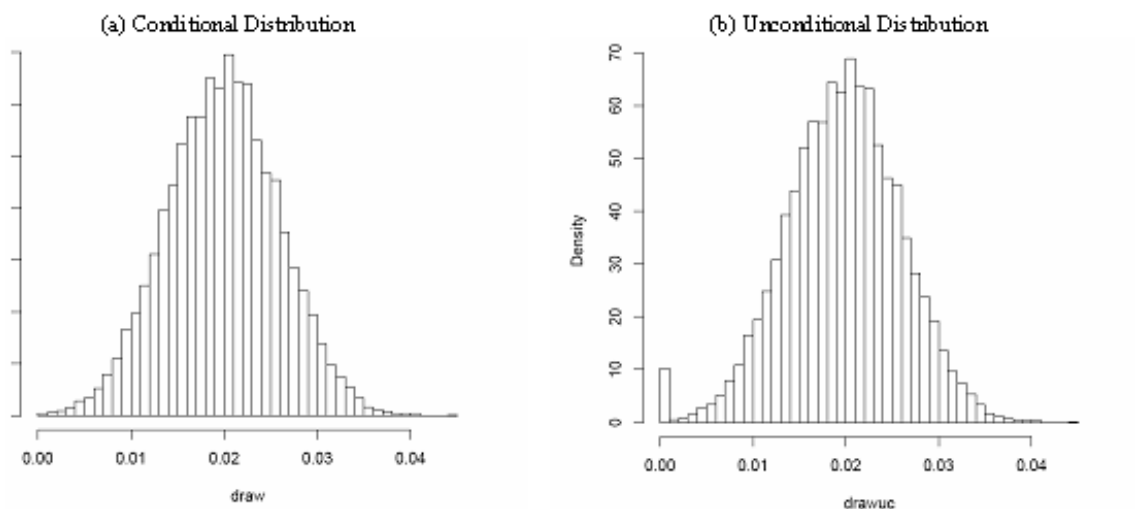
Notes:

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)).

E.1.1.4.6 Expert F

Figure E-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 μg/m^3

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 E-6. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert F (continued)
(2) Results for the range >7-30 μg/m^3
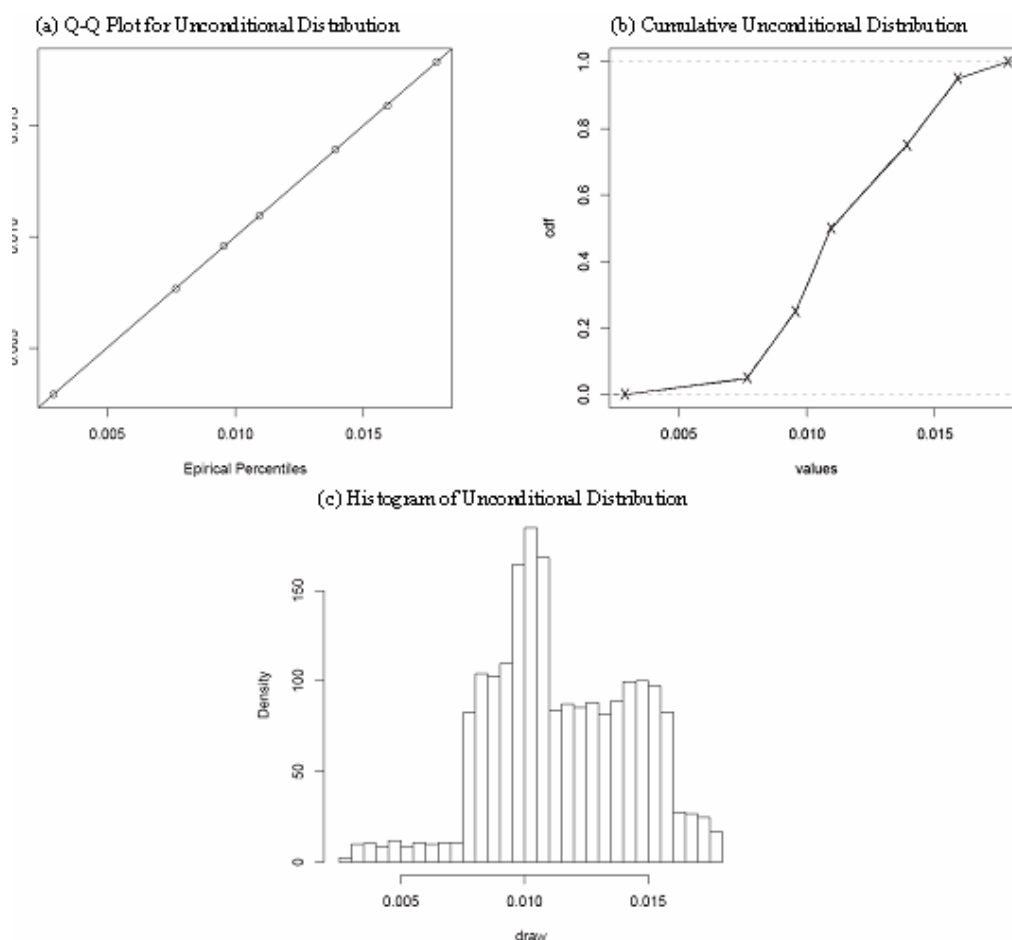
Notes:

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)).

E.1.1.4.7   Expert G

Figure E-7. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert G
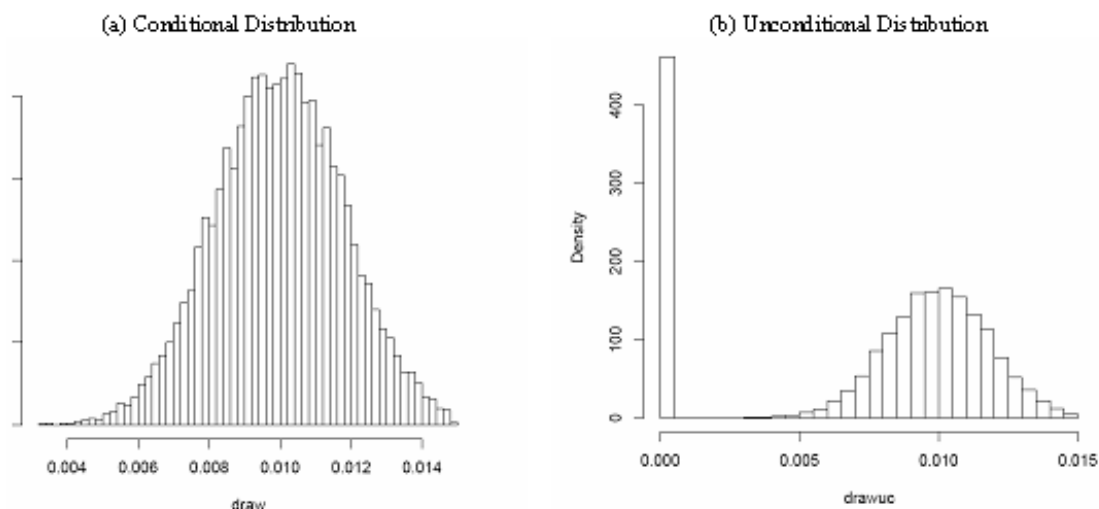
Notes:

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)).

E.1.1.4.8 Expert H

Figure E-8. Characteristics of the Random Draw from the Approximated Distribution of the PM$_{2.5}$ Effect Specified by Expert H
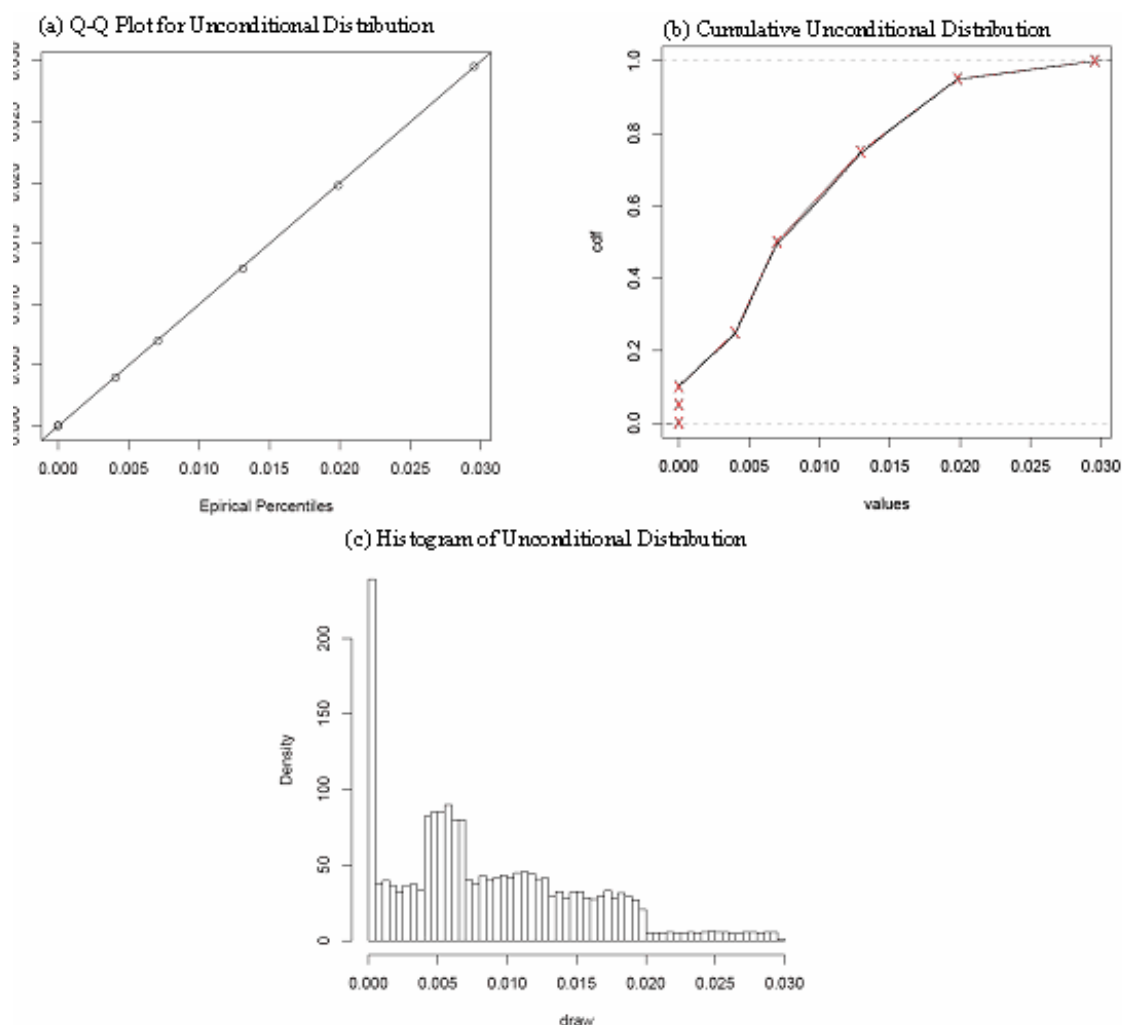
Notes:

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)).

E.1.1.4.9 Expert I

Figure E-9. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert I
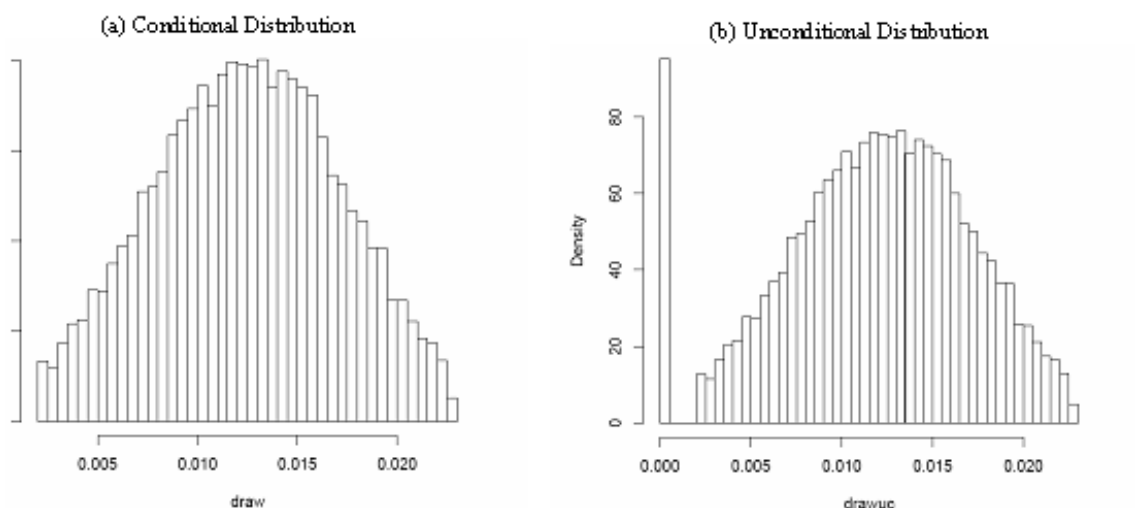
Notes:

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)).

E.1.1.4.10 Expert J

Figure E-10. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert J
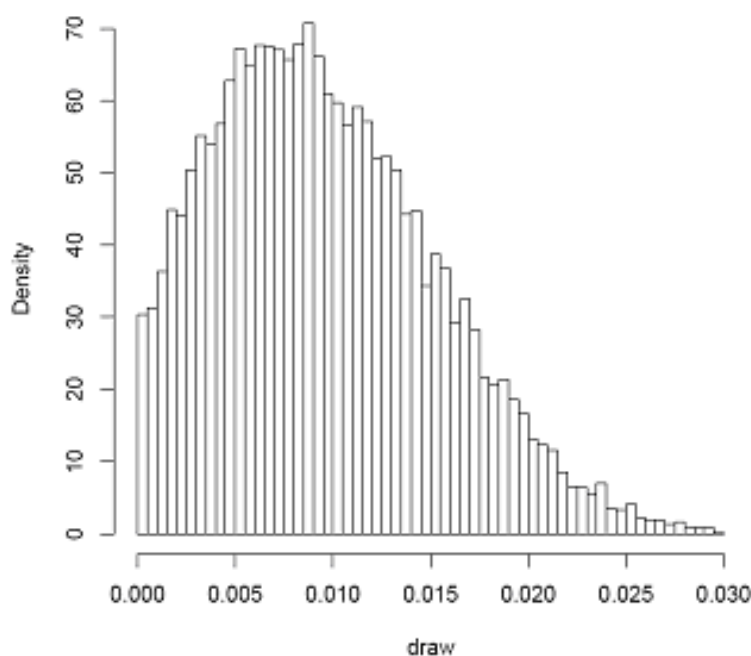
Notes:

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)).

E.1.1.4.11 Expert K

Figure E-11. Histogram of the Random Draw from the Distribution of the PM$_{2.5}$ Effect Specified by Expert K
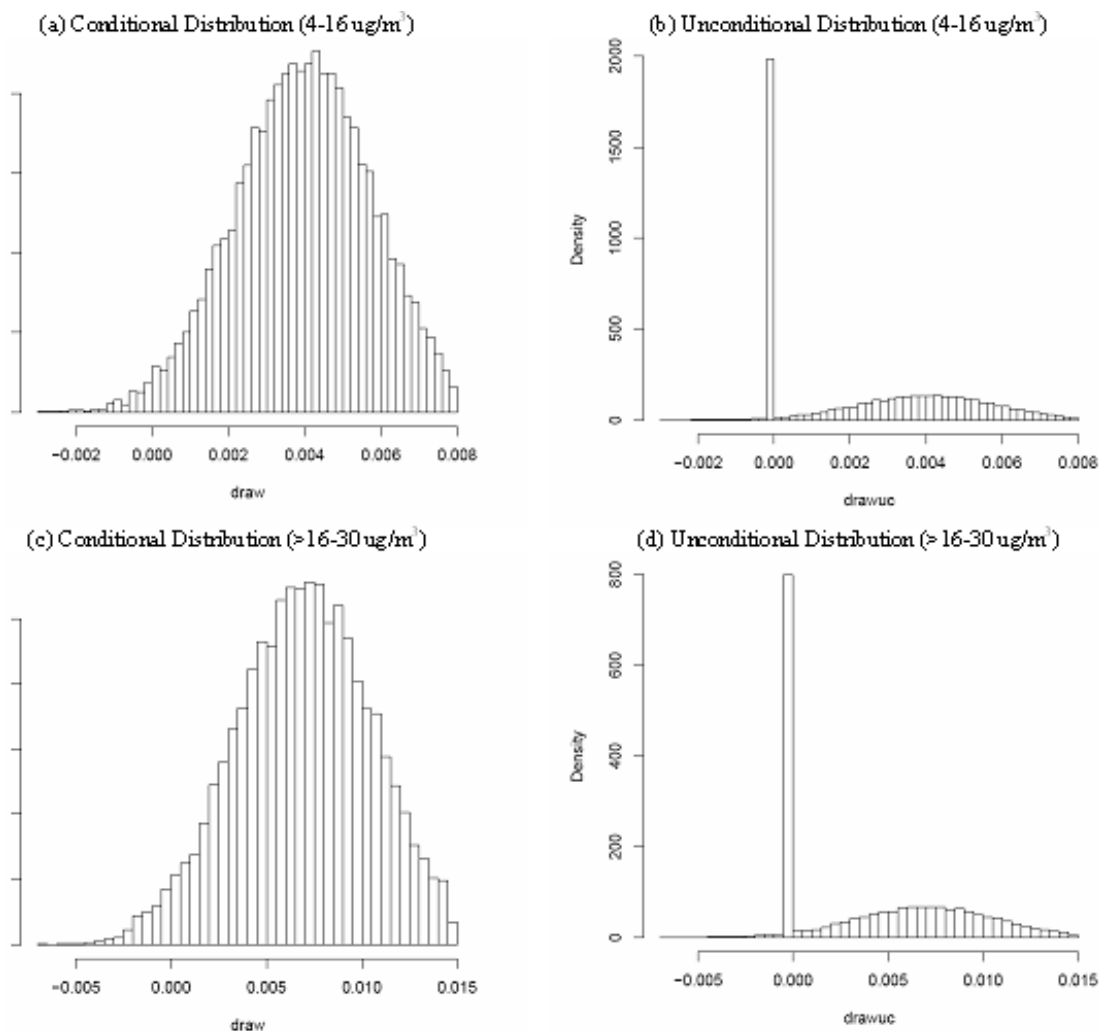
Notes:

Expert K specified a truncated Normal Distribution for two ranges separately (4-16 µg/m^3 and >16-30 µg/m^3). 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)).

E.1.1.4.12 Expert L

Figure E-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 µg/m^3
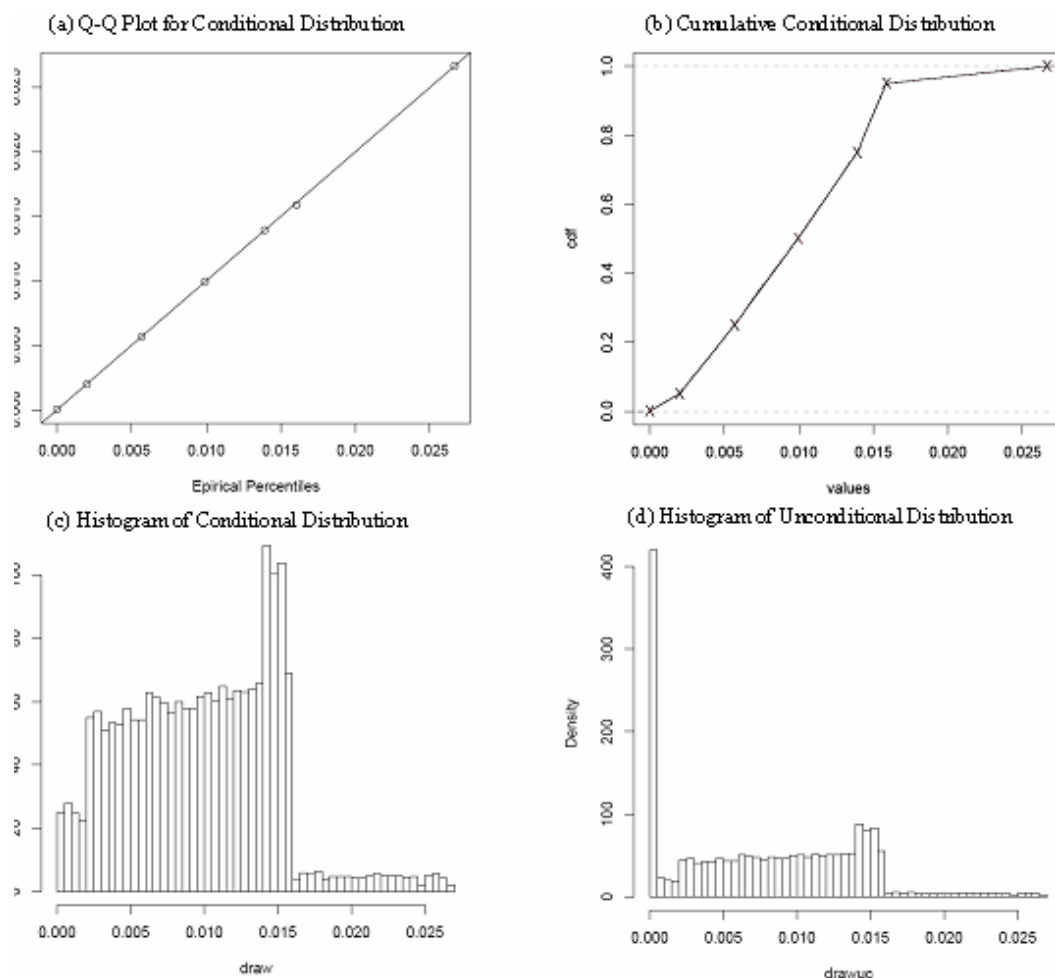
Notes:

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 E-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 µg/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)).

E.1.2 Krewski et al. (2009)

This cohort study consists of approximately 360,000 participants residing in areas of the country that have adequate monitoring information on levels of $PM_{2.5}$ for 1980 and about 500,000 participants in areas with adequate information for 2000. The causes of death that were analyzed included all causes, cardiopulmonary disease (CPD), ischemic heart disease (IHD), lung cancer, and all remaining causes. Data for 44 personal, individual-level covariates, based on participants' answers to a 1982 enrollment questionnaire, were also used for the analyses. The authors also collected data for seven ecologic (neighborhood-level) covariates, each of which represents local factors known or suspected to influence mortality, such as poverty level, level of education, and unemployment (at both Zip Code and city levels). Long-term average exposure variables were constructed for $PM_{2.5}$ from monitoring data for two periods: 1979-1983 and 1999-2000. Similar variables were constructed for long-term exposure to other pollutants of interest from single-year (1980) averages, including total suspended particles, ozone, nitrogen dioxide, and sulfur dioxide. Exposure was averaged for all monitors within a metropolitan statistical area (MSA) and assigned to participants according to their Zip Code area (ZCA) of residence. The authors chose the standard Cox proportional-hazards model (and a variation to allow for random effects) to calculate hazard ratios for various cause-of-death categories associated with the levels of air pollution exposure in the cohort. They extended the random effects Cox model to accommodate two levels of information for clustering and for ecologic covariates. Three main analyses were conducted: a Nationwide Analysis, Intra-Urban Analyses in the New York City (NYC) and Los Angeles (LA) regions, and an analysis designed to investigate whether critical time windows of exposure to pollutants might have affected mortality in the cohort.

Mortality, All-Cause

In a random effects Cox model, the coefficient and standard error in BenMAP-CE are estimated from the relative risks (1.06) and 95% confidence intervals (95% CI: 1.04-1.08) for a 10 µg/m^3 increase in the average of $PM_{2.5}$ exposure level for 1999-2000 (Krewski, et al., 2009, Commentary Table 4). The results were adjusted for the 44 individual-level covariates and the 7 ecologic covariates at the MSA & DIFF levels.

E.1.3 Lepeule et al. (2012)

Lepeule et al. (2012) evaluated the sensitivity of previous Six Cities results to model specifications, lower exposures, and averaging time using eleven additional years of cohort follow-up that incorporated recent lower exposures. The authors found significant associations between $PM_{2.5}$ exposure and increased risk of all-cause, cardiovascular and lung cancer mortality. The authors also concluded that the concentration-response relationship was linear down to $PM_{2.5}$ concentrations of 8 µg/m3, and that mortality rate ratios for $PM_{2.5}$ fluctuated over time, but without clear trends, despite a substantial drop in the sulfate fraction.

E.1.4 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 $PM_{2.5}$ to infants born in California in 1999 and 2000 using maternal addresses for mothers who lived within 5 miles of a $PM_{2.5}$ 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 $PM_{2.5}$ 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 $PM_{2.5}$, 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/m^3 increase in $PM_{2.5}$ 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 the 95% confidence interval (0.93-1.24) associated with a change in annual mean exposure of 10.0 μg/m^3 (Woodruff et al., 2006, p. 786).

E.2 Chronic/Severe Illness

Table E-6 below summarizes the health impacts functions used to estimate the relationship between $PM_{2.5}$ and chronic / severe health effects. We present a brief summary of each of the studies and any items that are unique to the study.

Table E-6. Core Health Impact Functions for Particulate Matter and Chronic Illness

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Acute Myocardinal Infractin, Nonfatal	Peters et al.	2001	Boston, MA	18-99		D24HourMean	0.024121	0.009285	Logistic	
Acute Myocardial Infraction, Nonfatal	Pope et al.	2006	Greater Salt Lake City, UT	0-99		D24HourMean	0.0048	0.0019	Logistic	Index MI and unstable angina
Acute Myocardial Infraction, Nonfatal	Sullivan et al.	2005	King County, WA	0-99		D24HourMean	0.0019	0.0022	Logistic	
Acute	Zanobetti and	2006	Greater Boston,	0-99		D24HourMean	0.0053	0.0022	Logistic	Age range

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Myocardial Infraction, Nonfatal	Schwartz		MA							adjusted Admissions through ER visits only.
Acute Myocardial Infraction, Nonfatal	Zanobetti et al.	2009	26 U.S. Comm	0-99		D24HourMean	0.0022	0.0006	Log-linear	Age range adjusted. All Seasons.

E.2.1 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 was 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 µg/m³ increase in twenty-four hour average $PM_{2.5}$ (Peters et al., 2001, Table 4, p. 2813).

Incidence Rate: We use the county-specific daily AMI hospitalization rate (ICD-9 code 410) for the population of individuals aged 18 years and older as the estimate for the incidence rate of nonfatal heart attack, assuming all heart attacks that are not instantly fatal will result in a hospitalization. We did not adjust for fatal AMIs in the incidence rate estimation, due to the way that the epidemiological studies are designed. Those studies consider total admissions for AMIs, which includes individuals living at the time

the studies were conducted. Therefore, we use the definition of AMI that matches the definition in the epidemiological studies.

Population: Population of ages 18 and older

Adjustment: As some fraction of the admitted individuals die in the hospital, we apply a survival rate of 93% in calculating the avoided cases of AMI in order to avoid double counting (once in the calculation of AMI cases and once in the calculation of PM-related mortality).

E.2.2 Pope et al. (2006)

Pope et al. (2006) evaluated the association between short-term exposure to $PM_{2.5}$ and acute ischemic heart disease events, including acute nonfatal myocardial infarction, all acute coronary events, and subsequent myocardial infarctions in individuals living in greater Salt Lake City, Utah. In a case-crossover study, these ischemic events were assessed in relation to a 10 μg/m^3 increase in $PM_{2.5}$. The researchers determined that a 10 μg/m^3 increase in $PM_{2.5}$ resulted in a 4.5% increase (95% CI: 1.1-8.0) in unstable angina and myocardial infarction.

Acute Myocardial Infarction, Nonfatal

In a single-pollutant model the coefficient and standard error were estimated from the percent increase (4.81%) and 95% confidence interval (95% CI: 0.98-8.79) for a 10 μg/m^3 increase in daily 24-hour mean $PM_{2.5}$ (Pope et al., 2006, Table 3).

Incidence Rate: AMI hospital admission rate for all ages. See the incidence rate discussion under Peters et al. (2001) in Section E.2.1.

Population: All ages

Adjustment: See the adjustment description in Section E.2.1.

E.2.3 Sullivan et al. (2005)

Sullivan et al. (2005) studied the relationship between onset time of acute myocardial infarction and the preceding hourly $PM_{2.5}$ concentrations in 5,793 confirmed cased of myocardial infarction through King County, Washington. In this case-crossover study from 1988-1994, air pollution exposure levels averaged 1 hour, 2 hours, 4 hours, and 24 hours before onset of myocardial infarction were compared to a set of time-stratified referent exposures from the same day of the week in the month of the case event. The authors estimated that an associated risk of 1.01 (95% CI: 0.98-1.05) for myocardial infarction onset could be attributed to a 10 μg/m^3 increase in PM 2.5 the hour before the MI onset. No increased risk was found in all cases with preexisting cardiac diseases with an odds ratio of 1.05 (95% CI: 0.95-1.16). Furthermore, stratification for hypertension, diabetes, and smoking status did not modify the association between $PM_{2.5}$ and onset of myocardial infarction.

Acute Myocardial Infarction, Nonfatal

In a single-pollutant model the coefficient and standard error were estimated from the odds ratio (1.02) and 95% confidence interval (95% CI: 0.98-1.07) for a 10 μg/m^3 increase in daily 24- hour mean PM$_{2.5}$ lagged 1 day (Sullivan et al., 2005, Table 3).

Incidence Rate: AMI hospital admission rate for all ages. See the incidence rate discussion under Peters et al. (2001) in Section E.2.1.

Population: All ages

Adjustment: See the adjustment description in Section E.2.1.

E.2.4 Zanobetti and Schwartz (2006)

Zanobetti and Schwartz (2006) analyzed hospital admissions through emergency department for myocardial infarction (ICD-9 code 410) and pneumonia (ICD-9 codes 480-487) for associations with fine particulate air pollution, ozone, black carbon, nitrogen dioxide, PM not from traffic, and CO in the greater Boston area from 1995-1999. The authors used a case- crossover analysis with control days matched on temperature. Significant associations were detected for NO$_2$ with a 12.7% increase 95% CI: 5.8-18.0), PM$_{2.5}$ with an 8.6% increase (95% CI: 1.2-15.4), and black carbon with an 8.3% increase (95% CI: 0.2-15.8) in emergency myocardial infarction hospitalizations. Similarly, significant associations were identified for PM$_{2.5}$ with a 6.5% increase (95% CI: 1.1-11.4) and CO with a 5.5% increase (95% CI: 1.1-9.5) in pneumonia hospitalizations.

Acute Myocardial Infarction, Nonfatal

The study looked at hospital admissions of AMI through the ER. Under the assumption that all heart attacks will end in hospitalization, we consider the endpoint as heart attack events to be consistent with other studies. In a single-pollutant model, the coefficient and standard error are estimated from the percent change in risk (8.65%) and 95% confidence interval (95% CI: 1.22-15.38%) for a 16.32 μg/m^3 increase in daily 24-hour mean PM$_{2.5}$ for an average of the 0- and 1-day lag (Zanobetti A. and Schwartz, 2006, Table 4).

Incidence Rate: AMI hospital admission rate for all ages. See the incidence rate discussion under Peters et al. (2001) in Section E.2.1.

Population: All ages. Note that although Zanobetti and Schwartz (2006) reports results for the 65-99 year old age range, for comparability to other studies, we apply the results to all ages. Since the vast majority of AMIs occur among population 65-99, over-counting may not be an issue when applying the risk coefficient to all ages.

Adjustment: See the adjustment description in Section E.2.1.

E.2.5 Zanobetti et al. (2009)

Zanobetti et al. (2009) examined the relationship between daily PM$_{2.5}$ levels and emergency hospital admissions for cardiovascular causes, myocardial infarction, congestive heart failure, respiratory disease, and diabetes among 26 U.S. communities from 2000-2003. The authors used meta-regression to examine how this association was modified by season- and community-specific PM$_{2.5}$ composition while controlling for seasonal temperature as a substitute for ventilation. Overall, the authors found that PM$_{2.5}$ mass higher in Ni, As, and Cr as well as Br and organic carbon significantly increased its effects on hospital admissions. For a 10 μg/m^3 increase in 2-day averaged PM$_{2.5}$, a 1.89% (95% CI: 1.34-2.45) increase in cardiovascular disease admissions, a 2.25% (95% CI: 1.10-3.42) increase in myocardial infarction admissions, a 1.85% (95% CI: 1.19-2.51) increase in congestive heart failure admissions, a 2.74% (95% CI: 1.30-4.20) increase in diabetes admissions, and a 2.07% (95% CI: 1.20-2.95) increase in respiratory admissions were observed. The relationship between PM$_{2.5}$ and cardiovascular admissions was significantly modified when the mass of PM$_{2.5}$ was high in Br, Cr, Ni, and sodium ions, while mass high in As, Cr, Mn, organic carbon, Ni and sodium ions modified the myocardial infarction relationship and mass high in As, organic carbon, and sulfate ions modified the diabetes admission rates.

Acute Myocardial Infarction, Nonfatal

The study looked at hospital admissions of AMI through ER. Under the assumption that all heart attacks will end in hospitalization, we consider the endpoint as heart attack events to be consistent with other studies. In a single-pollutant model the coefficient and standard error are estimated from the percent change in risk (2.25%) and 95% confidence interval (95% CI: 1.10-3.42) for a 10 μg/m^3 increase in 2-day averaged PM$_{2.5}$ (Zanobetti et al., 2009, Table 3).

Incidence Rate: AMI hospital admission rate for all ages. See the incidence rate discussion under Peters et al. (2001) in Section E.2.1.

Population: All ages. Note that although Zanobetti et al. (2009) reports results for the 65-99 year old age range, for comparability to other studies, we apply the results to all ages. Since the vast majority of AMIs occurs among population 65-99, over-counting may not be an issue when applying the risk coefficient to all ages.

Adjustment: See the adjustment description in Section E.2.1.

E.3   Hospitalizations

Table E-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 E-7. Core Health Impact Functions for Particulate Matter and Hospital Admissions

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma	Babin et al.	2007	Washington, DC	0-17		D24HourMean	0.0020	0.0043	Log-linear	Age range adjusted Admission from emergency department only.
All Cardiovascular (less Myocardial Infarctions)	Bell et al.	2008	202 U.S. Counties	65-99		D24HourMean	0.0008	0.0001	Log-linear	Urgent admission only. Yearly national estimates
All Cardiovascular (less Myocardial Infarctions)	Bell et al.	2012	187 U.S. Counties	65-99		D24HourMean	0.0008	0.0001071	Log-linear	
All Respiratory	Kloog et al.	2012	New England	65-99		D24HourMean	0.0007	0.0010	Log- linear	
All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.001400	0.000341	Log-linear	
All Cardiovascular (less Myocardial Infarctions)	Peng et al.	2008	108 U.S. Counties	65-99		D24HourMean	0.0007	0.0001	Log-linear	Emergency HA
All Cardiovascular (less Myocardial Infarctions)	Peng et al.	2009	119 U.S. Counties	65-99		D24HourMean	0.0007	0.0002	Log-linear	Urgent or emergency HA
Asthma	Sheppard	2003	Seattle, WA	0-64		D24HourMean	0.003324	0.001045	Log-linear	
All Cardiovascular (less Myocardial Infarctions)	Zanobetti et al.	2009	26 U.S. Communities	65-99		D24HourMean	0.00019	0.0003	Log-linear	All seasons
All Respiratory	Zanobetti et al.	2009	26 U.S. Communities	65-99		D24HourMean	0.0021	0.0004	Log-linear	All seasons

E.3.1   Babin et al. (2007)

Babin et al. (2007) examined pediatric asthma-related emergency room (ER) visits and hospital admissions (ICD-9 code 493) in Washington, D.C. from 2001-2004 and their short-term associations with ozone, particulate matter, socioeconomic status, and age group. The association between $PM_{2.5}$ and asthma hospitalization was found statistically insignificant.

Hospital Admissions, Asthma (ICD-9 code 493)

In a single-pollutant model, the coefficient and standard error are estimated from the average percent increase in risk (0.2%) and 95% confidence interval (95% CI: -0.6% -

1.1%) for a 1 μg/m^3 increase in same-day daily 24-hour mean PM$_{2.5}$ (Babin et al., 2007, Table 2).

Note that although Babin et al. (2007) reports results for the 1-17 year old age range, for comparability to other studies, we apply the results to the population of ages 0 to 17.

E.3.2   Bell et al. (2008, 2012)

Bell et al. (2008) evaluated the association between short-term exposure to PM$_{2.5}$ and the risk of cardiovascular (ICD-9 codes 410-414, 26-427, 428, 429, 430-438, and 440-449) and respiratory (ICD-9 codes 464-466, 480-487, and 490-492) hospital admissions among Medicare enrollees =65 years old varied by season and geographic region in 202 U.S. counties with populations greater than 200,000 from 1999-2005. Three time-series models were used to provide three key variables: consistent PM effects across the year, different PM effects by season, and smoothly varying PM effects throughout the year. A two-stage Bayesian hierarchical model was used to estimate the association between PM$_{2.5}$ and hospitalization rates, with the first stage estimating the association within a single county and the second stage combining county- specific estimates. The authors found statistically significant evidence of seasonal and regional variation. Respiratory hospitalizations were highest in winter with a 1.05% increase (95%PI: 0.29-1.82) in hospitalizations per 10 μg/m^3 increase in same-day PM$_{2.5}$. A 1.49% increase (95% PI: 1.09-1.89) in cardiovascular hospital admissions were also found for the winter season, and associations were observed in other seasons as well. The strongest association was for the northeast for both respiratory and cardiovascular admissions.

Hospital Admissions, Cardio-, Cerebro- and Peripheral Vascular Disease (ICD-9 codes 426-427, 428, 430-438, 410-414, 429; 440-449)

For different seasons (i.e., autumn, spring, summer, winter, and all-year) and regions (i.e., southwest, northwest, southeast, southwest, and nationwide), the coefficient and standard error are estimated from the average percent increase in risk and 95% confidence interval for a 10 μg/m^3 increase in same-day (lag 0) daily 24-hour mean PM$_{2.5}$ (Bell et al., 2008, Table 2).

Note that Bell et al. (2008) considered a broader range of ICD-9 codes and estimated the risk of both cardiovascular events and cerebro- and peripheral vascular disease. For comparability to other studies, EPA decided to apply a baseline hospitalization rate for ICD-9 codes 390-409 and 411-429 when using this C-R function in quantifying impacts.

E.3.3   Kloog et al. (2012)

Kloog et al. (2012) examined the relationship between both short-term and long-term PM$_{2.5}$ exposure and emergent hospital admissions for respiratory (ICD-9 codes 460-519) and cardiovascular (ICD-9 codes 390-429) diseases in the New-England states of Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont. The study used hospital admission data from the U.S. Medicare program for

only patients that were at least 65 years old. Short-term exposure models looked at the mean 24-hour PM$_{2.5}$ concentration on the day of a patient's hospital admission, while long-term exposure was defined to be the mean PM$_{2.5}$ concentration over the entire study period (2000-2006). Results of the study showed an association between PM$_{2.5}$ and hospital admissions for all disease types in both the short and long term.

Hospital Admissions, Respiratory (ICD Codes 460-519)

In a single pollutant model for patients over the age of 65, the coefficient and standard error were estimated from the percent change (0.70%) and 95% confidence interval (0.35%-1.05%) for a 10 µg/m3 increase in same day (0 lag) 24-hour mean PM$_{2.5}$ concentration.

E.3.4 Moolgavkar (2000b), Cardiovascular

Moolgavkar (2000b) examined the association between air pollution and cardiovascular hospital admissions (ICD 390-429) 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, 411-429)

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/m^3 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.

E.3.5 Peng et al. (2008)

Peng et al. (2008) examined the risk of hospital admissions for cardiovascular and respiratory diseases in relation to particulate matter (PM$_{10}$-2.5 and PM$_{2.5}$). To accomplish this, the authors utilized a database of 108 U.S. counties with daily

emergency hospital admission rates for cardiovascular and respiratory diseases among Medicare enrollees living 9 miles from air monitors, temperature, and dew-point temperature. PM_{10}-2.5 and $PM_{2.5}$ concentrations were calculated by using monitoring data from January 1, 1999 through December 31, 2005. Overall, there were 3.7 million cardiovascular disease and 1.4 million respiratory disease-related hospital admissions for the time period assessed. The authors found that a 10 μg/m³ increase in PM_{10}-2.5 was associated with a 0.36% increase (95% PI: 0.05-0.68%) in cardiovascular disease admissions on the same day, and a 0.25% increase (95% PI: -0.11-0.60%) after adjusting for $PM_{2.5}$. For respiratory disease admissions, a 10 μg/m³ increase in PM_{10}-2.5 was found to be associated with an unadjusted 0.33% increase in respiratory disease admissions (95% PI: -0.21- 0.86%) and an adjusted 0.26% increase (95% PI: -0.32-0.84%) in emergency admissions. Also, unadjusted associations of $PM_{2.5}$ with cardiovascular and respiratory disease admissions were 0.71% (95% PI: 0.45-0.96%) for same-day exposure and 0.44% (95% PI: 0.06-0.82%) for exposure lagged by 2 days prior to hospital admission.

Hospital Admissions, Cardio-, Cerebro-, and Peripheral Vascular Disease (ICD-9 codes 426-427, 428, 430-438, 410-414, 429, 440-448)

In a single-pollutant model, the coefficient and standard error are estimated from the percent change in daily admission (0.44%) and 95% posterior interval (95% PI: 0.06-0.82%) for a 10 μg/m³ increase in daily 24-hour mean $PM_{2.5}$ concentrations for the same day (Peng et al., 2008, page 2175).

Note that Peng et al. (2008) considered a broader range of ICD-9 codes and estimated the risk of both cardiovascular events and cerebro- and peripheral vascular disease. For comparability to other studies, EPA decided to apply a baseline hospitalization rate for ICD-9 codes 390-409 and 411-429 when using this C-R function in quantifying impacts.

E.3.6   Peng et al. (2009)

Peng et al. (2009) investigated the relationship between hospital admissions for cardiovascular and respiratory disease and the chemical components of $PM_{2.5}$ across 119 U.S. urban communities for 12 million Medicare enrollees using log-linear Poisson regression models. This was achieved using a national database with daily data from 2000-2006 on emergency hospital admissions of cardiovascular and respiratory outcomes, ambient levels of $PM_{2.5}$ components and weather variables. Bayesian hierarchical statistical models were used to estimate the associations. Three scenarios for $PM_{2.5}$ exposure were assessed which were as follows: 1) for the period 2000-2006 and including only days with available measurements for all 7 $PM_{2.5}$ components from the Speciation Trends network (STN); 2) $PM_{2.5}$ measured by the STN for the period 2000-2006 and including only days with available measurements for all 7 $PM_{2.5}$ components from the STN and 3) $PM_{2.5}$ estimated as the sum of the 7 largest components of $PM_{2.5}$ mass for the period 2000-2006. Results of percent increases in emergency admissions associated with $PM_{2.5}$ at lag 0 under these scenarios were showed in Figure 2 and the results for the components of $PM_{2.5}$ from both single and multi-pollutant models were showed in Figure 3. In multi-pollutant models that

adjusted for the levels of other pollutants, the authors found that an interquartile range increase in elemental carbon was associated with a 0.80% increase (95% PI: 0.34-1.27%) in risk of same-day cardiovascular admissions. Similarly, an interquartile range increase in organic carbon matter was associated with a 1.01% increase (95% PI: 0.04-1.98%) risk of respiratory admissions on the same day.

Hospital Admissions, Cardio-, Cerebro-, and Peripheral Vascular Disease (ICD-9 codes 426-427, 428, 430-438, 410-414, 429, 440-448)

In a single-pollutant model, the coefficient and standard error are estimated from the percent change in daily admission (0.68%) and 95% posterior interval (95% PI: 0.26-1.10%) for a 10 μg/m^3 increase in daily 24-hour mean $PM_{2.5}$ concentrations for the same day (Peng et al., 2009, page 960).

Note that Peng et al. (2009) considered a broader range of ICD-9 codes and estimated the risk of both cardiovascular events and cerebro- and peripheral vascular disease. For comparability to other studies, EPA decided to apply a baseline hospitalization rate for ICD-9 codes 390-409 and 411-429 when using this C-R function in quantifying impacts.

E.3.7   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. $PM_{2.5}$ 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/m^3 increase in $PM_{2.5}$ in the 1-day lag GAM stringent model (Sheppard, 2003, pp. 228-229).

E.3.8   Zanobetti et al. (2009)

Zanobetti et al. (2009) examined the relationship between daily $PM_{2.5}$ levels and emergency hospital admissions for cardiovascular causes, myocardial infarction, congestive heart failure, respiratory disease and diabetes among 26 U.S. communities

from 2000-2003. The authors used meta-regression to examine how this association was modified by season- and community-specific $PM_{2.5}$ composition while controlling for seasonal temperature as a substitute for ventilation. Overall, the authors found that $PM_{2.5}$ mass higher in Ni, As, and Cr as well as Br and organic carbon significantly increased its effects on hospital admissions. For a 10 µg/m³ increase in 2-day averaged $PM_{2.5}$, the authors found a 1.89% (95% CI: 1.34-2.45) increase in cardiovascular disease admissions, a 2.25% (95% CI: 1.10-3.42) increase in myocardial infarction admissions, a 1.85% (95% CI: 1.19-2.51) increase in congestive heart failure admissions, a 2.74% (95% CI: 1.30-4.20) increase in diabetes admissions, and a 2.07% (95% CI: 1.20-2.95) increase in respiratory admissions. The relationship between $PM_{2.5}$ and cardiovascular admissions was significantly modified when the mass of $PM_{2.5}$ was high in Br, Cr, Ni, and sodium ions, while mass high in As, Cr, Mn, organic carbon, Ni and sodium ions modified the myocardial infarction relationship and mass high in As, organic carbon, and sulfate ions modified the diabetes admission rates.

Hospital Admissions, All Cardiovascular (ICD-9 codes 390-429)

In a single-pollutant model, the coefficient and standard error are estimated from the percent change in risk (1.89%) and 95% confidence interval (1.34%-2.45%) for a 10 µg/m³ increase in 2-day averaged $PM_{2.5}$ (Zanobetti et al., 2009, Table 3).

Note that Zanobetti et al. (2009) report results for ICD codes 390-429. In the benefit analysis, avoided nonfatal heart attacks are estimated separately. In order to avoid double counting heart attack hospitalizations, we have excluded ICD code 410 from the baseline incidence rate used in this function.

Hospital Admissions, All Respiratory (ICD-9 codes 460-519)

In a single-pollutant model, the coefficient and standard error are estimated from the percent change in risk (2.07%) and 95% confidence interval (1.2% - 2.95%) for a 10 µg/m³ increase in 2-day averaged $PM_{2.5}$ (Zanobetti et al., 2009, Table 3).

E.4   Emergency Room Visits

Table E-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 E-8. Core Health Impact Functions for Particulate Matter and Emergency Room Visits

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma	Glad et al.	2012	Pittsburgh, PA	0-99		D24HourMean	0.00392	0.00284	Logistic	All Races
Asthma	Mar et al.	2010	Greater Tacoma, WA	0-99		D24HourMean	0.0056	0.0021	Log-linear	
Asthma	Slaughter	2005	Spokane, WA	0-99		D24HourMean	0.0029	0.0027	Log-	

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
	et al.								linear	

E.4.1 Glad et al. (2012)

Glad et al. (2012) examined the relationship between air pollution and emergency department visits for asthma (ICD-9 code 493) in six Pittsburgh, PA hospitals between 2002 and 2005. The study includes a total of 6,979 individuals with a primary discharge diagnosis of asthma. Using a case-crossover methodology, particulate matter with an aerodynamic diameter ≤2.5μm ($PM_{2.5}$) had an effect both on the total population on day 1 after exposure (1.036, p <0.05), and on African Americans for day-1 lag (OR =1.055; 95% CI, 1.001–1.112), day-2 lag (OR=1.067; 95% CI, 1.015–1.122), day-3 lag (OR=1.053; 95% CI, 1.002–1.106), and the 6-day average (OR =1.088; 95% CI, 1.001–1.184). $PM_{2.5}$ had no significant effect on Caucasian Americans alone. The disparity in risk estimates by race may reflect differences in residential characteristics, exposure to ambient air pollution, or a differential effect of pollution by race.

Emergency Room Visits, Asthma (ICD-9 Code 493) – by Age and Racial Group

In a single pollutant model for patients of all ages and all races, the coefficient and standard error were estimated from the odds ratio (1.04) and 95% confidence interval (0.984 – 1.10) for a 10 μg/m3 increase in 24-hour mean $PM_{2.5}$ concentration averaged over lags 0-6 days (Glad et al., 2012, Table 3).

E.4.2 Mar et al. (2010)

Mar et al. (2010) assessed the effect of particulate matter air pollution, including emissions from diesel generators, on emergency room visits for asthma in the greater Tacoma, Washington area from January 3, 1998, to May 30, 2002, using Poisson regression models. Health data were collected for individuals of all ages from 6 Tacoma hospitals. The authors also assessed the impacts of diesel generator use on emergency room visits for asthma from January 24, 2001, to June 2, 2001. Overall, the researchers found an association between daily $PM_{2.5}$ levels and emergency room visits for asthma at lag days 2 and 3, with a relative risk for lag day 2 of 1.04 (95% CI: 1.01-1.07) and a relative risk for lag day 3 of 1.03 (95% CI: 1.0-1.06). No significant association between emergency room visits for asthma and increased use of the diesel generators was observed.

Emergency Room Visits, Asthma (ICD-9 code not reported)

In a single-pollutant model, the coefficient and standard error are estimated from the relative risk (1.04) and 95% confidence interval (95% CI: 1.01-1.07) for a 7 μg/m^3 increase in daily 24-hour mean $PM_{2.5}$ at lag day 2 (Mar et al., 2010, Table 4).

E.4.3   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 µg/m^3 increase in $PM_{2.5}$ (Norris et al., 1999, p. 491).

E.4.4   Slaughter et al. (2005)

Slaughter et al. (2005) examined the short-term association of particulate matter (PM1, $PM_{2.5}$, PM_{10}, and PM_{10}-2.5) and carbon monoxide with hospital admissions and emergency room visits for respiratory and cardiac outcomes and mortality in Spokane, Washington, from January 1995 to June 2001 using a log-linear generalized linear model. The authors found no association between respiratory emergency room visits and any size fraction of PM, but there was a suggestive relationship between fine PM and respiratory effects when compared to coarse PM. No association between cardiac hospital admissions or mortality and any size fraction of PM or CO was observed at the 0- to 3-day lag. CO, on the other hand, was found to be associated with all respiratory emergency room visits and visits for asthma at the 3-day lag.

Emergency Room Visits, Asthma (ICD-9 code 493)

In a single-pollutant model, the coefficient and standard error are estimated from the relative risk (1.03) and 95% confidence interval (95% CI: 0.98-1.09) for a 10 µg/m^3 increase in daily 24-hour mean $PM_{2.5}$ at 1-day lag (Slaughter et al., 2005, Table 4).

E.5   Minor Effects

Table E-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 E-9. Core Health Impact Functions for Particulate Matter and Minor Effects

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form
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
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	0.019012	0.006005	Logistic

E.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 PM2.1 and PM_{10} were marginally significantly related to bronchitis. The original study measured PM2.1, however when using the study's results we use $PM_{2.5}$. This makes only a negligible difference, assuming that the adverse effects of PM2.1 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 PM15. 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

(PM2.1 = 20.7 μg/m^3) versus the least polluted city (PM2.1 = 5.8 μg/m^3) (Dockery et al., 1996, Tables 1 and 4). The original study used PM2.1, however, we use the PM2.1 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.

E.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:

$$\sigma_\beta^2 = \mathrm{var}\left(\frac{\sum\limits_{i-1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = \mathrm{var}\left(\frac{\sum\limits_{i-1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\gamma} \right) = \sum\limits_{i-1976}^{1981} \mathrm{var}\left(\frac{\beta_i}{\sigma_{\beta_i}^2 \times \gamma} \right)$$

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

E.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 the period 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:

$$\beta = \left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = 0.00741.$$

The standard error of the coefficient is calculated as follows, assuming that the estimated year-specific coefficients are independent:

$$\sigma_\beta^2 = \text{var}\left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = \text{var}\left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\gamma} \right) = \sum\limits_{i=1976}^{1981} \text{var}\left(\frac{\beta_i}{\sigma_{\beta_i}^2 \times \gamma} \right)$$

This reduces down to:

$$\sigma_\beta^2 = \frac{1}{\gamma} \Rightarrow \sigma_\beta = \sqrt{\frac{1}{\gamma}} = 0.00036.$$

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

E.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

E.6   Asthma-Related Effects

Table E-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 E-10. Core Health Impact Functions for Particulate Matter and Asthma-Related Effects

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Cough	Mar et al.	2004	Spokane, WA	6-18		D24HourMean	0.0191	0.0098	Logistic	Uses incidence rate from Ostro et al. (2001). Age range adjusted.
Shortness of Breath	Mar eta al.	2004	Spokane, WA	6-18		D24HourMean	0.0122	0.0138	Logistic	Uses incidence rate from Ostro et al. (2001). Age range adjusted.
Cough	Ostro et al.	2001	Los Angeles, CA	6-18		D24HourMean	0.000985	0.000747	Logistic	
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	0.001942	0.000803	Logistic	
Upper Respiratory Symptoms	Pope et al.	1991	Utah Valley	9-11		D24HourMean	0.003600	0.001500	Logistic	

E.6.1   Mar et al. (2004)

Mar et al. (2004) studied the effects of various size fractions of particulate matter on respiratory symptoms of adults and children with asthma, monitored over many months. The study was conducted in Spokane, Washington, a semiarid city with diverse sources of particulate matter. Data on respiratory symptoms and medication use were recorded daily by the study's subjects, while air pollution data was collected by the local air agency and Washington State University. Subjects in the study consisted of 16 adults—the majority of whom participated for over a year—and nine children, all of whom were studied for over eight months. Among the children, the authors found a strong association between cough symptoms and several metrics of particulate matter, including $PM_{2.5}$. However, the authors found no association between respiratory symptoms and PM of any metric in adults. Mar et al. therefore concluded that the discrepancy in results between children and adults was due either to the way in which air quality was monitored, or a greater sensitivity of children than adults to increased levels of PM air pollution.

Asthma Exacerbation, Cough

In a single-pollutant model, the coefficient and standard error are estimated from the odds ratio (1.21) and 95% confidence interval (1.00-1.47) for a 10.0 μg/m³ increase in 1-day lagged concentration of $PM_{2.5}$ (Mar et al., 2004, Table 7).

Incidence Rate: Daily cough rate per person = 14.5%. Mar et al. (2004) did not report the incidence rate for each type of asthma exacerbation. The daily cough rate from Ostro et al. (2001, p.202) is applied here.

Population: The study reported results for population ages 7-12. For comparability to other studies, we apply the results to the population of ages 6 to 18. We treat this as two groups based on the available information from American Lung Association (2010b, Table 7). Asthmatic population ages 6 to 17 = 10.70% of population ages 6 to 17 and asthmatic population age 18 = 7.19% of population age 18. The American Lung Association (2010b, Table 7) estimates asthma prevalence for children 5- 17 and adults 18-44 at 10.70% and 7.19% respectively (based on data from the 2008 National Health Interview Survey).

Asthma Exacerbation, Shortness of Breath

In a single-pollutant model, the coefficient and standard error are estimated from the odds ratio (1.13) and 95% confidence interval (0.86-1.48) for a 10.0 $\mu g/m^3$ increase in current-day concentration of $PM_{2.5}$ (Mar et al., 2004, Table 7).

Incidence Rate: Daily shortness of breath rate per person = 7.4%. Mar et al. (2004) did not report the incidence rate for each type of asthma exacerbation. The daily rate of shortness of breath from Ostro et al. (2001, p.202) is applied here.

Population: See the population description for "Asthma Exacerbation, Cough" from Mar et al. (2004).

E.6.2   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 µg/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: The study reported results for African-American population ages 8-13. For comparability to other studies, we apply the results to the African-American population of ages 6 to 18. We treat this as two groups based on the available information from American Lung Association (2010b, Table 9). Asthmatic African-American population ages 6 to 17 = 17.76% of African-American population ages 6 to 17 and asthmatic African-American population age 18 = 7.52% of African-American population age 18. The American Lung Association (2010b, Table 9) estimates asthma prevalence for children 5- 17 and adults 18-44 at 17.76% and 7.52% respectively (based on data from the 2008 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 µg/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: The study reported results for African-American population ages 8-13. For comparability to other studies, we apply the results to the African-American population of ages 6 to 18. We treat this as two groups based on the available information from American Lung Association (2010b, Table 9). Asthmatic African-American population ages 6 to 17 = 17.76% of African-American population ages 6 to 17 and asthmatic African-American population age 18 = 7.52% of African-American population age 18. The American Lung Association (2010b, Table 9) estimates asthma prevalence for children 5- 17 and adults 18-44 at 17.76% and 7.52% respectively (based on data from the 2008 National Health Interview Survey).

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 µg/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 = 17.76% of African-American population ages 8 to 13. (Described above.)

E.6.3   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 = 10.70% of population ages 9 to 11. (The American Lung Association (2010b, Table 7) estimates asthma prevalence for children ages 5 to 17 at 10.70%, based on data from the 2008 National Health Interview Survey.)

Appendix F.   Core Ozone Health Impact Functions in U.S. Setup

In this Appendix, we present the core health impact functions used to estimate ozone-related adverse health effects, i.e., the functions that, as of the current release, U.S. EPA routinely uses in its regulatory analyses. 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 C 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 D presents a description of the sources for the incidence and prevalence data used in the health impact functions.

F.1   Short-term Mortality

Table F-1 summarizes the core 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 F-1. Core Health Impact Functions for Ozone and Short-Term Mortality

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Non-Accidental	Smith et al.	2009	98 U.S. cities	0-99	PM10	D8HourMax	0.000258	0.000167	Log-linear	Ozone season
All-Cause	Zanobetti and Schwartz	2008 b	48 Cities	0-99		D8HourMax	0.00050	0.00012	Log-linear	0-3 day lag, June-August, 1989-2000

F.1.1   Smith et al. (2009)

Smith et al. (2009) analyzed the relationship between daily mortality and ambient ozone concentrations through re-examination of evidence using the National Morbidity, Mortality, and Air Pollution Study (NMMAPS), which collected daily data on mortality, meteorology, and air pollutant concentrations for 100 U.S. cities from 1987-2000. The authors examined the sensitivity of city-specific ozone-mortality estimates to treatment of meteorology and co-pollutants, dependence on different ozone metrics, use of air conditioning, regional and spatial variability, and non-linear exposure-response relationships.

Non-Accidental Mortality

Assuming a 10 ppb change in daily 8-hour maximum ozone concentration, Smith et al. (2009) reported a 0.40 (0.22)% population-weighted change in non-accidental

mortality in a model without PM10, and 0.27 (0.22)% population-weighted change in non-accidental mortality in a model with PM10.

F.1.2 Zanobetti and Schwartz (2008b)

Zanobetti and Schwartz (2008b) examined the issue of "mortality displacement" (i.e., deaths are occurring in frail individuals and exposure is only moving the day of death to a day slightly earlier) in 48 U.S. cities during the warm season (i.e., June-August) for the years 1989-2000. The authors estimated a series of lagged specifications and found an effect size of 0.53% (95%C.I.: 0.28-0.77) for lag 0-3 days for all-cause mortality.

F.2   Hospital Admissions

Table F-2 summarizes the core health impact 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 F-2. Core Health Impact Functions for Ozone and Hospital Admissions

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
All Respiratory	Katsouyanni et al.	2009	14 U.S. cities	65-99		D8HourMax	0.000637	0.000400	Log-linear	Summer, penalized splines, 8df
All Respiratory	Katsouyanni et al.	2009	14 U.S. cities	65-99		D8HourMax	0.000614	0.000406	Log-linear	Summer, natural splines, 8df

F.2.1 Katsouyanni et al. (2009)

Katsouyanni et al. (2009) assessed the relationship between air pollution and hospital admissions from 1985 to 1994. Specifically, the authors examined hospitalizations due to respiratory diseases, conditions, or infections (ICD-9 codes 460-519) among U.S. citizens age 65 or older. The U.S. dataset included 14 cities with each city having data for 4 to 10 years from 1985-1994 and 7 cities having only summer ozone data. The investigators used a three-stage hierarchical model to account for within-city, within region, and between region variability. In the U.S. in adults 65 and older, authors found a .26% (.01, .51) increase in respiratory hospital admissions with 10 μg/m³ increase in ozone concentration using penalized splines with 8 df/year using lag 1. Authors found a .33% (.08, .58) increase in respiratory hospital admissions with 10 μg/m³ increase in ozone concentration using natural splines with 8 df/year using distributed lags.

F.3   Emergency Room Visits

Table F-3 summarizes the core 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 F-3. Core Health Impact Functions for Ozone and Emergency Room Visits

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form
Asthma	Glad et al.	2012	Pittsburgh, PA	0-99		D8HourMax	0.003057	0.0011709	Logistic
Asthma	Ito et al.	2007	New York, NY	0-99		D8HourMax	0.005213	0.000909	Log-linear
Asthma	Ito et al.	2007	New York, NY	0-99	PM2.5	D8HourMax	0.003976	0.000979	Log-linear
Asthma	Mar and Koenig	2009	Seattle, WA	18-99		D8HourMax	0.007696	0.0028374	Log-linear
Asthma	Mar and Koenig	2009	Seattle, WA	0-17		D8HourMax	0.010436	0.0043576	Log-linear
Asthma	Peel et al.	2005	Atlanta, GA	0-99		D8HourMax	0.000870	0.000529	Log-linear
Asthma	Sarnat et al.	2013	Atlanta, GA	0-99		D8HourMax	0.001113	0.0002828	Log-linear
Asthma	Wilson et al.	2005	Portland, ME	0.99		D8HourMax	0.003000	0.001000	Log-linear

F.3.1   Glad et al. (2012)

Glad et al. (2012) is summarized in section E.4.1.

Emergency Room Visits, Asthma

Using a case-crossover methodology, which controls for the effects of subject-specific covariates such as gender and race, a 2.5% increase was observed in asthma ED visits for each 10 ppb increase in the 1-hour maximum ozone level on day 2 (odds ratio [OR] = 1.025, p <0.05).

F.3.2   Ito et al. (2007)

The authors assessed the effect of multi-collinearity among gaseous co-pollutants of $PM_{2.5}$ and weather variables. The authors compiled daily meteorological and pollutant data in New York City for the years 1999 to 2002 and analyzed the effect of ozone on asthma emergency department visits for all year, warm months and cold months (all excluding September and October due to fall peaks in asthma emergency department visits). Data on ED visits were obtained from the 11 New York City Health and Hospitals Corporation medical centers with emergency receiving facilities.

F.3.3   Mar and Koenig (2009)

Mar and Koenig (2009) evaluated the relationship between outdoor ozone in the summer and asthma aggravation. The authors used hospital data on daily asthma cases from 1998 to 2002 in Seattle with local monitored $PM_{2.5}$ and ozone concentrations to assess the association between asthma visits to the emergency department and air pollutants. They analyzed 1-hour and 8-hour max ozone concentrations at 2 monitors in Greater Seattle. Asthma ED visits were analyzed at 0 through 5-day lags. The authors

found that ozone exposure exacerbates asthma in people in the Seattle area, especially in children. Authors found that in adults during the warmer months between May and September, a 10 ppb increase in 8-hour maximum ozone concentration is associated with relative risk of asthma-related ED visits of 1.08 (1.02, 1.14) with a 4-day lag. In children, during the warmer months, a 10 ppb increase in 8-hour maximum ozone concentration is associated with relative risk of asthma-related ED visits of 1.11 (1.02, 1.21) with a 3-day lag. The difference in lag and relative risk between children and adults suggests that children are more immediately responsive to the adverse effects of ozone exposure.

F.3.4   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: PM_{10}, ozone, NO_2, CO, and SO_2 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, NO_2, CO, and PM_{10} were associated with 1-3% increases in URI visits; a 2 microg/m increase of $PM_{2.5}$ organic carbon was associated with a 3% increase in pneumonia visits; and standard deviation increases of NO_2 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, NO_2, 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.

F.3.5   Sarnat et al. (2013)

The authors examined the influence of indoor air exchange rates as an effect modifier of associations between air pollutants and emergency department visits for asthma. The study analyzed associations between urban air pollutants (CO, NO_x, O_3, and $PM_{2.5}$) and asthma ED visits (ICD-9 codes 493 and 786.07) over a four-year period between 1999 and 2002 for residents of186 zip codes in metropolitan Atlanta. A spatial surface of daily ambient air concentrations of pollutants was generated using AERMOD air quality

modeling of 2002 NEI emissions combined with spatially interpolated background concentrations. AER was estimated at the zip code level based on an empirical model of direct and indirect predictors of AER collected from field surveys of building and meteorological characteristics. Effects of air pollution on asthma ED visits were estimated using a spatially resolved time series approach with Poisson generalized linear models, including stratified analyses for low or high AER areas. The authors found significant or near-significant positive effect modification of AER on the asthma ED visit effects of CO, NO_x, and $PM_{2.5}$. Ozone was a strong predictor of asthma ED visits but appeared unaffected by the AER. Overall, the study provides the first indication of short-term effect modification of air pollution risks with changes in AER. The HIF in BenMAP-CE is based on the association between ozone exposures and asthma ED visits for all AERs and all poverty levels (RR = 1.050, 95%CI 1.024-1.075) for a 26 ppb change in daily average ozone (3-day moving average of 0,1, and 2 day lags). The beta value was adjusted to use the daily 8-hour max metric as described below.

F.3.6   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 SO_2, 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 SO_2 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 O_3 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 SO_2 and O_3 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).

F.4   Minor Effects

Table F-4 summarizes the core 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 F-4. Core Health Impact Functions for Ozone and Minor Effects

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
School Loss Days, All Cause	Chen et al.	2000	Washoe Co, NV	5-17	PM_{10}, CO	D8HourMax	0.015763	0.004985	Linear	All year, 8-hour max from 1-hour max.
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.
Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	$PM_{2.5}$	D8HourMax	0.002596	0.000776	Log-linear	8-hour max from 1-hour max.

F.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 County 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).)

F.4.2   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 calculate 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:

$$Duration = \frac{TotalAbsences}{NewAbsences}$$

$$\Delta TotalAbsences = -[incidence \times (e^{-\beta \times O_3} - 1)] \times duration \times pop$$

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).)

F.4.3   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 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).

$$\beta = \left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right)$$

The standard error of the coefficient is calculated as follows, assuming that the estimated year-specific coefficients are independent:

$$\sigma_\beta^2 = \mathrm{var}\left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\sum\limits_{i=1976}^{1981} \dfrac{1}{\sigma_{\beta_i}^2}} \right) = \left(\frac{\sum\limits_{i=1976}^{1981} \dfrac{\beta_i}{\sigma_{\beta_i}^2}}{\gamma} \right) = \sum\limits_{i=1976}^{1981} \mathrm{var}\left(\frac{\beta_i}{\sigma_{\beta_i}^2 \times \gamma} \right)$$

This reduces down to:

$$\sigma_\beta^2 = \frac{1}{\gamma} \Rightarrow \sigma_\beta = \sqrt{\frac{1}{\gamma}} = 0.000658.$$

F.5 Asthma-Related Effects

Table F-5 summarizes the core health impacts functions used to estimate the relationship between ozone and asthma exacerbation. Below, we present a brief summary of each of the studies and any items that are unique to the study. Based on advice from the SAB-HES (U.S. EPA-SAB 2004), regardless of the age ranges included in the source epidemiology studies, we extend the applied population to ages 6 to 18, reflecting the common biological basis for the effect in children in the broader age group.

Table F-5. Core Health Impact Functions for Ozone and Asthma-Related Effects

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma Exacerbation, One or More Symptoms	Mortimer et al.	2002	Eight U.S. urban areas	6-18		D8HourMax	0.009288	0.00387	Logistic	Uses incidence rate of 0.116

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Asthma Exacerbation, One or More Symptoms	O'Connor et al.	2008	Seven U.S. urban communities	6-18	PM2.5, NO2	D8HourMax	0.0009661	0.002991	Log-linear	Uses incidence rate of 0.207
Asthma Exacerbation, One or More Symptoms	Schildcrout et al.	2006	Eight U.S. urban areas	6-18		D8HourMax	0.00222	0.002822	Logistic	Uses incidence rate of 0.52

F.5.1 Mortimer et al. (2002)

Mortimer et al. (2002) examined the effects of daily levels of PM10, SO2, NO2 and ozone on peak expiratory flow rate and asthma symptoms among 846 asthmatic children aged four to nine living in eight urban areas in the Northeast and Midwest. Exposures were estimated using EPA's AIRS database and averaging pollutant concentrations at all monitors in each subject's county. Effects were analyzed using mixed linear models and generalized estimating equation models. The authors found an association in single-pollutant models between each pollutant and self-reported morning asthma symptoms. The health impact function in BenMAP is based on the odds ratio for a 15 ppb increase in 4-day average 8-hour mean ozone between June and August (OR=1.16, 95%CI 1.02-1.30), adjusted to 8-hour max exposure metric as described below.

F.5.2 O'Connor et al. (2008)

O'Connor et al. (2008) studied the association between asthma symptoms and daily outdoor air pollutant concentrations (NO_2, SO_2, CO, ozone, and $PM_{2.5}$) among 937 asthmatic children aged five to 12 in seven inner-city areas in the U.S. Exposure data were estimated from monitor data from EPA's AIRS database for the years 1998 to 2001. Outcome data were collected via telephone interview every two weeks. The effect of pollutant levels on frequency of asthma symptoms was assessed using generalized estimating equation models. The health impact function in BenMAP is based on the relative risk for a 26.7 ppb increase in 19-day average 8-hour mean ozone between June and August (OR=1.04, 95%CI 0.82-1.32), adjusted to the 8-hour max exposure metric as described below.

Incidence Rate: Daily incidence rate for days with wheeze, tightness in chest, cough 21% (Table 1).

F.5.3 Schildcrout et al. (2006)

Schildcrout et al. (2006) investigated the relation between ambient concentrations of the five criteria pollutants (PM_{10}, O_3, NO_2, SO_2, and CO) and asthma exacerbations (daily symptoms and use of rescue inhalers) among 990 children in eight North American cities during the 22-month prerandomization phase (November 1993-September 1995) of the Childhood Asthma Management Program. Short-term effects of CO, NO_2, PM_{10}, SO_2, and warm-season O_3 were examined in both one-pollutant and two-pollutant models, using lags of up to 2 days in logistic and Poisson regressions.

When modeling ozone, the authors limited the study period to warm months (May through September).

Asthma Excerbation, One or More Symptoms

In a single-pollutant model, Schildcrout et al.(2006, Figure 1) reported an odds ratio of 1.06 (95% CI: 0.92, 1.23) for daily asthma symptoms associated with 30 ppb change in 24-hr mean of ozone at lag 0.

Incidence Rate: Daily incidence rate for one or more symptoms (symptom score>0) = 52% (Schildcrout, et al., 2006, Table 1).

F.6   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.

Appendix G. Additional Health Impact Functions in U.S. Setup

In this Appendix, we present additional health impact functions for estimating PM$_{2.5}$ and Ozone-related adverse health effects. Unlike Appendices E and F, these functions are included in the U.S. Setup but are not currently used by the U.S. EPA in regulatory impact analyses. For the health impact functions currently used by EPA, see the following page: https://www.epa.gov/benmap/benmap-community-edition. For Ozone Health Impact Functions, click the "U.S. EPA approach for quantifying and valuing ozone effects" link. For PM$_{2.5}$ Health Impact Functions, click the "U.S. EPA approach for quantifying and valuing PM effects" link.

G.1   Additional PM$_{2.5}$ Health Impact Functions

Tables G-1 and G-2 summarize the additional health impact functions for PM$_{2.5}$ included in BenMAP-CE.

Table G-1. Additional Health Impact Functions for PM$_{2.5}$ and Mortality

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Mortality, Lung Cancer	Krewski et al.	2009	116 U.S. cities	30-99		Annual	0.013103	0.003795	Log-linear	Mortality, Lung Cancer
Mortality, IHD	Krewski et al.	2009	116 U.S. cities	30-99		Annual	0.021511	0.002058	Log-linear	Mortality, IHD
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	0.005827	0.002157	Log-linear	
Mortality, All Cause	Woodruff et al.	1997	86 cities	Infant		Annual	0.003922	0.001221	Logistic	

Table G-2. Additional Health Impact Functions for PM$_{2.5}$ and Other Effects

Effect	Author	Year	Location	Age	Co-Poll	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	
All Cardiovascular (less Myocardial Infarctions)	Bell et al.	2012	187 U.S. Counties	65-99		D24HourMean	0.0008	0.0001071	Log-linear	
Congestive Heart Failure	Ito	2003	Detroit, MI	65-99		D24HourMean	0.003074	0.001292	Log-linear	
Dysrhythmia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.001249	0.002033	Log-linear	
Ischemic Heart (less Myocardial Infarctions)	Ito	2003	Detroit, MI	65-99		D24HourMean	0.001435	0.001156	Log-linear	

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Chronic Lung	Ito	2003	Detroit, MI	65-99		D24HourMean	0.001169	0.002064	Log-linear	
Pneumonia	Ito	2003	Detroit, MI	65-99		D24HourMean	0.003979	0.001659	Log-linear	
Hospital Admissions, All Cardiovascular (less Myocardial Infarctions)	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.001580	0.000344	Log-linear	
Hospital Admissions, Chronic Lung	Moolgavkar	2003	Los Angeles, CA	65-99		D24HourMean	0.001850	0.000524	Log-linear	
Hospital Admissions, Chronic Lung (less Asthma)	Moolgavkar	2000	Los Angeles, CA	18-64		D24HourMean	0.002200	0.000733	Log-linear	
Emergency Room Visits, Asthma	Norris et al.	1999	Seattle, WA	0-17	NO2, SO2	D24HourMean	0.016527	0.004139	Log-linear	

G.1.1 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 $\mu g/m^3$ 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 $\mu g/m^3$ 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 (2010a, Table 4) reports a chronic bronchitis prevalence rate for ages 18 and over of 4.37%.

G.1.2 Bell et al. (2012)

See explanation under Bell et al. (2008) in Appendix E, Section E.3.2.

G.1.3 Integrated Exposure Response Function (Cohen et al., 2017)

Burnett et al. (2014) developed relative risk (RR) functions that can be used to estimate the global burden of disease attributable to ambient fine particle ($PM_{2.5}$) exposure over the entire global exposure range (0-1,000 µg/m³). Functions were developed for causes of death including ischemic heart disease (IHD), cerebrovascular disease (CEV), chronic obstructive pulmonary disease (COPD), and lung cancer (LC), as well as mortality and lost-years of healthy life due to acute lower respiratory infection (ALRI) in children less than five years of age. These functions were developed because the bulk of the epidemiological studies linking $PM_{2.5}$ and mortality to date have been conducted in areas where annual average ambient fine particulate concentrations are lower than those observed in many developing countries. The original Integrated Exposure Response risk function (the 2013 IER) was estimated for each health endpoint by fitting a curve to RR data combined from studies on ambient air pollution, secondhand tobacco smoke, household solid cooking fuel, and active smoking, where the latter categories provide risk estimates associated with higher concentration $PM_{2.5}$ exposures. More recently, Cohen et al. (2017) updated the IER analysis done in Burnett et al. (2014) (the 2015 IER). The updated analysis included RR estimates from more recent epidemiological studies and refined statistical estimation techniques. In addition, the IER function for ALRI was expanded to estimate mortality impacts. The 2015 IER adult mortality RR curves have been adapted into a set of linear health impact functions that can be used in BenMAP-CE to estimate the health effects associated with changes in $PM_{2.5}$ concentrations occurring between 0-300 µg/m³. Each linear segment associated with a particular endpoint estimates the effects over a particular range of $PM_{2.5}$ concentrations defined by the A and B parameters of the corresponding HIF. Therefore, when using these functions in a health impacts analysis, all linear segments for a given cause-of-death should be run to ensure all effects are being captured. These component functions can then be pooled together to produce a single point estimate for an endpoint using the Sum Dependent method during Step 3 of a BenMAP-CE analysis (Aggregate, Pool & Value).

G.1.4 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 µg/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 µg/m^3 in the 1-day lag GAM stringent model (Ito, 2003, Table 7).

Dysrhythmia (ICD-9 code 427)

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 µg/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 µg/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 µg/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.

G.1.5 Krewski et al. (2009)

See full study explanation in Appendix E, Section E.1.2.

Mortality, Lung Cancer (ICD-10 code C30-C39)

In a random effects Cox model, the coefficient and standard error are estimated from the relative risks (1.14) and 95% confidence intervals (95% CI: 1.06-1.23) for a 10 $\mu g/m^3$ increase in the average of $PM_{2.5}$ exposure level for 1999-2000 (Krewski, et al., 2009, Commentary Table 4). The results were adjusted for the 44 individual-level covariates and the 7 ecologic covariates at the MSA & DIFF levels.

Mortality, Ischemic Heart Disease (ICD-10 code I20-I25)

In a random effects Cox model, the coefficient and standard error are estimated from the relative risks (1.24) and 95% confidence intervals (95% CI: 1.19-1.29) for a 10 $\mu g/m^3$ increase in the average of $PM_{2.5}$ exposure level for 1999-2000 (Krewski, et al., 2009, Commentary Table 4). The results were adjusted for the 44 individual-level covariates and the 7 ecologic covariates at the MSA & DIFF levels.

G.1.5 Laden et al. (2006)

A large body of epidemiologic literature has found an association of increased fine particulate air pollution ($PM_{2.5}$) 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 $PM_{2.5}$ 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 $PM_{2.5}$ concentrations were measured between 1979-1988, and estimated for later years from publicly available data. E xposure was defined as (1) city-specific mean $PM_{2.5}$ during the two follow-up periods, (2) mean $PM_{2.5}$ in the first period and change between these periods, (3) overall mean $PM_{2.5}$ across the entire follow-up, and (4) year-specific mean $PM_{2.5}$. 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 $\mu g/m^3$ increase in $PM_{2.5}$ 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). $PM_{2.5}$ 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 $PM_{2.5}$ (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 $PM_{2.5}$ concentrations. Reduced $PM_{2.5}$ 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 of (1.07-1.26) associated with a change in annual mean exposure of 10.0 $\mu g/m3$ (Laden et al., 2006, p. 667).

G.1.6 Moolgavkar (2000a)

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.

G.1.7 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 a 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.

G.1.8 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 μg/m^3 increase in PM$_{2.5}$ (Norris et al., 1999, p. 491).

G.1.9 Pope et al. (2002)

The Pope et al. (2002) analysis is a longitudinal cohort tracking study that uses the same American Cancer Society 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 (1979-1983 PM data, conducted for 61 metropolitan areas with 359,000 individuals), annual mean PM from the end of the tracking period (1999-2000, 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 1999-2000 than in 1979-1983 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, 1979-1983 Exposure

The coefficient and standard error for $PM_{2.5}$ using the 1979-1983 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 1979-1983 and 1999-2000 Exposure

The coefficient and standard error for $PM_{2.5}$ using the average of 1979-1983 and 1999-2000 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).

G.1.10 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 the 95% confidence interval (1.02-1.07) associated with a 10 μg/m^3 change in PM_{10} (Woodruff et al., 1997, Table 3).

G.2   Additional Ozone Health Impact Functions

Tables G-3 through G-5 summarize the additional health impact functions for Ozone included in BenMAP-CE.

Table G-3. Additional Health Impact Functions for Ozone and Mortality*

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
Non-Accidental	Bell et al.	2004	95 US cities	0-99		D8HourMax	0.000261	0.000089	Log-linear	Warm season. 8-hour max from24-hour mean
Non-Accidental	Bell et al.	2004	95 US cities	0-99		D24HourMean	0.000520	0.000128	Log-linear	All year
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
All Cause	Bell et al.	2005	US & non-US	0-99		D24HourMean	0.001500	0.000401	Log-linear	Warm season
Cardio-pulmonary	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.
Cardio-pulmonary	Huang et al.	2005	19 US cities	0-99		D24HourMean	0.001250	0.000398	Log-linear	Warm season
Non-Accidental	Ito and Thurston	1996	Chicago, IL	18-99	PM 10	D1HourMax	0.000634	0.000251	Log-linear	

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
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		D1HourMax	0.000400	0.000066	Log-linear	1-hour max
Non-Accidental	Ito et al.	2005		0-99		M24HourMean	0.001750	0.000357	Log-linear	Warm season. 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
Long Term Mortality, Respiratory	Jerrett et al.	2009	86 urban areas	30-99	$PM_{2.5}$	Annual	0.003922	0.001325	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Northeast	30-99		Annual	-0.001005	0.003853	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Industrial Midwest	30-99		Annual	0.000000	0.004604	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Southeast	30-99		Annual	0.011333	0.003193	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Upper Midwest	30-99		Annual	0.013103	0.026212	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Northwest	30-99		Annual	0.005827	0.003118	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Southwest	30-99		Annual	0.019062	0.007583	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	Southern California	30-99		Annual	0.000995	0.002767	Log-linear	
Long Term Mortality, Respiratory	Jerrett et al.	2009	86 urban areas	30-99	$PM_{2.5}$	Annual	0.004471	0.001510	Log-linear	
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.
All Cause	Levy et al.	2005	US and non-US	0-99		D1HourMax	0.000841	0.000134	Log-linear	Warm season.

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
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 U.S. cities	0-99		D8HourMax	0.000426	0.000150	Logistic	Warm season. 8-hour max from 1-hour max.
Non-Accidental	Schwartz	2005	14 U.S. cities	0-99		D1HourMax	0.000370	0.000130	Logistic	Warm season
Non-Accidental	Smith et al.	2009	98 U.S. cities	0-99		D8HourMax	0.000322	0.000084	Log-linear	Ozone season

*Unless otherwise stated, mortality is short-term.

Table G-4. Additional Health Impact Functions for Ozone and Hospital Admissions

Effect	Author	Year	Location	Age	Co-Poll	Metric	Beta	Std Err	Form	Notes
All Respiratory	Burnett et al.	2001	Toronto, CAN	0-1	PM$_{2.5}$	D1HourMax	0.007301	0.002122	Log-linear	Warm season
All Respiratory	Burnett et al.	2001	Toronto, CAN	0-1	PM$_{2.5}$	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

Table G-5. Additional Health Impact Functions for Ozone and Other Effects

Effect	Author	Year	Location	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	Gilliland et al.	2001	Southern California	5-17		D8HourMean	0.008150	0.004630	Log-linear	
Acute Respiratory Symptoms, Minor Restricted Activity Days	Ostro and Rothschild	1989	Nationwide	18-64	PM$_{2.5}$	D1HourMax	0.002200	0.000658	Log-linear	
Emergency Room Visits, Asthma	Wilson et al.	2005	Manchester, NH	0-99		D8HourMax	-0.001000	0.002000	Log-linear	

G.2.1   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.2.2   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.2.3   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.2.4   Chen et al. (2000)

See Appendix F, Section F.4.1, for an explanation of this study.

G.2.5   Gilliland et al. (2001)

See Appendix F, Section F.4.2, for an explanation of this study.

G.2.6   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 PM_{10}, but are robust to: (i) the adjustment for long-term trends, other gaseous pollutants (NO_2, SO_2 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.2.7 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 -- PM_{10}, ozone, CO, SO_2, and visual range-derived extinction coefficient -- both PM_{10} 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 PM_{10} 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.2.8 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 1-hour daily maximum, 24-hour daily average metrics, and 8-hour maximum from the 24-hour mean. We include the warm season 8-hour maximum from the 24-hour mean. 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%.

<u>One-hour Max Function</u>

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%.

<u>Daily Average Function</u>

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.2.9   Jerret et al. (2009)

Jerrett et al. (2009) examined the potential contribution of long-term ozone exposure to the risk of death from cardiopulmonary causes and specifically to death from respiratory causes. Data from the study cohort of the American Cancer Society Cancer Prevention Study II were correlated with air-pollution data from 96 metropolitan statistical areas in the United States. Associations between ozone concentrations and the risk of death were evaluated with the use of standard and multilevel Cox regression models. In single-pollutant models, increased concentrations of either $PM_{2.5}$ or ozone were significantly associated with an increased risk of death from cardiopulmonary causes. In two-pollutant models, $PM_{2.5}$ was associated with the risk of death from cardiovascular causes, whereas ozone was associated with the risk of death from respiratory causes. The estimated relative risk of death from respiratory causes that was associated with an increment in ozone concentration of 10 ppb was 1.040 (95% confidence interval, 1.010 to 1.067). The association of ozone with the risk of death from respiratory causes was insensitive to adjustment for confounders and to the type of statistical model used. The authors concluded that they were not able to detect an effect of ozone on the risk of death from cardiovascular causes when the concentration of $PM_{2.5}$ was taken into account. But they did demonstrate a significant increase in the risk of death from respiratory causes in association with an increase in ozone concentration.

Mortality, Respiratory (ICD-9 code 460-519) - 86 U.S. urban areas

In a two-pollutant model the coefficient and standard error are estimated from the relative risk (1.040) and 95% confidence interval (95% CI: 1.013-1.067) for a 10 ppb

increase in ambient ozone concentration measured from April to September during the years from 1977 to 2000 in 86 MSAs (Jerrett, et al., 2009, Table 3).

Mortality, Respiratory (ICD-9 code 460-519) - by region

In single-pollutant models the coefficient and standard error for different regions are estimated from the relative risks and 95% confidence intervals for a 10 ppb increase in ambient ozone concentration measured from April to September during the years from 1977 to 2000 (Jerrett, et al., 2009, Table 4).

Mortality, Respiratory (ICD-9 code 460-519) - adjusted daily metric

Based on the coefficients estimated from the two-pollutant model in the 86 urban areas using daily 1-hour max metric, the coefficients were adjusted for daily 8-hour max metric using a ratio of 1.14 (Anderson & Bell table 2).

G.2.10 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 $PM_{2.5}$.

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 μg/m^3 change in ozone. We converted μg/m^3 to ppb with an assumed relationship of 1.96 μg/m^3 per 1.0 ppb.

G.2.11 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.2.12 Moolgavkar et al. (1997)

Moolgavkar et al. (1997) examined the relationship between air pollution and hospital admissions (ICD-9 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.2.13 Ostro and Rothschild (1989)

See Section F.4.3 for an explanation of the study.

G.2.14 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.2.15 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.2.16 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.2.17 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 (SO_2) 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 SO_2, inhalable particles (PM_{10}), 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 PM_{10} associations were little changed by control for either ozone or SO_2. The ozone association was likewise independent of the other pollutants. The SO_2 association was substantially attenuated by control for ozone in both cities, and by control for PM_{10} 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/m^3 increase in PM_{10}, 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 PM_{10}, 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 µg/m^3 increase in average daily ozone levels (Schwartz, 1995, Table 6, p. 535).To calculate the coefficient, a conversion of 1.96 µg/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 µg/m^3 increase in average daily ozone levels (Schwartz, 1995, Table 3, p. 534).To calculate the coefficient, a conversion of 1.96 µg/m^3 per ppb was used, based on a density of ozone of 1.96 grams per liter (at 25 degrees Celsius).

G.2.18 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.2.19 Smith et al. (2009)

See Appendix F, Section F.1.1, for an explanation of the study

G.2.20 Wilson et al. (2005)

See Appendix F, Section F.3.6, for an explanation of the study.

Appendix H.   Core Health Valuation Functions in U.S. Setup

This appendix presents the core unit values that are available in BenMAP for each of the health endpoints included in the current suite of health impact functions. Specifically, this appendix includes the values currently used by U.S. EPA in regulatory impact analyses. For the .apvx files summarizing current EPA practices, see:

https://www.epa.gov/benmap/benmap-community-edition

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 are 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 value
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:

$$\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta}$$

This appendix also presents EPA methods for developing income growth adjustment factors that allow BenMAP-CE users to adjust the WTP estimates to account for the growth in income over time.

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 $8.7 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 mirrors that of Viscusi (1992), and uses the same criteria as Viscusi in his review of value-of-life studies. The $8.7 million estimate is consistent with Viscusi's conclusion (updated to 2015$) that "most of the reasonable estimates of the value of life are clustered in the $5.2 to $12.3 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. Table H-2 presents the unit values for the 26 value-of-life studies.

Table H-2. Core Unit Values for VSL based on 26-value-of-life studies

Basis for Estimate *	Age Range at Death		Unit Value (VSL) (2015$)	Distribution of Unit Value	Parameters of Distribution	
	Min	Max			P1	P2
VSL, based on 26 value-of-life studies	0	99	8,705,114	Weibull	9,648,168	1.509588

* The original value of a statistical life was calculated in 1990$. We have used a factor of 1.8134, based on the All-Items CPI-U.

H.2 Hospital Admissions & Emergency Room Visits

This section presents the core 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.2.1 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 National Inpatient Sample (NIS) database (2007). The NIS is the largest inpatient care database in the United States, and it is the only national hospital database containing charge information on all patients. It contains data from a very large nationally representative sample of about eight million hospital discharges, and therefore provides the best estimates of mean hospital charges and mean lengths of stay available, with negligible standard errors. The sampling frame for the 2007 NIS is a sample of hospitals that comprises approximately 90 percent of all hospital discharges in the United States. Since the NIS is based on discharge samples, the discharge-level weight was used to weight discharges in order to produce national estimates. The principle diagnoses were used to define the health endpoints.

Since most pollution-related hospital admissions are likely unscheduled, the unit values of avoided hospital admissions used in BenMAP are based solely on unscheduled hospitalizations. 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.

County-specific median annual income divided by (52*5) was used to estimate county-specific median daily wage. The data source for median annual income is the 2015 American Community Survey (ACS). ACS provided data for median annual income for all individuals over 16 years old in 819 counties. For all other counties, ACS provided a

five-year estimate of median annual income for the years 2010-2014. We calculated the ratio of state-specific median annual income in 2015 to state-specific median annual income during this five-year interval (2010-2014). This ratio was then applied to the 2010-2014 county-specific median annual income to obtain an estimate of 2015 county-specific income for the 2,323 counties without 2015 one-year estimates from ACS. Because wage data used in BenMAP are county-specific, the unit value for a hospital admission varies from one county to another.

Although the data for hospital charges are from year 2007, the default hospital admission unit values in BenMAP are in year 2015 dollars to be consistent with the unit values of other health endpoints in BenMAP. This was done by inflating the medical costs (2007 dollars) to 2015 dollars using BenMAP's inflation index.

The hospital admission outcomes that the EPA uses in its regulatory analyses are given in Table H-3. 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.

Table H-3. Core Unit Values Available for Hospital Admissions

| Endpoint | ICD Codes | Age Range | | Mean Hospital Charge (2015 $) | Mean Length of Stay (days) | Total Cost of Illness (Unit Value in 2015$)* |
		Min	Max			
HA, All Cardiovascular	390-429	18	64	$45,659	4.12	$46,371
HA, All Cardiovascular	390-429	65	99	$42,642	4.88	$43,485
HA, All Respiratory	460-519	65	99	$35,402	6.07	$36,451
HA, Asthma	493	0	64	$16,655	3.00	$17,174
HA, Chronic Lung Disease	490-496	18	64	$21,989	3.90	$22,663

* The opportunity cost of a day spent in the hospital was estimated, for the above exhibit, at the median daily wage of all workers, regardless of age. The median daily wage was calculated by dividing the median weekly wage ($864 in 2015$) by 5. The median weekly wages for 2015 were obtained from the U.S. Census Bureau's 2015 American Community Survey, "Selected Economic Characteristics: 2015 American Community Survey 1-Year Estimates."

H.2.2 Emergency Room Visits for Asthma

Two unit values are currently available for use in BenMAP for asthma emergency room (ER) visits. One is $533.69, 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 $533.69 in 2015$ (using the CPI for medical care to adjust to 2015$). 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 $533.69, on the interval [$395.14, $738.19].

A second unit value is $446.52 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-4.

Table H-4. Core Unit Values Available for Asthma-Related ER Visits

	Age Range		Unit Value (2015$)	Distribution of Unit Value	Parameters of Distribution	
Basis for Estimate	**Min**	**Max**			**P1**	**P2**
COI: Smith et al. (1997)	0	99	$534	Triangular	$395	$738
COI: Standford et al. (1999)	0	99	$447	Normal	8.95	--

H.3   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 acute respiratory symptoms (minor restricted activity days), asthma exacerbation, and school loss days. Tables H-5 and H-6 summarize the values used by EPA in their regulatory impact analyses.

Table H-5. Additional Unit Values Available for Myocardial Infarction

	Age Range			Opportunity	
Basis of Estimate	**Min**	**Max**	**Medical Cost ***	**Cost *****	**Total Cost**
COI: 5 yrs med, 5 yrs wages, 3% DR, Wittels (1990)	0	24	$187,530	$0	$187,530
	25	44	$187,530	$13,301	$200,831
	45	54	$187,530	$19,604	$207,134
	55	65	$187,530	$113,316	$300,846
	66	99	$187,530	$0	$187,530
COI: 5 yrs med, 5 yrs wages, 3% DR, Russell (1998)	0	24	$38,253	$0	$38,253
	25	44	$38,253	$13,301	$51,554
	45	54	$38,253	$19,604	$57,857
	55	65	$38,253	$113,316	$151,569
	66	99	$38,253	$0	$38,253
COI: 5 yrs med, 5 yrs wages,	0	24	$187,530	$0	$187,530

Basis of Estimate	Age Range		Medical Cost *	Opportunity Cost **	Total Cost
	Min	Max			
7% DR, Wittels (1990)	25	44	$187,530	$11,908	$199,438
	45	54	$187,530	$17,552	$205,082
	55	65	$187,530	$101,451	$288,981
	66	99	$187,530	$0	$187,530
COI: 5 yrs med, 5 yrs wages, 7% DR, Russell (1998)	0	24	$36,167	$0	$36,167
	25	44	$36,167	$11,908	$48,075
	45	54	$36,167	$17,552	$53,719
	55	65	$36,167	$101,451	$137,618
	66	99	$36,167	$0	$36,167

* 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 $187,530; Russell et al. estimated first- year direct medical costs and annual costs thereafter. The resulting 5-year cost is $38,253, using a 3% discount rate, and $36,167, using a 7% discount rate. Medical costs were inflated to 2015$ using CPI for medical care.
** From Cropper and Krupnick (1999). Present discounted value of 5 yrs of lost earnings, at 3% and 7% discount rate, adjusted from 1977$ to 2015$ using CPI-U "all items".

Table H-6. Core Unit Values Available for Acute Symptoms and Illnesses

Health Endpoint	Basis for Estimate	Age Range		Unit Value (2015$)	Distribution of Unit Value	Parameters of Distribution	
		Min	Max			P1	P2
Acute Bronchitis	WTP: 6 day illness, CV studies	0	17	$490	Uniform	144.60	834.98
Minor Restricted Activity Days	WTP: 1 day, CV studies	18	99	$70	Triangular	28.51	110.62
Upper Respiratory Symptoms	WTP: 1 day, CV studies	0	17	$34	Uniform	12.29	59.34
Lower Respiratory Symptoms	WTP: 1 day, CV studies	0	17	$21	Uniform	9.56	33.68
Work Loss Days *	Median daily wage, county-specific	18	65	$173	None	N/A	N/A
School Loss Days	Described in text	0	17	$106	None	N/A	N/A

* 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-7. Core Unit Values Available for Asthma-related Acute Symptoms and Illnesses

Health Endpoint	Basis for Estimate	Age Range		Unit Value (2015$)	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)	0	17	$59	Uniform	21.42	97.56

H.3.1 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 (1999), and a three percent discount rate, we estimated the following present discounted values in lost earnings over 5 years due to a heart attack (2015$): $13,301 for someone between the ages of 25 and 44, $19,604 for someone between the ages of 45 and 54, and $113,316 for someone between the ages of 55 and 65. The corresponding age-specific estimates of lost earnings using a seven percent discount rate are $11,908, $17,552, and $101,451, 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 $187,530 in year 2015$. 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 $85,052 in 2015$ for MI patients, using statistical prediction (regression) models to estimate inpatient costs. Only inpatient costs (physician fees and hospital costs) were included. Estimates from Eisenstein et al. are included in Appendix I.4 as they are not used in EPA impact analyses.

Russell et al. (1998) estimated first-year direct medical costs of treating nonfatal MI of $15,540 (in 1995$), and $1,051 annually thereafter. Converting to year 2015$, that would be $38,253 for a 5-year period, using a three percent discount rate, or $36,167, 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 above.

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.

H.3.2 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. The unit value that the EPA uses in their benefit analyses reflects that acute bronchitis typically lasts 6 or 7 days. To generate this unit value, the original unit value of $81.63 could be multiplied by 6 or 7. A unit value of $490 (=$81.63 x 6) was therefore derived. For more information about the original, one-day unit value, see Appendix I, Section I.4.2. The unit value that the EPA uses can be found in table H-6.

H.3.3 Minor Restricted Activity Days (MRADs)

Two unit values are currently available in BenMAP for MRADs associated with acute respiratory symptoms. No studies are reported to have estimated WTP to avoid a minor restricted activity day (MRAD). 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.

IEc (1993) derived an estimate of WTP to avoid a 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 $).

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 $61. Adjusting to 2015$, this is a triangular distribution centered at $69.58, ranging from $29 to $111.

The estimate for the MRADs that is used in EPA benefits analyses can be found in Table H-6.

H.3.4 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.

Currently, the EPA only uses the unit value for asthma exacerbation in children from Rowe and Chestnut (1986) for avoiding a "bad asthma day". There are two other unit values for children and two unit values for adults included in BenMAP but not currently used by the EPA. These are discussed further in Appendix I, Section I.4.3. In Rowe and Chestnut (1986), the mean of the four average WTPs is $32 (1990 $), or $59 in 2015$. 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 [$21, $98] in 2015$. Table H-7 summarizes the unit value utilized by EPA.

H.3.5 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-8 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-8. Median WTP Estimates and Derived Midrange Estimates (in 2015$)

Symptom*	Dickie et al.	Tolley et al. (1986)	Loehman et al. (1979)	Mid-Range Estimate
Throat congestion	6.84	29.65	-	18.14
Head/sinus congestion	7.98	31.94	14.87	18.14
Coughing	2.29	25.11	9.03	12.70
Eye irritation	-	28.50	-	28.50
Headache	2.29	45.63	-	18.14
Shortness of breath	0.00	-	19.16	9.06
Pain upon deep inhalation (PDI)	8.01	-	-	8.01
Wheeze	4.57	-	-	4.57
Coughing up phlegm	4.99	-	-	4.99
Chest tightness	11.42	-	-	11.42

* 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, $33.91, is just an average of the seven estimates of mean WTP for the different URS complexes. This unit value can be found in Table H-6.

H.3.6 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-8), 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 unit value for LRS that EPA uses for its benefits analyses, $21.43, is just an average of the eleven estimates of mean WTP for the different LRS symptom clusters.

H.3.7 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. The unit value for WLDs can be found in Table H-6.

H.3.8 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 Labor Statistics (2015) we obtained the rate of participation in the workforce of women with children under 18 years of age. We multiplied this rate (69.9%) by the estimated daily lost wage (if a mother must stay at home with a sick child), based on the median full-time weekly wage among women 25 and older in 2015. This median weekly wage is $759 (2015$). Dividing by 5 gives an estimated median daily wage of $152. 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 =

69.9% of $152, or $106. We currently have insufficient information to characterize the uncertainty surrounding this estimate.

A unit value based on the approach described above is likely to understate the value of a school loss day in three 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. Third, the approach may use a wage rate that is too low by assuming that men do not stay at home with sick children. The unit value of $106 is therefore considered an "interim" value until such time as alternative means of estimating this unit value become available. The unit value available can be found in table H-6 above.

H.4   Developing Income Growth Adjustment Factors

Chapter 4 of the BenMAP-CE User Manual provides instructions for formatting and adding income growth data. These values are used to adjust WTP estimates for growth in real income. As discussed in that chapter, evidence and theory suggest that WTP should increase as real income increases. When reviewing the economic literature to develop income growth adjustment factors, it is important to have an economist assist. For an overview of valuation, see Chapter 7: Aggregating, Pooling, and Valuing.

Adjusting WTP to reflect growth in real income requires three steps:

1. *Identify relevant income elasticity estimates from the peer-reviewed literature.*

2. *Calculate changes in future income.*

3. *Calculate adjustments to WTP based on changes in future income and income elasticity estimates.*

1. Identifying income elasticity estimates

Income elasticity estimates relate changes in demand for goods to changes in income. Positive income elasticity suggests that as income rises, demand for the good also rises. Negative income elasticity suggests that as income rises, demand for the good falls. BenMAP-CE does not adjust Cost-of-Illness (COI) estimates according to changes in income elasticity due to the fact that COI estimates the direct cost of a health outcome; instead we adjust this metric using inflation factors described above. BenMAP-CE includes income elasticity estimates specific to the type of health endpoint associated with the WTP estimate. BenMAP-CE contains elasticity estimates for three types of health effects: minor, severe and premature mortality. Minor health effects are those of short duration. Severe, or chronic, health effects are of longer duration. Consistent with economic theory, the peer reviewed literature indicates that income elasticity varies according to the severity of the health effect. Below we summarize the health endpoints considered minor and severe within the default *United States* setup in BenMAP-CE.

Minor
   ▪ Asthma exacerbation

- Acute bronchitis
- Acute respiratory symptoms (minor restricted activity days)
- Lower respiratory symptoms
- Upper respiratory symptoms

Severe
- Chronic bronchitis
- Chronic asthma

A review of the literature revealed a range of income elasticity estimates that varied across the studies and according to the severity of health endpoint. Table H-8 summarizes the income elasticity estimates found in BenMAP-CE to adjust minor health effects, severe health effects and premature mortality. Here we have provided a lower-, upper- and central-estimate for each type of health endpoint.

Table H-7. Income Elasticity Estimates

Health Endpoint	Lower Bound	Central Estimate	Upper Bound
Minor Health Effect	0.04	0.15	0.30
Severe and Chronic Health Effects	0.25	0.45	0.60
Premature Mortality	0.08	0.40	1.00

2. Calculating changes in future income

The next input to the WTP adjustment is annual changes in future income. The Congressional Budget Office's (2016) ten year projections of US Gross Domestic Product (GDP) are used to estimate changes in future income. Historical GDP data came from the U.S. Bureau of Commerce's Bureau of Economic Analysis. GDP values were adjusted for inflation as needed using the Implicit Price Deflator annual index, published by the Economic Research Division of the Federal Reserve Bank of St. Louis. We divided the projected change in GDP by the Woods & Poole (2015) projected change in total US population to produce an estimate of the future GDP per capita.

3. Calculating changes in WTP

The income elasticity estimates from Table H-7 and the estimated changes in future income may then be used to estimate changes in future WTP for each health endpoint. The adjustment formula follows four steps:

1)
$$\varepsilon = \frac{\frac{\Delta WTP}{WTP}}{\frac{\Delta I}{I}} = \frac{(WTP_2 - WTP_1) \times (I_2 + I_1)}{(I_2 - I_1) \times (WTP_2 + WTP_1)}$$

2)
$$\varepsilon I_2 WTP_2 + \varepsilon I_2 WTP_1 - \varepsilon I_1 WTP_2 - \varepsilon I_1 WTP_1$$
$$= I_2 WTP_2 + I_1 WTP_2 - I_2 WTP_1 - I_1 WTP_1$$

3)
$$WTP_2 \times (\varepsilon I_2 - \varepsilon I_1 - I_2 - I_1) = WTP_1 \times (\varepsilon I_1 - \varepsilon I_2 - I_1 - I_2)$$

4)
$$WTP_2 = WTP_1 \times \frac{\varepsilon I_1 - \varepsilon I_2 - I_1 - I_2}{\varepsilon I_2 - \varepsilon I_1 - I_2 - I_1}$$

Table H-8 summarizes the income-based WTP adjustments used within BenMAP-CE for minor health endpoints, severe health endpoints, and premature mortality.

Table H-8. Income-Based WTP Adjustments by Health Effect and Year

	Minor Health Endpoint			Severe Health Endpoint			Mortality		
Year	Low	Mid	Upper	Low	Mid	Upper	Low	Mid	Upper
1990	1	1	1	1	1	1	1	1	1
1991	0.99943614	0.997887194	0.995778859	0.99648118	0.993674994	0.991575539	0.998872638	0.994375765	0.985998511
1992	1.00027895	1.0010463	1.002093554	1.00174439	1.003141999	1.004191637	1.000557899	1.002792478	1.006995797
1993	1.00083518	1.003135324	1.006280541	1.005231023	1.009435654	1.01260066	1.001670957	1.008382797	1.021089673
1994	1.001928926	1.007252812	1.014558434	1.012117267	1.021917462	1.029330373	1.003861666	1.019458413	1.049366593
1995	1.00252676	1.009508491	1.019107819	1.01589787	1.028799176	1.038584113	1.005059958	1.025558352	1.065144777
1996	1.003553152	1.013389945	1.026960373	1.022416711	1.040715098	1.054657698	1.00711906	1.036109447	1.092786551
1997	1.004830718	1.018236518	1.036808729	1.030580044	1.055725813	1.074997663	1.00968492	1.049381137	1.128201008
1998	1.006105304	1.023088813	1.046717048	1.038779855	1.070903659	1.095668674	1.012248039	1.062778115	1.164700031
1999	1.007476926	1.028329492	1.057473063	1.047665954	1.087466359	1.118347049	1.01500988	1.077371955	1.205345869
2000	1.008633733	1.032765627	1.066622734	1.055212379	1.101626635	1.137836933	1.017342448	1.089827657	1.24079144
2001	1.008626103	1.032736301	1.066562176	1.05516243	1.101532698	1.137707233	1.01732707	1.089745045	1.240554214
2002	1.008962274	1.034028053	1.069233894	1.057363987	1.105680108	1.143433094	1.018005252	1.093389511	1.251058578
2003	1.009722471	1.036953092	1.075297475	1.062356591	1.115113139	1.156486392	1.019539833	1.10167253	1.275163889
2004	1.010831594	1.041232586	1.084201217	1.069678664	1.129017591	1.175803781	1.021781206	1.11386621	1.31124568
2005	1.011770725	1.044866562	1.091792464	1.075912952	1.140922546	1.192415953	1.023680925	1.12429142	1.342671752
2006	1.012430668	1.047425628	1.097154975	1.080312252	1.149360299	1.204231024	1.025016785	1.131672144	1.36525023
2007	1.01275146	1.048671246	1.099770427	1.08245647	1.153483868	1.210017681	1.025666475	1.135276675	1.376378059
2008	1.012262344	1.04677248	1.095784903	1.079188704	1.147202373	1.201206207	1.024675965	1.12978518	1.359451532
2009	1.01079917	1.041107416	1.083940148	1.069464207	1.12860918	1.175235033	1.021715641	1.113508224	1.310176253
2010	1.011459589	1.043661356	1.089271665	1.073843718	1.136964202	1.186884999	1.023051262	1.120826483	1.332167268
2011	1.011786222	1.044926405	1.091917515	1.076015592	1.141119123	1.192690849	1.023712158	1.124463439	1.343194842
2012	1.012354255	1.047128916	1.096532583	1.079801798	1.148379683	1.202856183	1.024862051	1.130814672	1.362613082
2013	1.012712717	1.048520923	1.099454761	1.082197666	1.15298593	1.209318399	1.025588155	1.134841442	1.375030875
2014	1.013344169	1.050976157	1.104618788	1.086428881	1.161142945	1.220786929	1.026867747	1.141967177	1.397211194
2015	1.014022827	1.053619623	1.110193372	1.090992332	1.169972539	1.233237505	1.02824378	1.149673343	1.421498775
2016	1.014274955	1.054602981	1.11227107	1.092692018	1.173269868	1.237896562	1.028755188	1.152549148	1.430643797
2017	1.014827728	1.056761265	1.116838455	1.096426606	1.180531025	1.248175263	1.029876947	1.158878565	1.45092845
2018	1.015322924	1.058697701	1.120945096	1.099781871	1.187074065	1.257459402	1.030882478	1.164577603	1.469381213
2019	1.015639186	1.059935808	1.123574972	1.101929545	1.191271544	1.263426423	1.031524897	1.168231487	1.48130703
2020	1.015908599	1.060991406	1.125819683	1.103761792	1.194858909	1.268532872	1.032072306	1.171352863	1.491554141
2021	1.016283393	1.062461257	1.128949761	1.106315613	1.199867845	1.275673389	1.032834172	1.175709248	1.505947828
2022	1.016681671	1.064024925	1.132284641	1.109035254	1.205213666	1.283308029	1.03364408	1.180356026	1.521420956
2023	1.017086029	1.065614223	1.135679722	1.11180234	1.210665345	1.291108131	1.034466624	1.185091853	1.53731966
2024	1.017486334	1.067189336	1.139050007	1.114547729	1.216086745	1.298879623	1.035281181	1.189798594	1.553251266
2025	1.017879009	1.068736196	1.142365217	1.117246628	1.221428633	1.306551456	1.036080599	1.194433689	1.56906867
2026	1.018263578	1.070252538	1.145620227	1.119895339	1.226682782	1.314110994	1.036863804	1.198989868	1.584743023

Appendix I. Additional Health Valuation Functions in U.S. Setup

In this Appendix, we present additional health valuation functions. Unlike those in Appendix H, these functions are included in the U.S. Setup but are not currently used by the U.S. EPA in regulatory impact analyses. For the health valuation functions currently used by EPA, see the following page: https://www.epa.gov/benmap/benmap-community-edition. For Ozone Health Valuation Functions, click the "U.S. EPA approach for quantifying and valuing ozone effects" link. For PM$_{2.5}$ Health Valuation Functions, click the "U.S. EPA approach for quantifying and valuing PM effects" link.

I.1 Mortality

I.1.1 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 in Appendix H, section H.1.1, we have included three alternatives based loosely on the results of work by Mrozek and Taylor (2002) and Viscusi and Aldy (2003). Each of these three alternatives has a mean value of $7.6 million (2015$), but with a different distribution: normal, uniform, triangular, and beta. Table H-10 presents the distribution parameters for these additional valuations in BenMAP.

Table I-1. Additional Unit Values for VSL for additional value-of-life studies

Basis for Estimate *	Age Range at Death		Unit Value (VSL) (2015$)	Distribution of Unit Value	Parameters of Distribution	
	Min	Max			P1	P2
VSL, based on 2015$ range from $1.38 million to $13.76 million – 95% CI of assumed normal distribution	0	99	7,570,229	Normal	3,160,172	–
VSL based on 2015$ range from $1.38 million to $13.76 million – assumed uniform distribution	0	99	7,570,229	Uniform	1,376,405	13,764,053
VSL based on 2015$ range from $1.38 million to $13.76 million – assumed triangular distribution	0	99	7,570,229	Triangular	1,376,405	13,764,053

* The original value of a statistical life was calculated in 1990 $. We have used a factor of 1.8134, based on the All-Items CPI-U.

I.2 Hospital Admissions

This sub-section presents the unit values for hospital admissions that are not used by the EPA in regulatory impact analyses but are included in BenMAP. See Appendix H,

Section H.2 for more information about these unit values. Table I-2 includes the unit values for hospital admissions for endpoints included in BenMAP but not used by the EPA in their regulatory impact analyses.

Table I-2. Additional Unit Values Available for Hospital Admissions

Endpoint	ICD Codes	Age Range Min	Age Range Max	Mean Hospital Charge (2015 $)	Mean Length of Stay (days)	Total Cost of Illness (Unit Value in 2015$)*
HA, All Cardiovascular	390-429	0	99	$33,063	4.59	$33,856
HA, Congestive Heart Failure	428	65	99	$33,734	5.32	$34,654
HA, Dysrhythmia	427	0	99	$33,063	3.72	$33,706
HA, Ischemic Heart Disease	410-414	65	99	$55,591	4.61	$56,388
HA, All Respiratory	460-519	0	1	$16,929	3.19	$17,480
HA, All Respiratory	460-519	0	99	$32,563	5.35	$33,488
HA, Asthma	493	65	99	$26,153	4.79	$26,981
HA, Asthma	493	0	99	$18,590	3.37	$19,172
HA, Chronic Lung Disease	490-496	65	99	$25,413	4.79	$26,241
HA, Chronic Lung Disease	490-496	0	99	$22,312	4.10	$23,021
HA, Chronic Lung Disease (less Asthma)	490-492, 494-496	18	64	$23,980	4.23	$24,711
HA, Chronic Lung Disease (less Asthma)	490-492, 494-496	65	99	$25,254	4.79	$26,082
HA, Chronic Lung Disease (Less Asthma)	490-492, 494-496	0	99	$24,834	4.59	$25,627
HA, Pneumonia	480-487	65	99	$30,229	5.77	$31,226
HA, Pneumonia	480-487	0	99	$29,046	5.25	$29,953

* The opportunity cost of a day spent in the hospital was estimated, for the above exhibit, at the median daily wage of all workers, regardless of age. The median daily wage was calculated by dividing the median weekly wage ($864 in 2015$) by 5. The median weekly wages for 2015 were obtained from the U.S. Census Bureau's 2015 American Community Survey, "Selected Economic Characteristics: 2015 American Community Survey 1-Year Estimates."

I.3   Chronic Illness

This sub-section presents the unit values developed for chronic bronchitis, chronic asthma, and non-fatal myocardial infarctions.

I.3.1   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.

I.3.1.1 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 Viscusi study, and WTP to avoid a less severe case. The estimated relationship (see Table 4 in Krupnick and Cropper) can be written as:

$$\ln(WTP) = \alpha + \beta \times sev$$

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

$$\ln(WTP_{13}) = \alpha + \beta \times 13$$

and

$$\ln(WTP_x) = \alpha + \beta \times x$$

Subtracting one equation from the other,

$$\ln(WTP_{13}) - \ln(WTP_x) = \beta \times (13 - x)$$

or

$$1n\left(\frac{WTP_{13}}{WTP_x}\right) = \beta \times (13 - x)$$

Exponentiating and rearranging terms,

$$WTP_x = WTP_{13} \times e^{-\beta \bullet (13 - x)}$$

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 $470,000 (2015$), but not adjusted for the growth of income). Therefore, if you request that the analysis be carried out in "point estimate" mode, that is the unit value that is used.

I.3.1.2   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 (1999). 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 2015$ using the CPI-U for medical care; lost earnings (opportunity costs) were inflated to 2015$ 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 is shown in Table I-3.

Table I-3. Additional 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	$468,641	Custom
COI: med costs + wage loss, 3% DR	27	44	$32,478	$199,479	$231,947	None
	45	64	$40,699	$111,959	$152,658	None
	65	99	$18,993	$0	$18,993	None
COI: med costs + wage loss, 7% DR	27	44	$13,509	$118,460	$131,969	None
	45	64	$24,651	$87,732	$112,383	None
	65	99	$15,468	$0	$15,468	None

I.3.2   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 $206,765.

I.3.3   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 $65,568 (in 2015$).

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 $41,646 (in 2015$). A unit value, based on a three percent discount rate, is the average of the two estimates, or $53,607. 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 I-4 below.

Table I-4. Additional 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	$53,607	Triangular	$41,646	$65,568
WTP: 7% DR	27	99	$34,901	Triangular	$27,114	$42,689

I.4   Acute Symptoms and Illness Not Requiring Hospitalization

See Appendix H, Section H.3 for a general explanation of acute symptoms and illness not requiring hospitalization. Table I-5 summarizes unit values for acute bronchitis in children, acute respiratory symptoms (minor restricted activity days), any of 19 respiratory symptoms, upper respiratory symptoms, lower respiratory symptoms, and work loss days (WLDs)

Table I-5. Additional Unit Values Available for Acute Symptoms and Illnesses (in 2015 $)

Health Endpoint	Basis for Estimate *	Age Range		Unit Value	Distribution of Unit Value	Parameters of Distribution	
		Min	Max			P1	P2
Acute Bronchitis	WTP: 1 day illness, CV studies	0	17	$82	Uniform	24.10	139.16
	WTP: 28 symptom-days, Dickie and Ulery	0	17	$529	Lognormal	6.27	0.0907
Minor Restricted Activity Days	WTP: 3 symptoms 1 day, Dickie and Ulery (2002)	18	99	$138	Lognormal	4.93	0.0649
Lower Respiratory Symptoms	WTP: 2 symptoms 1 day, Dickie and Ulery (2002)	0	17	$264	Lognormal	5.58	0.07048
	WTP: 2x1 day, CV studies	0	17	$42	Uniform	19.12	67.36
Upper	WTP: 2 symptoms 1 day,	0	17	$264	Lognormal	5.58	0.07048

Health Endpoint	Basis for Estimate *	Age Range		Unit Value	Distribution of Unit Value	Parameters of Distribution	
Respiratory Symptoms	Dickie and Ulery (2002)						
	WTP: 2x1 day, CV studies	0	17	$68	Uniform	24.58	118.68
Any of 19 Respiratory Symptoms	WTP: 1 day illness, CV studies	1	65	$33	Uniform	0	66.41

* 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.

I.4.1 Non-Fatal Myocardial Infarctions (Heart Attacks)

See Appendix H, Section H.3.1 for an explanation of this valuation function. Table I-6 shows the unit values not used by the EPA.

Table I-6. Additional Unit Values Available for Myocardial Infarction

Basis of Estimate	Age Range		Medical Cost *	Opportunity Cost **	Total Cost
	Min	Max			
COI: 10 yrs med, 5 yrs wages, 3% DR, Eisenstein (2001)	0	24	$85,052	$0	$85,052
	25	44	$85,052	$13,301	$98,353
	45	54	$85,052	$19,604	$104,656
	55	65	$85,052	$113,316	$198,368
	66	99	$85,052	$0	$85,052
COI: 10 yrs med, 5 yrs wages, 7% DR, Eisenstein (2001)	0	24	$85,052	$0	$85,052
	25	44	$85,052	$11,908	$96,960
	45	54	$85,052	$17,552	$102,604
	55	65	$85,052	$101,451	$186,503
	66	99	$85,052	$0	$85,052

* 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 $187,530; Russell et al. estimated first- year direct medical costs and annual costs thereafter. The resulting 5-year cost is $38,253, using a 3% discount rate, and $36,167, using a 7% discount rate. Medical costs were inflated to 2015$ using CPI for medical care.
** From Cropper and Krupnick (1999). Present discounted value of 5 yrs of lost earnings, at 3% and 7% discount rate, adjusted from 1977$ to 2015$ using CPI-U "all items".

I.4.2 Acute Bronchitis in Children

In previous benefit analyses, EPA used a unit value of $81.63 (adjusted to 2015$). 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. See Appendix H, Section H.3.2 for the unit value the EPA uses for their benefit analyses.

As discussed in Appendix H, Section H.3.2, 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 $529. 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. Estimates for Acute Bronchitis are available in Table I-5.

I.4.3   Minor Restricted Activity Days (MRADs)

See Appendix H, Section H.3.3 for more information.

In addition to the estimate of WTP to avoid a MRRAD used in EPA benefits analyses, 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 $138.

The estimate for the MRADs that is not used in EPA benefits analyses can be found in Table I-5.

I.4.4   Asthma Exacerbation

See Appendix H, Section H.3.4, for more information about Asthma Exacerbation. Table I-7 below describes the unit values for Asthma-related Acute Symptoms and Illnesses included in BenMAP but not used by the EPA for their regulatory impact analyses. All unit values for this section are summarized in Table I-7.

The first unit value for adult is based on willingness to pay to avoid an asthma exacerbation from four WTP estimates from Rowe and Chestnut (1986) for avoiding a "bad asthma day." See Section H.3.4 for further description of Rowe and Chestnut

(1986). The 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 $104, 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 $221.

Table I-7. Additional Unit Values Available for Asthma-related Acute Symptoms and Illnesses

Health Endpoint	Basis for 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	$59	Uniform	21.42	97.56
	1 symptom-day, Dickie and Ulery (2002)	18	99	$104	Lognormal	4.64	0.0957
	2 x bad asthma day, Rowe and Chestnut (1986)	0	17	$118	Uniform	42.84	195.12
	1 symptom-day, Dickie and Ulery (2002)	0	17	$221	Lognormal	5.39	0.0925
Cough; Moderate or Worse; One or more symptoms; Shortness of Breath; Wheeze	Bad asthma day, Rowe and Chestnut (1986)	0	17	$59	Uniform	21.42	97.56

* 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.

I.4.5   Upper Respiratory Symptoms (URS) in Children

See Appendix H, Section H.3.5 for an explanation of URS valuation in children.

In addition to the one unit value that EPA uses for their benefit analyses, two other unit values can be found in BenMAP. 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. This second unit value of $67.82 (=2 x $33.91) is derived from the unit value found in Appendix H, Section H.3.4.

Another 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 $264.

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. Unit values can be found in Table I-5.

I.4.6   Lower Respiratory Symptoms (LRS) in Children

See Appendix H, Section H.3.6 for more discussion of LRS.

In addition to the original value used by the EPA in their benefits analyses, BenMAP contains a second unit value is twice the original unit value, or $42.86. This value is 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 $264. These two additional unit values can be found in Table I-5.

I.4.7   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:

$$(0.40)(\$33.91) + (0.40)(\$21.43) + (0.20)(\$33.91 + \$21.43) = \$33.20.$$

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. The value for any of 19 respiratory symptoms can be found in Table I-5.

Appendix J. Population & Other Data in U.S. Setup

This section describes the population, monitor, and demographic data in the United States setup. It consists of the following three subsections:

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, $PM_{2.5}$, PM_{10}, lead, NO_2, and SO_2 monitor data for the years 2000-2007. Data for CO are available at the BenMAP website: http://www.epa.gov/air/benmap/.

Demographic Datasets. This subsection describes the various datasets in the U.S. setup related to demography: household size, poverty rates, and educational attainment.

J.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 2010 Census block data. A separate application 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. The PopGrid program is available on the BenMAP-CE website here: www.epa.gov/benmap

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). Table J-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 you require.

Table J-1. Demographic Groups and Variables Available in BenMAP

Racial Group	Ethnicity	Age	Gender
White, African American, Asian, American Indian	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 need.

J.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 Table J-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:

$$age_{3-12} = \frac{1}{2} \times age_{1-4} + age_{5-9} + \frac{3}{5} \times age_{10-14}$$

To estimate population levels for the years after the last Census in 2010, BenMAP scales the 2010 Census-based estimate with the ratio of the county-level forecast for the future year of interest over the 2010 county-level population level. Woods & Poole (2015) 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 2020, BenMAP calculates:

$$age_{4-9,g,2020} = age_{4-9,g,2010} \times \frac{age_{4-9,county,2020}}{age_{4-9,county,2010}}$$

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 2010 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 2010:

$$age_{4-9,g \ in \ county_c,2010} = age_{4-9,g,2010} \times fraction \ age_{4-9,g \ in \ county_c}$$

To then forecast the population in 2020, we scale the 2010 estimate with the ratio of the county projection for 2020 to the county projection for 2010:

$$age_{4-9, g \text{ in } county_c, 2020} = age_{4-9, g \text{ in } county_c, 2010} \times \frac{age_{4-9, county_c, 2020}}{age_{4-9, county_c, 2010}}$$

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 2020 as follows:

$$age_{4-9, g, 2020} = \sum_{c=1}^{n} age_{4-9, g, 2010} \times \frac{total \ pop_{g \ in \ county_c}}{total \ pop_g} \times \frac{age_{4-9, county_c, 2020}}{age_{4-9, county_c, 2010}}$$

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:

$$age_{4-9, g, 2020} = \sum_{c=1}^{n} age_{4-9, g, 2010} \times \frac{total \ pop_{g \ in \ county_c}}{total \ pop_g} \times \frac{age_{4-9, county_c, 2020}}{age_{4-9, county_c, 2010}}$$

$$age_{10-14, g, 2020} = \sum_{c=1}^{n} age_{10-14, g, 2010} \times \frac{total \ pop_{g \ in \ county_c}}{total \ pop_g} \times \frac{age10-14_{10-14, county_c, 2020}}{age_{10-14, county_c, 2010}}$$

$$age_{4-12, g, 2020} = age_{4-9, g, 2020} + \frac{3}{5} \times age_{10-14, g, 2020}$$

J.1.2 Data Needed for Forecasting

Underlying the population forecasts in BenMAP there are block-level databases used to provide year 2010 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 2010 estimates for a grid definition of interest (e.g., 12 kilometer CMAQ grid). The output from PopGrid with the year 2010 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 year (2009, 2011-2050) and year 2010 population data for each county and each of the 304 race-ethnicity-gender-age population groups.

We describe the development of each databases below.

J.1.2.1 Block-Level Census 2010

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 J-2 summarizes the initial set of variables and the final desired set of variables.

Table J-2. Race, Ethnicity and Age Variables in 2010 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 2010 block-level and tract-level files (Summary File 1)
Data: http://www2.census.gov/census_2010/04-Summary_File_1/
Docs: http://www.census.gov/prod/cen2010/doc/sf1.pdf

Census 2000 MARS national-level summary
Docs: http://www.census.gov/popest/archives/files/MRSF-01-US1.pdf

The SF1 tract-level and MARS data, as described below, are needed to reorganize the variables that come initially in the block-levelSF1 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 U. S. Census.)

The steps in preparing the data are as follows:

1. Adjust Age-classifications:

We combined some age groups in the block-level 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 tract-level SF1 data, which gave us the fraction of 0-4 year-olds who are <1.

2. Fill in Missing Racial-Ethnic Interactions:

We used the tract-level SF1 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.

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, using state fractions of these races in each age-sex-race-block- level datum.

Exhibit J-3: Summary of Modified Race and Census 2000 Race Distributions for the United States

Subject	Modified Race		Census 2000	
	Number	Percent	Number	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

J.1.2.2   County-Level Forecasts

Woods & Poole (2015) developed county-level forecasts for each year from 2000 through 2050, by age and gender for non-Hispanic White, African-American, Asian-American, and Native-American and for all Hispanics. The detailed documentation can be found at http://www. woodsandpoole.com/pdfs/CED15.pdf. As discussed below, the

adjustments necessary to prepare the data for use in BenMAP are relatively straightforward.

For each non-Hispanic subset of the population and each year from 2000-2050, we divided the Woods and Poole population for that year by the Woods and Poole population for that subset in 2010. 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.[2]

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 2010 Census and those in 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 2010

There are a small number of cases were the 2010 county population for a specific demographic group is zero, so the ratio of any future year to the year 2010 data is undefined. In these relatively rare cases, we prepared statewide and national totals and used ratios at the higher levels of geographic aggregation when the more local ratios caused divide-by-zero errors.

J.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.

---

[2] Previous versions of the BenMAP-CE program used a different methodology whereby population estimates for 2000 – 2009 were adjusted using the ratio of 2000 Woods & Poole estimated population and 2000 Census population.

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 2010 population estimate from the Census Bureau multiplied by the ratio of future-year to year 2010 county population estimates from Woods & Poole. BenMAP assumes that all age-gender groups within a given race-ethnic group have the same geographic distribution.

J.1.3.1 How to Use PopGrid

After installing PopGrid, double-click on the PopGrid executable "PopGrid4.exe." The following screen will appear:
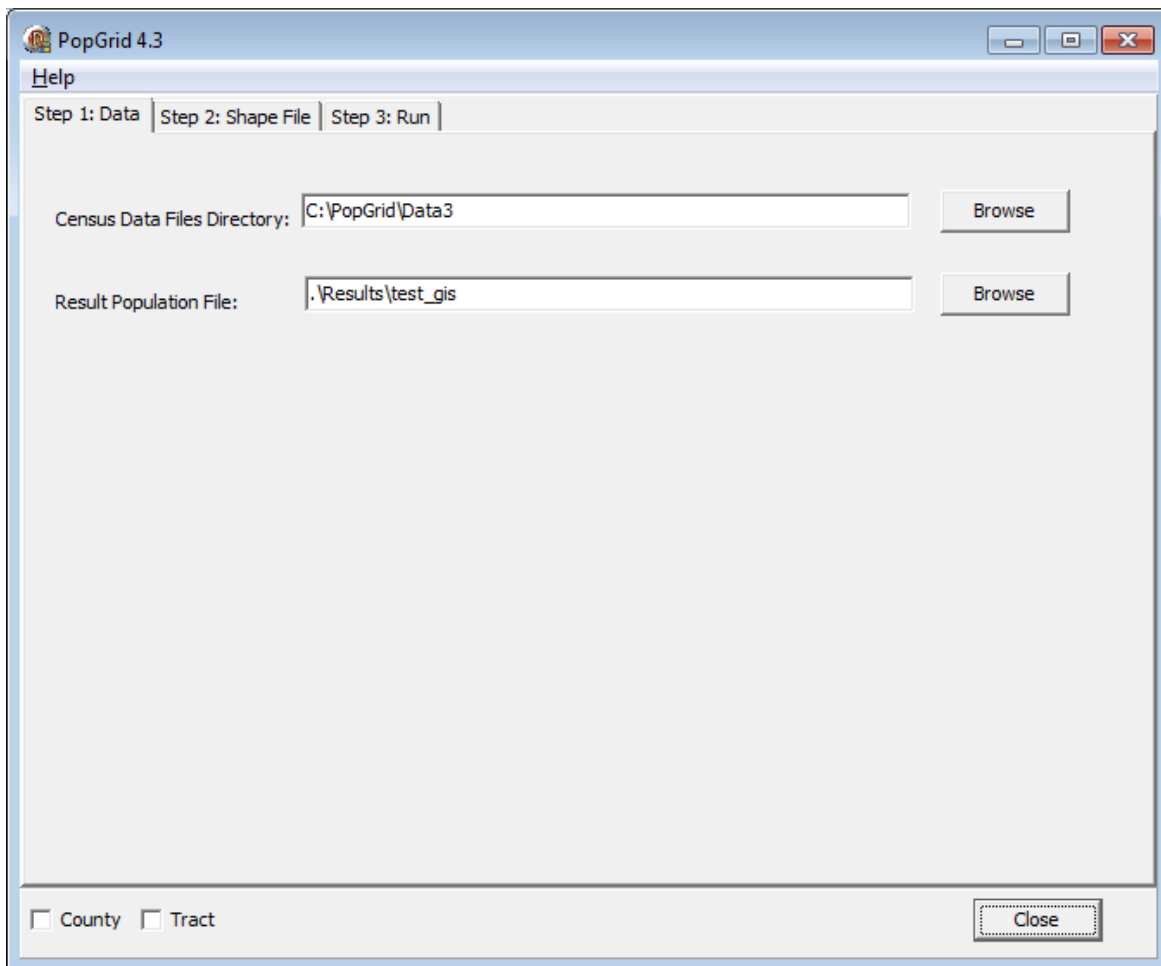
The **Census Data Files Directory** box points PopGrid to where the block data are located that PopGrid uses. Make sure that the files in this directory are unzipped. This data folder should look something like the following:
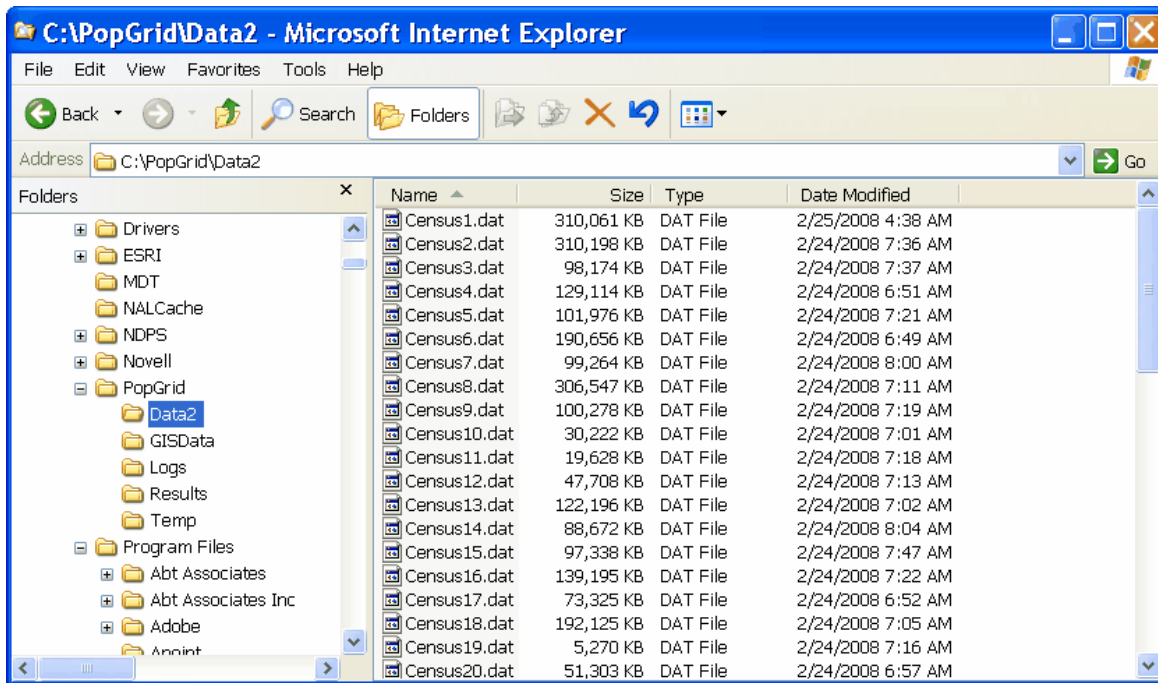
The **Result Population File** box provides the path and the name of the file that you want to create. In the example above, PopGrid is being used to estimate population for the intersection of air basins and counties in California (CA_AirBasin_by_County).

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:
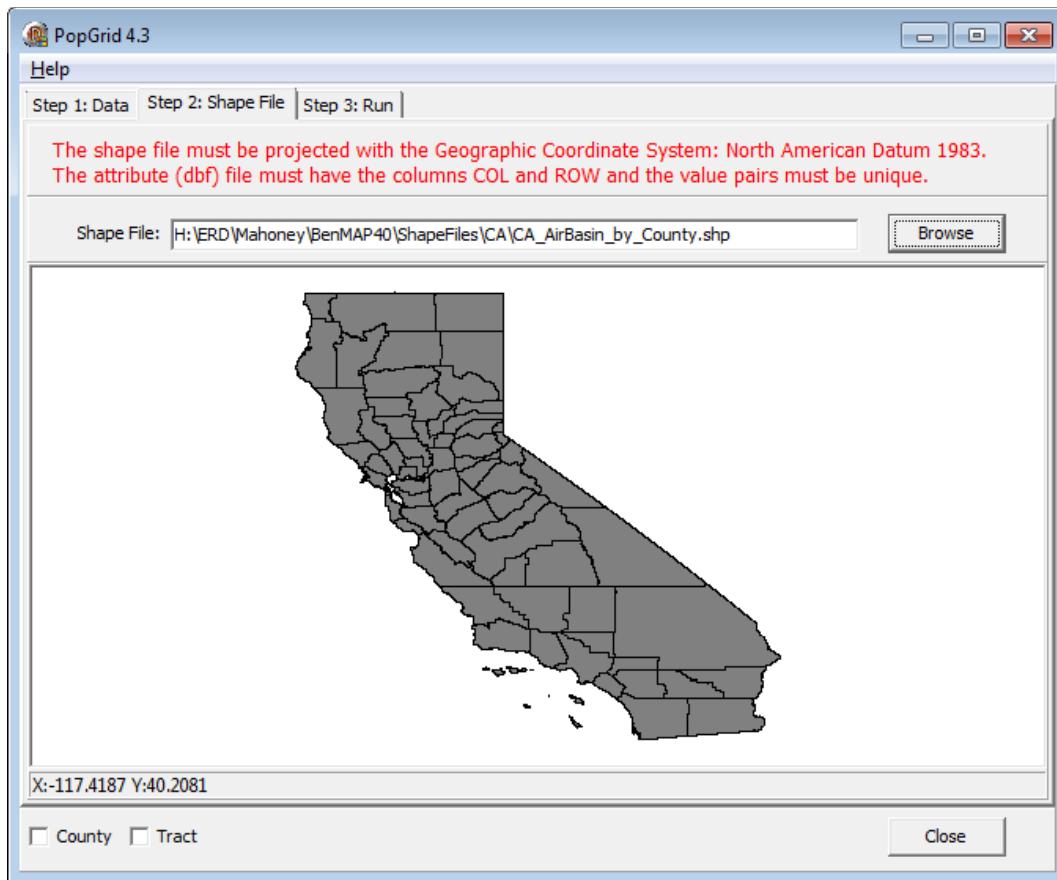
After choosing your shapefile, which must contain a column/row index, go to the **Step 3: Run** tab, which should look as follows:
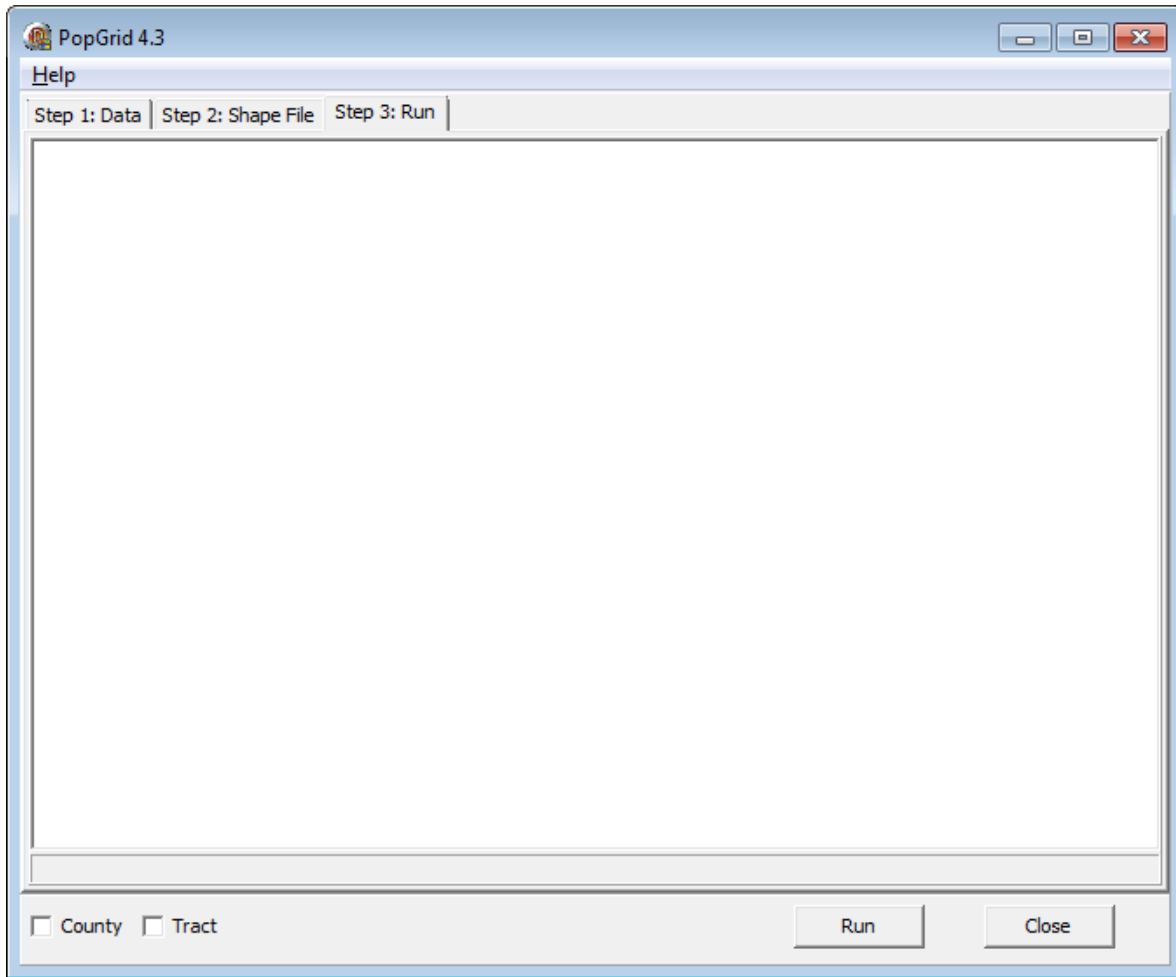
Click **Run**. PopGrid will now begin processing. It can take a very long time to run. When PopGrid has finished running, check the log file. The log file notes the start time, the files that PopGrid used, and the end time. Also, at the very end of the log file, PopGrid notes the number of people that 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").

H.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 J-5 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 J-5. 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

Row	Column	Year	Population	Race	Ethnicity	Gender	AgeRange
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	30TO$_3$4
58	81	2000	0.08	BLACK	HISPANIC	MALE	30TO$_3$4

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 J-5 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 J-6, 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 J-6. 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

SourceCol	SourceRow	TargetCol	TargetRow	Race	Ethnicity	Value	Year
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

J.2   Monitor Data in U.S. Setup

BenMAP-ready data files were created from 2000 through 2013, as reported to the U.S. Environmental Protection Agency's (EPA) Air Quality System (AQS) for $PM_{2.5}$ and ozone. Table J-7 summarizes the data sources and vintage of the processed data.

Table J-7. Underlying Data Sources for BenMAP Air Quality Data File

Pollutant	AQS Parameter Code	Year	Data Source
$PM_{2.5}$	88101	2000-2013	https://www.epa.gov/outdoor-air-quality-data
Ozone	44201	2000-2013	

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 J-8. 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 J-8. Pollutant-Specific QC Checks Performed in AQA

Pollutant	AQS Parameter Code	Acceptable Concentration Range	Maximum Rate of Change
$PM_{2.5}$	88101	> = 0 µg/m3	–
Ozone	44201	> = 0 ppb	60 ppb

J.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: $µg/m^3$ for aerosols, and ppb for ozone.

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, NO_2, SO_2, and CO.

J.2.2   Output Files

Table H-8 lists the number of monitors by pollutant and year, represented in the resulting BenMAP-ready data files. The AQS Parameter Code for $PM_{2.5}$ is 88101 and for Ozone is 44201.

Table J-9. Number of Monitors by Pollutant and Year Included in BenMAP

Year	Number of Monitors by Year	
	$PM_{2.5}$	Ozone
2000	1,295	1,138
2001	1,336	1,184
2002	1,331	1,191
2003	1,311	1,210
2004	1,222	1,207
2005	1,264	1,195
2006	1,202	1,201
2007	1,156	1,228
2008	1,188	1,236
2009	1,265	1,244
2010	1,289	1,272
2011	1,283	1,325
2012	1,278	1,314
2013	1,306	1,310

J.3　Demographic Datasets in U.S. Setup

The U.S. setup includes county-level data on household size, poverty status, educational attainment, unemployment, health insurance coverage, and occupational status. These datasets are included in the *EPA Standard Variables* dataset under *Variable Datasets*. We describe the data sources and processing methodology for each dataset below.

J.3.1　Household Size

To generate average household size for each county, we utilize ACS 5-year estimates for 2012 to 2016. Average household size was provided by ACS at the county level for all counties except for two, for which data were not available.[3] For these counties, we applied the state level average household size.

J.3.2　Educational Attainment

We use data from the ACS to provide county-level summaries of educational attainment stratified by race and ethnicity. These data represent 5-year average ACS estimates from 2012 to 2016. Specifically, the data included in BenMAP span two broad education categories: no high school diploma, and high school diploma (or equivalency) and above. The latter category includes individuals with a high school diploma (or equivalency), some college, college degree, or post-graduate degree.[4]

For both education groups (with/without HS diploma), we estimate the fraction of the race- and origin-specific county population (ages 25 years and above) in each education group. Thus, the two estimates sum to one for each county (within a given race category). To generate these values, we summed population estimates across genders and education groups for four race categories (White; Black or African American; American Indian and Alaska Native; and Asian and Pacific Islanders) and two ethnicity categories (Hispanic and Non-Hispanic). To generate the Asian and Pacific Islander group, we had to sum population estimates from (1) the Asian dataset and (2) the Pacific Islander dataset. For Non-Hispanic estimates, we subtracted Hispanic population from the total population to generate point estimates. Both the Hispanic and Non-Hispanic categories span the other race categories.

All estimates were generated at the county level for 3,109 counties in the contiguous United States.[5] For each estimate, we generate a coefficient of variation (CV) equal to the ratio of the standard error to the point estimate. For counties with a CV greater than 0.3, we impute the county-level estimate with a state-level estimate following Census

---

[3] The two counties without data were Shannon County, South Dakota (FIPS Code 46113) and Bedford, Virginia (FIPS Code 51515).

[4] Before BenMAP-CE version 1.4, the program included five education categories: no high school diploma, high school diploma (or equivalency), some college, college degree, and post graduate degree.

[5] In 2010, Bedford city, Virginia was deleted from the list of counties in the U.S. Due to BenMAP's grid definition, we continue to include Bedford city (FIPS code 51515) in this update. We impute the value for this county with the state level estimate. For further information, please see the Census website: https://www.census.gov/geo/reference/county-changes.html

guidance, which defines any estimate with a CV greater than 0.3 as low reliability and to be used with extreme caution (King et al. 2015).

In addition to the race- and origin-specific files, aggregated county-level files are included in BenMAP that represent the fraction of the *total* county population in each education category.

J.3.3   Poverty Status

To determine the poverty status among different age, race and ethnicity groups at the county level, we utilize ACS 5-year estimates from 2012 to 2016. Each resulting dataset represents the fraction of the demographic group (defined by age, race, and ethnicity) in the county that falls below the federal poverty line. Following the imputation methodology described above, we generate poverty estimates for combinations of four race and two ethnicity categories (White; Black or African American; American Indian and Alaska Native; Asian and Pacific Islanders; Hispanic; and Non-Hispanic) and three age categories (0 to 17 years old, 18 to 64 years old, and 65 years old and up). We also include aggregated statistics for the county overall, for each age group, for each race, and for each ethnicity. The EPA Standard Variables dataset also includes one aggregated dataset (poor200) representing the fraction of the county-level population below 200% of the poverty line.

J.3.4   Unemployment Rates

We use monthly county-level unemployment rates from the Bureau of Labor Statistics for the period from February 2017 to February 2018. The value calculated in BenMAP-CE represents the average employment rate across all months within this time period. The unemployment rate was calculated by dividing the number of unemployed workers in each county by the county's total labor force. We note that the labor force does not include discouraged workers, people who have permanently left the labor force (e.g., retirees), and individuals who have not yet entered the labor force. Thus, the rates do **not** reflect the fraction of the population in BenMAP-CE that is not employed. We impute values for two counties using state unemployment rates, calculated by taking the total number of unemployed individuals within the state for each month, dividing by the total labor force within the state for the month, and averaging over all months.

J.3.5   Health Insurance

We use 2016 data from the Small Area Health Insurance Estimates (SAHIE) collected by the U.S. Census Bureau to calculate the percentage of individuals without health insurance in each county. Total numbers of individuals with and without health insurance by county were downloaded for the year 2016. The SAHIE data provides the number of individuals with and without health insurance by county, disaggregated into three age groups: 18 to 65, 40 to 65, and under 65. Using these data, we calculated the percent of people uninsured by county for the age groups 1 to 17, 18-39, 40-64, and under 65. As an example of this calculation, we subtracted the 18 to 65 uninsured population from the under 65 uninsured population to generate the 1 to 17 uninsured population. This value is then divided by the total 1 to 17 population (derived in the same manner) to estimate the uninsured rate.

There were two identical rows of data for the Washington, D.C. area (FIPS Code 11000 and 11001); we use FIPS code 11001. Shannon County, SD changed FIPS code in 2015 from 46113 to 46102; we changed the code to 46113 to match the BenMAP-CE county grid definition. Lastly, Bedford City, VA merged with Bedford County, VA in 2013, so data for Bedford County, VA (FIPS 51019) were used for Bedford City, VA (FIPS 51515).

J.3.6   Blue Collar Workers

We use five-year estimates (2012-2016) from the ACS to estimate the fraction of each county's labor force employed in white collar and blue collar occupations. The dataset includes the number of employed individuals over 16 that work within five occupation categories. We assign each of these five occupations to either the blue collar or white collar designation, as shown in Table J-10.

Table J-10. Mapping Occupations to Blue Collar and White Collar Designations

Occupation	Designation
Management, business, science and arts	White collar
Service	White collar
Sales and office	White collar
Natural resources, construction and maintenance	Blue collar
Production, transportation and material moving	Blue collar

We calculate the fraction of each county in blue collar professions by dividing the total number of individuals employed in blue collar jobs by the total number of employed individuals within each county. The same calculation is done for white collar professions. We adjust FIPS codes to match the BenMAP-CE county grid definition, as described in section J.3.5.

Appendix K.   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.

K.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.

K.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 *n*th percentile point is equal to:

$$(n-1) \times \frac{100}{N} + \frac{100}{2N}$$

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 2.5th, 7.5th, ..., 97.5th. 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.

K.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 0.5th, 1.5th, ..., and 99.5th 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.

K.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) to derive a pooled distribution. This pooled distribution incorporates information from all the studies used in the pooling procedure.

K.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.

K.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.

K.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.

K.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:

$$\beta_{fe} = \sum w_i \times \beta_i$$

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

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

Table K-1 shows the relevant calculations for this pooling for three sample studies.

Table K-1. Example of Fixed Effects Model Calculations

Study	B_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			?=508.16	?=1.000	?=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, vfe, is the inverse of the sum of the variances, or 0.00197. (The sums of the βi and vi are not shown, since they are of no importance. The sum of the 1/vi is S, used to calculate the weights. The sum of the weights, wi, i=1, ..., n, is 1.0, as expected.)

K.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 βi 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 PM_{10} on mortality, for example, if the composition of PM_{10} varies among study locations the underlying relationship between mortality and PM_{10} may be different from one study location to another. For example, fine particles make up a greater fraction of PM_{10} 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:

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

Under the null hypothesis that there is a single underlying parameter, β, of which all the βi'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 Q_w, the between-study variance, η^2, is:

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

It can be shown that the denominator is always positive. Therefore, if the numerator is negative (i.e., if 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 + η2.

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

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

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

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

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 β'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 β'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 β's. The formulas above for calculating fixed effects weights, for testing the fixed effects hypothesis, and for calculating random effects weights can all be used by substituting the mean incidence change for the ith health impact function for βi and the variance of incidence change for the ith health impact function for vi.

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 β'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 β's. The formulas above for calculating fixed effects weights, for testing the fixed effects hypothesis, and for calculating random effects weights can all be used by substituting the mean dollar benefit change for the ith valuation for βi 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.

K.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).

K.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.

K.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 L.   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.

L.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.

L.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!

L.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).

L.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

L.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.

BenMAP supports pollutant values present in the active setup. By default, these include Ozone and PM2.5 (no subscripts needed) for the United States setup. BenMAP Command Line will support additional pollutants as well, so long as they are present in

the active setup selected at the start of the Command Line Script. These values can be found in 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

L.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.

L.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

-V N A

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.

L.3.2.3   Monitor Rollback

[NOTE: Monitor Rollback is currently disabled in BenMAP Command Line. We are aware of this issue and working to include the functionality in upcoming BenMAP releases.]

// MonitorRollback

BaselineFilename   = -BaselineFilename;

// RollbackOptions

Percentage   = -Percentage;

Increment   = -Increment;

// RollbackToStandardOptions

Standard   = -Standard;

Metric   = -Metric;

Ordinality   = -Ordinality;

InterdayRollbackMethod = -InterdayRollbackMethod;

IntradayRollbackMethod = -IntradayRollbackMethod;

L.3.3 Run CFG

The command line version of BenMAP does not support creation of new .cfgx files, both because this would be quite cumbersome to do in plain text, and because it probably is not needed. Slight modifications of existing .cfgx 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 .cfgx file to run.

-ResultsFilename <filename>

Specifies the .cfgrx 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 .cfgx file.

-ControlAQG <filename>

Specifies the control air quality grid file to use when running the configuration - overrides whatever value is present in the .cfgx file.

-Year <Integer>

Year in which to run the configuration (this will affect the population numbers used) - overrides whatever value is present in the .cfgx 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 .cfgx file.

-Threshold <real>

Threshold to use when running the configuration - overrides whatever value is present in the .cfgx file.

L.3.4   Run APV

The command line version of BenMAP does not support creation of new .apvx files, both because this would be quite cumbersome to do in plain text, and because it probably is not needed. Slight modifications of existing .apvx 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 .apvx file to run.

-ResultsFilename <filename>

Specifies the .apvrx file to save the results in.

Optional Parameters

-CFGRFilename <filename>

Specifies the .cfgrx file to use when running the APV configuration - note that this file must contain the same set of results which the .cfgrx file originally used to generate the .apvx file contained. Overrides whatever value is present in the .apvx 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 .apvx 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 .apvx 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>

L.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 .cfgrx files; and APV Configuration Results Reports, which can be generated from .apvrx 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.

L.3.5.1   Audit Trail

Audit trail reports require only the parameters described in the "Generate Report" section.

L.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. Custom fields include DataSet, Endpoint Group, Endpoint, Pollutant, Metric, Seasonal Metric, Metric Statistic, Author, Year, Location, Other Pollutants, Qualifier, Reference, Race, Ethnicity, Gender, Start Age, End Age, Function, Incidence Dataset, Prevalence Dataset, Beta, Disbeta, P1Beta, P2Beta, A, NameA, B, NameB, C and NameC. 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, Baseline, Percent of Baseline, Standard Deviation, Variance, and Percentiles. If this parameter is not present, all fields will be included in the report.

L.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.

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).

L.4 Example 1

VARIABLES

%CFG%	C:\BenMAP\CommandLine\Configurations\PM25 Wizard.cfgx
%APV%	C:\BenMAP\CommandLine\Configuations\PM25 Wizard.apvx
%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.aqgx
-GridType	CMAQ 12km
-Pollutant	PM2.5

MonitorDirect

-InterpolationMethod	VNA
-MonitorDataType	Library
-MonitorDataSet	EPA Standard Monitors
-MonitorYear	2002
-MaxDistance	50

CREATE AQG

-Filename	%AQG%\PM25_2002Control_50km.aqgx
-GridType	CMAQ 12km
-Pollutant	PM2.5

MonitorRollback

```
-InterpolationMethod    VNA
-MonitorDataType        Library
-MonitorDataSet         EPA Standard Monitors
-MonitorYear            2002
-RollbackGridType       State
-MaxDistance            50
```

RollbackToStandardOptions

```
-Stanard                65
-Metric                 D24HourMean
-InterdayRollbackMethodIncremental
```

RUN CFG

```
-CFGFilename            %CFG%
-ResultsFilename        %RESULTSDIR%\PM25_2002_50km.aqgx
-BaselineAQG            %AQG%\PM25_2002Baseline_50km.aqgx
-ControlAQG             %AQG%\PM25_2002Control_50km.aqgx
```

RUN APV

```
-APVFilename            %APV%
-ResultsFilename        %RESULTSDIR%\PM25_2002_50km.apvrx
-CFGRFilename           %RESULTSDIR%\PM25_2002_50km.cfgrx
-IncidenceAggregation   Nation
-ValuationAggregation   Nation
```

GENERATE REPORT APVR

```
-InputFile              %RESULTSDIR%\PM25_2002_50km.apvrx
-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.apvrx
-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

L.5   Example 2

VARIABLES

%CFG%	C:\BenMAP\CommandLine\Configurations\PM25 Wizard.cfgx
%APV%	C:\BenMAP\CommandLine\Configurations\PM25 Wizard.apvx
%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.aqgx
-GridType	County
-Pollutant	PM2.5

MonitorDirect

-InterpolationMethod	VNA
-MonitorData Type	Library
-MonitorDataSet	EPA Standard Monitors
-MonitorYear	2004

CREATE AQG

-Filename	%AQG%\PM25_2004_Control.aqgx
-GridType	County
-Pollutant	PM2.5

MonitorRollback

-InterpolationMethod	VNA
-MonitorDataType	Library

```
-MonitorDataSet        EPA Standard Monitors
-MonitorYear           2004
-RollbackGridType      State
-MaxDistance           50
```

RollbackToStandardOptions

```
-Standard                 35
-Metric                   D24HourMean
-InterdayRollbackMethod   Incremental
```

RUN CFG
```
-CFGFilename           %CFG%
-ResultsFilename       %RESULTSDIR%\PM25_2004.cfgrx
-BaselineAQG           %AQG%\PM25_2004Baseline.aqgx
-ControlAQG            %AQG%\PM25_2004Control.aqgx
```

RUN APV
```
-APVFilename           %APV%
-ResultsFilename       %RESULTSDIR%\PM25_2004.apvrx
-CFGRFilename          %RESULTSDIR%\PM25_2004.cfgrx
-IncidenceAggregaton   Nation
-IncidenceAggregation  Nation
```

GENERATE REPORT APVR

```
-InputFile        %RESULTSDIR%\PM25_2004.apvrx
-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.apvrx
-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 M. Function Editor

The function editor is used to develop both health impact functions and valuation functions. This appendix describes the syntax of this editor.

M.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.

M.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.)

M.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"" ");

M.4   Operations

Arithmetical

+, −, ×, /;

div - integer division;

mod - modulo;

^ - power of;

- - negate;

Logical

<, <=, >=, >, <>, =;

and, or, xor, not;

Bitwise

and, or, xor;

~ - negate;

M.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

M.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 PM$_{10}$ 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 (PM$_{2.5}$) 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 (AHRQ). 2007. Healthcare Cost and Utilization Project. National Inpatient Sample (NIS) Rockville, Maryland. Available at: http://www.ahrq. gov/data/hcup/ [accessed February, 2012].

Agency for Healthcare Research and Quality (AHRQ). 2009. HCUPnet, Healthcare Cost and Utilization Project. Rockville, MD. Available on the Internet at http://hcupnet.ahrq.gov

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.

American Lung Association. 2010a. Trends in COPD (Chronic Bronchitis and Emphysema): Morbidity and Mortality. American Lung Association Epidemiology and Statistics Unit

Research and Program Services Division. February. Table 4. Based on NHIS data (CDC, 2008). Available at http://www.lungusa.org/finding-cures/our-research/trend-reports/copd- trend-report.pdf.

American Lung Association. 2010b. Trends in Asthma Morbidity and Mortality. American Lung Association Epidemiology and Statistics Unit Research and Program Services Division. February. Table 7. Based on NHIS data (CDC, 2008). Available at http://www.lungusa.org/ finding-cures/our-research/trend-reports/asthma-trend-report.pdf.

Babin, S. M., H. S. Burkom, et al. 2007. "Pediatric patient asthma-related emergency department visits and admissions in Washington, DC, from 2001-2004, and associations with air quality, socio-economic status and age group." Environ Health 6: 9.

Bell, M. L., K. Ebisu, et al. 2008. "Seasonal and Regional Short-term Effects of Fine Particles on Hospital Admissions in 202 US Counties, 1999-2005." American Journal of Epidemiology 168 (11): 1301-1310.

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.

Burnett, R. T., C. Arden Pope III, M. Ezzati, C. Olives, S. S. Lim, S. Mehta, H. H. Shin, G. Singh, B. Hubbell, M. Brauer, H. R. Anderson, K. R. Smith, J. R. Balmes, N. G. Bruce, H. Kan, F. Laden, A. Prüss-Ustün, M. C. Turner, S. M. Gapstur, W. R. Diver, and A. Cohen. 2014. An integrated risk function for estimating the global burden of disease attributable to ambient fine particulate matter exposure. Environmental Health Perspectives. Vol. 122 (4): 397-403.

Centers for Disease Control and Prevention (CDC). 2008. National Center for Health Statistics. National Health Interview Survey, 1999-2008.

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.

Cohen, A. J., M. Brauer, R. Burnett, H.R. Anderson, J. Frostad, K. Estep, K. Balakrishnan, B. Brunekreef, L. Dandona, R. Dandona, V. Feigin, G. Freedman, B. Hubbell, A. Jobling, H. Kan, L. Knibbs, Y. Liu, R. Martin, L. Morawska, C. A. Pope III, H. Shin, K. Straif, G. Shaddick, M. Thomas, R. van Dingenen, A. van Donkelaar, T. Vos, C. J. L. Murray, M. H. Forouzanfar. 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Disease Study 2015. Lancet. Vol. 389: 1907-1918.

Collet, D. 1994. Modelling Survival Data in Medical Research. Chapman & Hall: New York.

Cropper, Maureen R. and Alan J. Krupnick. 1999. The social costs of chronic heart and lung disease, in Maureen Cropper, Valuing Environmental Benefits: Selected Essays of Maureen Cropper, Edward Elgar: Northampton, MA.

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.

Dewanji, A. and S. H. Moolgavkar. 2000. A Poisson process approach for recurrent event data with environmental covariates. Environmetrics. Vol. 11: 665-673.

Dewanji, A. and S. H. Moolgavkar. 2002. Choices of stratification in Poisson process analysis of recurrent event data with environmental covariates. Stat. Med. Vol. 21: 3383-3393.

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.

Fung, K. Y. K., Daniel; Chen, Yue; Burnett, Rick; Cakmak, Sabit. 2003. Comparison of time series and case-crossover analyses of air pollution and hospital admission data. International Journal of Epidemiology. Vol. 32: 1064-1070.

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.

Glad, J.A., L.L. Brink, E.O. Talbott, P.C. Lee, X. Xu, M. Saul, and J. Rager. 2012. The Relationship of Ambient Ozone and PM2.5 Levels and Asthma Emergency Department Visits: Possible Influence of Gender and Ethnicity. Archives of Environmental & Occupational Health. Vol 62 (2): 103-108.

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 PM$_{2.5}$ 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 PM_{10}/mortality associations: an investigations of at-risk subpopulations. Journal of Exposure Analysis and Environmental Epidemiology. Vol. 6 (1): 79-95.

Ito, K., G. D. Thurston and R. A. Silverman. 2007. Characterization of $PM_{2.5}$, gaseous pollutants, and meteorological interactions in the context of time-series health effects models. J Expo Sci Environ Epidemiol. Vol. 17 Suppl 2: S45-60.

Jerrett, M., R. Burnett, A. I. Pope, K. Ito, G. Thurston, D. Krewski, Y. Shi, E. Calle and M. Thun. 2009. Long-Term Ozone Exposure and Mortality. The New England Journal of Medicine. Vol. 360 (11): 1085-95.

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.Katsouyanni, K and J.M. Samet. 2009. Air Pollution and Health: A European and North American Approach (APHENA). Health Effects Institute, Research Report 142.

Kennedy, P. 1990. A Guide to Econometrics. MIT Press: Cambridge, MA.

King, K., Starsinic, M., Hernandez Viver, A., Beaghen, M. 2015, "The Reliability of ACS 5-Year Estimates of Race Groups and American Indian and Alaska Native Populations", U.S. Census Bureau.

Kloog, I., B.A. Coull, A. Zanobetti, P. Koutrakis, J.D. Schwartz. 2012. Acute and Chronic Effects of Particles on Hospital Admissions in New-England. PLoS ONE. Vol 7 (4): 1-8.

Krewski, D., M. Jerrett, R. Burnett, R. Ma, E. Hughes, Y. Shi, M. C. Turner, C. A. I. Pope, G. Thurston, E. Calle and M. J. Thun. 2009. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality. HEI Research Report, 140, Health Effects Institute, Boston, 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.

Mar, T.F., J.Q. Koenig. 2009. Relationship between visits to emergency departments for asthma and ozone exposure in greater Seattle, Washington. Annals of Allergy, Asthma & Immunology, Vol. 103(6):474-479.

Mar, T. F., J. Q. Koenig, et al. 2010. Associations between asthma emergency visits and particulate matter sources, including diesel emissions from stationary generators in Tacoma, Washington. Inhal Toxicol 22 (6): 445-8.

Mar, T. F., T. V. Larson, et al. 2004. An analysis of the association between respiratory symptoms in subjects with asthma and daily air pollution in Spokane, Washington." Inhal Toxicol 16 (13): 809-15.

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.

Mortimer, K. M., L. M. Neas, D. W. Dockery, S. Redline and I. B. Tager. 2002. The effect of air pollution on inner-city children with asthma. Eur Respir J. Vol. 19 (4): 699-705.

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.

O'Connor, G.T., L. Neas, B. Vaughn, M. Kattan, H. Mitchell, E. F. Crain, R. Evans III, R. Gruchalla, W. Morgan, J. Stout, G. K. Adam, and M. Lippmann. 2008. Acute respiratory health effects of air pollution on children with asthma in US inner cities. Journal of Allergy and Clinical Immunology, Vol. 121(5): 1133-1139.

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.

Peng, R. D., M. L. Bell, et al. 2009. Emergency admissions for cardiovascular and respiratory diseases and the chemical composition of fine particle air pollution. Environ Health Perspect, Vol. 117 (6): 957-63.

Peng, R. D., H. H. Chang, et al. 2008. "Coarse particulate matter air pollution and hospital admissions for cardiovascular and respiratory diseases among Medicare patients." Jama, Vol. 299 (18): 2172-9.

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, J. B. Muhlestein, et al. 2006. "Ischemic heart disease events triggered by short-term exposure to fine particulate air pollution." Circulation, Vol. 114 (23): 2443-8.

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 PM_{10} 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 PM_{10} pollution in Utah Valley. Environ Res. Vol. 58 (2): 204-19.

Roger V.L., A.S. Go, D.M. Lloyd-Jones, E.J. Benjamin, J.D. Berry, W.B. Borden, D.M. Bravata, S. Dai, E.S. Ford, C.S. Fox, H.J. Fullerton, C. Gillespie, S.M. Hailpern, J.A. Heit, V.J. Howard, B.M. Kissel, S.J. Kittner, D.T. Lackland, J.H. Lichtman, L.D. Lisabeth, D.M. Makuc, G.M. Marcus, A. Marelli, D.B. Matchar, C.S. Moy, D. Mozaffarian, M.E. Mussolino, G. Nichol, N.P. Paynter, E.Z. Soliman, P.D. Sorlie, N. Sotoodehnia, T.N. Turan, S.S. Virani, N.D. Wong, D. Woo and M.B. Turner; on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics— 2012 update: a report from the American Heart Association. Circulation. 2012;125:e2-e220.

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.

Sarnat, J. A., S. E. Sarnat, W. D. Flanders, H. H. Chang, J. Mulholland, L. Baxter, V. Isakob, and H. Özkaynak. 2013. Spatiotemporally resolved air exchange rate as a modifier of acute air pollution-related morbidity in Atlanta. Journal of Exposure Science and Environmental Epidemiology 23, 606-615.

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.

Schildcrout, J. S., L. Sheppard, T. Lumley, J. C. Slaughter, J. Q. Koenig and G. G. Shapiro. 2006. Ambient air pollution and asthma exacerbations in children: an eight-city analysis. Am J Epidemiol. Vol. 164 (6): 505-17.

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., C. Spix, G. Touloumi, L. Bacharova, T. Barumamdzadeh, A. le Tertre, T. Piekarksi, A. Ponce de Leon, A. Ponka, G. Rossi, M. Saez and J.P. Schouten. 1996. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. J. Epidemiol. Community Health. Vol. 50 Suppl. 1: S3-11.

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.

Slaughter, J. C., E. Kim, et al. 2005. "Association between particulate matter and emergency room visits, hospital admissions and mortality in Spokane, Washington." J Expo Anal Environ Epidemiol 15(2): 153-9.

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.

Schildcrout, J. S., L. Sheppard, T. Lumley, J. C. Slaughter, J. Q. Koenig and G. G. Shapiro. 2006. Ambient air pollution and asthma exacerbations in children: an eight-city analysis. Am J Epidemiol. Vol. 164 (6): 505-17.

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., C. Spix, G. Touloumi, L. Bacharova, T. Barumamdzadeh, A. le Tertre, T. Piekarksi, A. Ponce de Leon, A. Ponka, G. Rossi, M. Saez and J.P. Schouten. 1996. Methodological issues in studies of air pollution and daily counts of deaths or hospital admissions. J. Epidemiol. Community Health. Vol. 50 Suppl. 1: S3-11.

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.

Slaughter, J. C., E. Kim, et al. 2005. "Association between particulate matter and emergency room visits, hospital admissions and mortality in Spokane, Washington." J Expo Anal Environ Epidemiol 15(2): 153-9.

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.

Sullivan, J., L. Sheppard, et al. 2005. "Relation between short-term fine-particulate matter exposure and onset of myocardial infarction." Epidemiology 16 (1): 41-8.

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: $PM_{2.5}$ 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.

U.S. EPA. 2012. Regulatory Impact Analysis for the Final Revisions to the National Ambient Air Quality Standards for Particulate Matter. EPA-452/R-12-003. Office of Air Quality Planning and Standards, Health and Environmental Impacts Division. December. Available at http://www.epa.gov/pm/2012/finalria.pdf.

U.S. EPA Science Advisory Board (U.S. EPA-SAB). 2004. Advisory on Plans for Health Effects Analysis in the Analytical Plan for EPA's Second Prospective Analysis—Benefits and Costs of the Clean Air Act, 1990–2020. EPA-SAB-COUNCIL-ADV-04-002. March. Available at http://yosemite.epa.gov/sab%5CSABPRODUCT.NSF/08E1155AD24F871C85256E5400433D5D/$File/council_adv_04002.pdf.

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 ($PM_{2.5}$) air pollution and selected causes of postneonatal infant mortality in California. Environmental Health Perspectives. Vol. 114: 786-790.

Woods & Poole Economics Inc. 2012. Complete Demographic Database. Washington, DC. http://www.woodsandpoole.com/index.php.

Zanobetti, A., M. Franklin, et al. 2009. Fine particulate air pollution and its components in association with cause-specific emergency admissions. Environmental Health 8: 58-60.

Zanobetti, A. and Schwartz, J., 2008. Mortality displacement in the association of ozone with mortality: an analysis of 48 cities in the United States. American Journal of Respiratory and Critical Care Medicine 177(2): 184-189.

Zanobetti A. and Schwartz, J. 2006. Air pollution and emergency admissions in Boston, MA. J Epidemiol Community Health 60 (10): 890-5.