Health Risk and Exposure Assessment for Ozone

Second External Review Draft

DISCLAIMER

This draft document has been prepared by staff from the Risk and Benefits Group, Health and Environmental Impacts Division, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency. Any findings and conclusions are those of the authors and do not necessarily reflect the views of the Agency. This draft document is being circulated to facilitate discussion with the Clean Air Scientific Advisory Committee to inform the EPA's consideration of the ozone National Ambient Air Quality Standards.

This information is distributed for the purposes of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent any Agency determination or policy.

Questions related to this preliminary draft document should be addressed to Dr. Bryan Hubbell, U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, C539-07, Research Triangle Park, North Carolina 27711 (email: hubbell.bryan@epa.gov).

EPA-452/P-14-004a
February 2014

Health Risk and Exposure Assessment for Ozone

Second External Review Draft

U.S. Environmental Protection Agency
Office of Air and Radiation
Office of Air Quality Planning and Standards
Health and Environmental Impacts Division
Risk and Benefits Group
Research Triangle Park, North Carolina 27711

This page left intentionally blank.

TABLE OF CONTENTS

TABLE OF CONTENTS ..... **v**

TABLE OF FIGURES ..... **xi**

TABLE OF TABLES ..... **xxv**

1 INTRODUCTION **1-1**

1.1 HISTORY 1-3

1.2 CURRENT RISK AND EXPOSURE ASSESSMENT: GOALS AND PLANNED APPROACH 1-5

1.3 ORGANIZATION OF DOCUMENT 1-6

2 OVERVIEW OF EXPOSURE AND RISK ASSESSMENT DESIGN **2-1**

2.1 POLICY-RELEVANT EXPOSURE AND RISK QUESTIONS 2-2

2.2 AIR QUALITY CHARACTERIZATION 2-4

2.2.1 O_3 Chemistry and Response to Changes in Precursor Emissions 2-5

2.2.2 Sources of O_3 and O_3 Precursors 2-6

2.2.3 Simulation of Meeting Existing and Alternative Standards 2-7

2.2.4 Consideration of Health Evidence 2-8

2.2.5 Exposures of Concern 2-9

2.2.6 Health Endpoints 2-9

2.2.7 Exposure and Concentration-response Functions for Health Endpoints 2-13

2.2.8 At-risk Populations 2-14

2.3 URBAN-SCALE MODELING OF INDIVIDUAL EXPOSURE 2-15

2.3.1 Microenvironmental O_3 Concentrations 2-16

2.3.2 Human Activity Patterns 2-17

2.3.3 Modeling of Exposures Associated with Simulating Just Meeting O_3 Standards 2-19

2.3.4 Considerations in Selecting Urban Case Study Areas for the Exposure Analysis 2-19

2.4 RISK ASSESSMENT 2-19

2.4.1 Attributable Risk 2-20

2.4.2 Modeling of Risk for Total Exposure to O_3 2-21

2.4.3 Distributions of Risk Across O_3 concentrations 2-22

2.5 MODELING OF RISKS ASSOCIATED WITH SIMULATING JUST MEETING O_3 STANDARDS 2-22

2.6 CONSIDERATIONS IN SELECTING URBAN CASE STUDY AREAS FOR THE RISK ANALYSIS 2-23

2.7 RISK CHARACTERIZATION 2-23

2.8 REFERENCES 2-25

3 SCOPE 3-1
3.1 OVERVIEW OF EXPOSURE AND RISK ASSESSMENTS FROM LAST REVIEW 3-2
3.1.1 Overview of Exposure Assessment from Last Review 3-2
3.1.2 Overview of Risk Assessment from Last Review 3-3
3.2 PLAN FOR THE CURRENT EXPOSURE AND RISK ASSESSMENTS 3-5
3.3 CHARACTERIZATION OF UNCERTAINTY AND VARIABILITY IN THE CONTEXT OF THE O_3 EXPOSURE AND RISK ASSESSMENT 3-7
3.4 AIR QUALITY CHARACTERIZATION 3-10
3.5 EXPOSURE ASSESSMENT 3-13
3.6 URBAN-SCALE LUNG FUNCTION RISK ANALYSES BASED ON APPLICATION OF RESULTS FROM CONTROLLED HUMAN EXPOSURE STUDIES 3-15
3.7 URBAN CASE STUDY AREA EPIDEMIOLOGY-BASED RISK ASSESSMENT .. 3-17
3.8 NATIONAL-SCALE MORTALITY RISK ASSESSMENT 3-22
3.9 PRESENTATION OF EXPOSURE AND RISK ESTIMATES TO INFORM THE O_3 NAAQS POLICY ASSESSMENT 3-24
3.10 REFERENCES 3-26
4 AIR QUALITY CONSIDERATIONS 4-1
4.1 INTRODUCTION 4-1
4.2 OVERVIEW OF O_3 MONITORING AND AIR QUALITY DATA 4-1
4.3 OVERVIEW OF URBAN-SCALE AIR QUALITY INPUTS TO RISK AND EXPOSURE ASSESSMENTS 4-4
4.3.1 Urban Case Study Areas 4-5
4.3.2 Recent Air Quality 4-9
4.3.3 Air Quality Adjustments for "Just Meeting" Existing and Potential Alternative O_3 Standards 4-14
4.4 OVERVIEW OF NATIONAL-SCALE AIR QUALITY INPUTS 4-30
4.5 UNCERTAINTIES IN MODELING OF RESPONSES TO EMISSION REDUCTIONS TO JUST MEET EXISTING AND POTENTIAL ALTERNATIVE STANDARDS .. 4-38
4.6 REFERENCES 4-52
5 CHARACTERIZATION OF HUMAN EXPOSURE TO O_3 5-1
5.0 OVERVIEW 5-1
5.1 SYNOPSIS OF O_3 EXPOSURE AND EXPOSURE MODELING 5-2
5.1.1 Human Exposure 5-2
5.1.2 Estimating O_3 Exposure 5-3
5.1.3 Modeling O_3 Exposure Using APEX 5-4
5.2 SCOPE OF THE EXPOSURE ASSESSMENT 5-6
5.2.1 Urban Areas Selected 5-6
5.2.2 Time Periods Simulated 5-7
5.2.3 Ambient Concentrations Used 5-9

5.2.4 Meteorological Data Used 5-10
5.2.5 Populations Simulated 5-11
5.2.6 Key Physiological Processes And Personal Attributes Modeled 5-15
5.2.7 Microenvironments Modeled 5-16
5.2.8 Model Output 5-17
5.3 EXPOSURE ASSESSMENT RESULTS 5-23
5.3.1 Overview 5-23
5.3.2 Exposure Modeling Results for Base Air Quality 5-24
5.3.3 Exposure Modeling Results for Simulations of Just Meeting Existing and Alternative O_3 Standards 5-25
5.4 TARGETED EVALUATION OF EXPOSURE MODEL INPUT AND OUTPUT DATA 5-35
5.4.1 Analysis of Time-Locaton-Activity Data 5-35
5.4.1.1 General Evaluation of CHAD Study Data: Historical and Recently Acquired Data 5-36
5.4.1.2 Exposure-Relevant Personal Attributes Included in CHAD and APEX Simulated Individuals 5-36
5.4.1.3 Evaluation of Afternoon Time Spent Outdoors for CHAD and Survey Participants 5-37
5.4.1.4 Evaluation of Afternoon Time Spent Outdoors for ATUS Survey Participants 5-38
5.4.1.5 Evaluation of Outdoor Time and Exertion Level for Asthmatics and Non-Asthmatics in CHAD 5-39
5.4.2 Characterization of Factors Influencing High Exposures 5-40
5.4.3 Exposure Results for additional at-risk populations and Lifestages, Exposure scenarios, and Air Quality Input Data Used 5-41
5.4.3.1 Exposures Estimated for All School-age Children During Summer Months, Neither Attending School or Performing Paid Work 5-41
5.4.3.2 Exposures Estimated for Outdoor Workers During Summer Months 5-43
5.4.3.3 Exposures Estimated for All School-age Children When Accounting for Averting Behavior 5-45
5.4.3.4 Comparison of APEX Estimated Exposures Using Three Different Base Case Air Quality Data Sets: AQS, VNA, and EVNA 5-46
5.4.3.5 Comparison of APEX Estimated Exposures Using Two Different Adjusted Air Quality Data Sets: Quadratic Rollback and HDDM 5-48
5.4.4 Limited Performance Evaluations 5-49
5.4.4.1 Personal Exposure Comparisons 5-49
5.4.4.2 Ventilation Rate Comparisons 5-51
5.4.4.3 Evaluation of Longitudinal Profile Methodology 5-55
5.5 VARIABILITY AND UNCERTAINTY 5-56

5.5.1 Treatment of Variability 5-56
5.5.2 Characterization of Uncertainty 5-57
5.6 KEY OBSERVATIONS 5-65
5.7 REFERENCES 5-71
6 CHARACTERIZATION OF HEALTH RISKS BASED ON CONTROLLED HUMAN EXPOSURE STUDIES 6-1
6.1 INTRODUCTION 6-1
6.1.1 Development of Approach for Current Risk Assessment 6-2
6.1.2 Comparison of Controlled Human Exposure- and Epidemiologic-based Risk Assessments 6-3
6.2 SCOPE OF LUNG FUNCTION HEALTH RISK ASSESSMENT 6-3
6.2.1 Selection of Health Endpoints 6-4
6.2.2 Approach for Estimating Health Risk Based on Controlled Human Exposure Studies 6-5
6.2.3 Controlled Human Exposure Studies 6-6
6.2.4 The McDonnell-Stewart-Smith (MSS) Model 6-8
6.2.5 The Exposure-Response Function Approach Used in Prior Reviews 6-16
6.3 O_3 RISK ESTIMATES 6-21
6.3.1 Lung Function Risk Estimates Based on the McDonnell-Stewart-Smith Model 6-22
6.3.2 Lung Function Risk Estimates Based on the Exposure-Response Functions Approach Used in Prior Reviews 6-28
6.3.3 Comparison of the MSS Model with the Exposure-Response Function Approach 6-29
6.4 EVALUATION OF THE MSS MODEL 6-35
6.4.1 Summary of Published Evaluations 6-35
6.4.2 Children 6-35
6.4.3 Threshold vs. Non-Threshold Models 6-36
6.5 CHARACTERIZATION OF UNCERTAINTY 6-37
6.5.1 Statistical Model Form 6-38
6.5.2 Convergence of APEX Results 6-41
6.5.3 Application of Model for All Lifestages 6-42
6.5.4 Application of Model for Asthmatic Children 6-43
6.5.5 Interaction Between O_3 and Other Pollutants 6-43
6.5.6 Qualitative Assessment of Uncertainty 6-43
6.6 DISCUSSION 6-46
6.7 REFERENCES 6-50
7 CHARACTERIZATION OF HEALTH RISK BASED ON EPIDEMIOLOGICAL STUDIES 7-1
7.1 GENERAL APPROACH 7-1
7.1.1 Basic Structure of the Risk Assessment 7-1
7.1.2 Calculating O_3-Related Health Effects Incidence 7-11
7.2 AIR QUALITY CONSIDERATIONS 7-13

7.3 SELECTION OF MODEL INPUTS AND ASSUMPTIONS 7-15
7.3.1 Selection of Urban Study Areas 7-15
7.3.2 Selection of Epidemiological Studies and Specification of Concentration-Response Functions 7-17
7.3.3 Baseline Health Effect Incidence and Prevalence Data 7-30
7.3.4 Population (demographic) Data 7-31
7.4 ADDRESSING VARIABILITY AND UNCERTAINTY 7-31
7.4.1 Treatment of Key Sources of Variability 7-34
7.4.2 Qualitative Assessment of Uncertainty 7-38
7.4.3 Description of Core and Sensitivity Analyses 7-45
7.5 URBAN STUDY AREA RESULTS 7-47
7.5.1 Assessment of Health Risk After Just Meeting the Existing 75 ppb standard 7-65
7.5.2 Assessment of Health Risk Associated with Simulating Meeting Potential Alternative Standards of 70, 65, and 60 ppb 7-67
7.5.3 Sensitivity Analyses Designed to Enhance Understanding of the Core Risk Estimates 7-72
7.6 KEY OBSERVATIONS REGARDING OVERALL CONFIDENCE IN THE RISK ASSESSMENT AND RISK ESTIMATES 7-83
7.7 REFERENCES 7-86
8 NATIONAL SCALE MORTALITY RISK BURDEN BASED ON APPLICATION OF RESULTS FROM EPIDEMIOLOGICAL STUDIES 8-1
8.1 NATIONAL-SCALE ASSESSMENT OF MORTALITY RELATED TO O_3 EXPOSURE 8-1
8.1.1 Methods 8-2
8.1.2 Results 8-6
8.1.3 Sensitivity Analysis 8-15
8.1.4 Discussion 8-19
8.2 EVALUATING THE REPRESENTATIVENESS OF THE URBAN STUDY AREAS IN THE NATIONAL CONTEXT 8-21
8.2.1 Analysis Based on Consideration of National Distributions of Risk-Related Attributes 8-22
8.2.2 Analysis Based on Consideration of National Distribution of O_3-Related Mortality Risk 8-45
8.2.3 Analysis Based on Consideration of National Responsiveness of O_3 Concentrations to Emissions Changes 8-49
8.2.5 Discussion 8-77
8.3 REFERENCES 8-78
9 SYNTHESIS 9-1
9.1 INTRODUCTION **9-2**
9.2 SUMMARY OF KEY RESULTS 9-2
9.2.1 Air Quality Considerations (Chapter 4) 9-2
9.2.2 Human Exposure Modeling (Chapter 5) 9-8

9.2.3 Health Risks Based on Controlled Human Exposure Studies (Chapter 6) 9-13
9.2.4 Health Risks Based on Epidemiological Studies (Chapters 7 and 8) 9-18
9.3 COMPARISON OF RESULTS ACROSS EXPOSURE, LUNG FUNCTION RISK, AND EPIDMEIOLOGY-BASED MOTALITY AND MORBIDITY RISK ANALYSES 9-27
9.3.1 Evaluation of Exposures and Risks After Just Meeting the Existing Standard ... 9-27
9.3.2 Reductions in Exposure and Risk Metrics after Just Meeting Alternative Standards 9-31
9.4 OVERALL ASSESSMENT OF REPRESENTATIVENESS OF EXPOSURE AND RISK RESULTS 9-36
9.4.1 Representativeness of Selected Urban Case Study Areas in Reflecting Areas Across the Nation with Elevated Risk 9-36
9.4.2 Representativeness of Selected Urban Case Study Areas in Reflecting Responsiveness of Risk to Just Meeting Existing and Alternative O_3 Standards 9-37
9.5 OVERALL ASSESSMENT OF CONFIDENCE IN EXPOSURE AND RISK RESULTS 9-38
9.5.1 Uncertainties in Modeling O_3 Responses to Meeting Standards 9-39
9.5.2 Uncertainties in Modeling Exposure and Lung-function Risk 9-40
9.5.3 Uncertainties in Modeling Epidemiological-based Risk 9-40
9.6 OVERALL INTEGRATED CHARACTERIZATION OF RISK IN THE CONTEXT OF KEY POLICY RELEVANT QUESTIONS 9-42
9.7 REFERENCES 9-47

TABLE OF FIGURES

Figure 2-1. Overview of Exposure and Risk Assessment Design. 2-2
Figure 2-2. Causal Determinations for O_3 Health Effects. 2-11
Figure 3-1. Conceptual Diagram for Air Quality Characterization in the Health REA. 3-10
Figure 3-2. Conceptual Diagram for Population Exposure Assessment..... 3-13
Figure 3-3. Conceptual Diagram of O_3 Lung Function Health Risk Assessment Based on Controlled Human Exposure Studies. 3-15
Figure 3-4. Conceptual Diagram of Urban Case Study Area Health Risk Assessment Based on Results of Epidemiology Studies. 3-18
Figure 3-5. Conceptual Diagram of National O_3 Mortality Risk Assessment Based on Results of Epidemiology Studies. 3-23
Figure 4-1. Map of Monitored 8-hour O_3 Design Values for the 2006-2008 Period 4-3
Figure 4-2. Map of Monitored 8-hour O_3 Design Values for the 2008-2010 Period..... 4-4
Figure 4-3. Flowchart of Air Quality Data Processing for Different Parts of the Urban-scale Risk and Exposure Assessments 4-5
Figure 4-4. Trends in Annual 4th Highest 8-hour Daily Maximum O_3 Concentrations in ppb for the 15 Urban Case Study Areas for 2006-2010. Urban areas are grouped into 3 regions: Eastern (top), Central (middle), and Western (bottom). 4-7
Figure 4-5a. Maps of the 5 Eastern U.S. Urban Case Study Areas Including O_3 Monitor Locations..... 4-11
Figure 4-5b. Maps of the 5 Central U.S. Urban Case Study Areas Including O_3 Monitor Locations..... 4-12
Figure 4-5c. Maps of the 5 Western U.S. Urban Case Study Areas Including O_3 Monitor Locations..... 4-13
Figure 4-6. Flowchart of HDDM adjustment methodology to inform risk and exposure assessment. 4-17
Figure 4-7. Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard. Values are based on the Zanobetti & Schwartz study areas for April-October of 2006-2008 4-21

Figure 4-8. Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard. Values are based on the Zanobetti & Schwartz study areas for June-August of 2006-2008 4-22

Figure 4-9. Distributions of composite monitor 8-hour daily maximum values for the 12 urban case study areas in the epidemiology-based risk assessment. Plots depict values based on ambient measurements (base), and values obtained with the HDDM adjustment methodology showing attainment of 75, 70, 65 and 60 ppb standards. Values shown are based on CBSAs for April-October of 2007. Note that the HDDM adjustment technique was not able to adjust air quality to show attainment of a 60 ppb standard in New York, so no boxplot is shown for that case..... 4-25

Figure 4-10. Distributions of composite monitor 8-hour daily maximum values for the 12 urban case study areas in the epidemiology-based risk assessment. Plots depict values based on ambient measurements (base), and values obtained with the HDDM adjustment methodology showing attainment of 75, 70, 65 and 60 ppb standards. Values shown are based on CBSAs for April-October of 2009. Note that Detroit air quality was meeting 75 ppb in 2008-2010, and the HDDM adjustment technique was not able to adjust air quality to show attainment of a 60 ppb standard in New York, so no boxplots are shown for those cases 4-26

Figure 4-11. Maps showing the 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in Atlanta based on 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids 4-28

Figure 4-12. Maps showing the 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in New York based on 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids 4-29

Figure 4-13. Maps of 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in Houston for 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids. 4-30
Figure 4-14. May-September average 8-hour daily maximum O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation..... 4-32
Figure 4-15. June-August average 8-hour daily 10am-6pm mean O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation..... 4-33
Figure 4-16. April-September average 1-hour daily maximum O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation..... 4-34
Figure 4-17. Frequency and Cumulative Distributions of the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells 4-35
Figure 4-18. 2006-2008 O_3 Design Values Versus 2006-2008 Fused Seasonal Average O_3 Levels for the CMAQ 12km Grid Cells Containing O_3 Monitors..... 4-37
Figure 5-1. Conceptual Framework Used for Estimating Study Area Population O_3 Exposure Concentrations 5-8
Figure 5-2. Percent of asthmatic school-age children in all study areas with at least one O_3 exposure at or above 60 ppb-8hr while at moderate or greater exertion using base air quality (2006-2010), stratified by year (top left panel) or by study area (bottom left panel). 5-20
Figure 5-3. Percent of asthmatic school-age children in Atlanta with at least one O_3 exposure at or above 60 ppb-8hr (left top panel), 70 ppb-8hr (middle top panel), and 80 ppb-8hr (right top panel while at moderate or greater exertion, years 2006-2010 air quality adjusted to just meet the existing and alternative O_3 standard levels. The multi-panel display (bottom) illustrates the same exposure results expanded to

reflect individual data points by year, standard averaging period, and benchmark level. 5-22
Figure 5-4. Percent of asthmatic school-age children in Atlanta with multiple O_3 exposures at or above 60 ppb-8hr while at moderate or greater exertion, years 2006-2010 air quality adjusted to just meet the existing and alternative O_3 standard levels. ... 5-23
Figure 5-5. Percent of all school-age children with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards 5-30
Figure 5-6. Percent of asthmatic school-age children with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards 5-31
Figure 5-7. Percent of all asthmatic adults with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards 5-32
Figure 5-8. Percent of all older adults with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards. 5-33
Figure 5-9. Percent of all school-age children with multiple daily maximum 8-hr average O_3 exposures at or above 60 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards. 5-34
Figure 5-10. Comparison of the percent of all school-age children having daily maximum 8-hr average O_3 concentration at or above 60 ppb during June, July, and August in Detroit 2007: using any available CHAD diary (“All CHAD Diaries”) or using CHAD diaries having no time spent in school or performing paid work (“No School/Work Diaries”) 5-42

Figure 5-11. Percent of workers between age 18-35 experiencing exposures at or above selected benchmark levels while at moderate or greater exertion using an outdoor worker approach (left panel) and a general population-based approach (right panel) considering air quality adjusted to just meet the existing standard in Atlanta, GA, Jun-Aug, 2006. 5-45
Figure 5-12. Percent of all school-age children (left panel) and asthmatic school-age children (right panel) having daily maximum 8-hr average O_3 concentration at or above benchmark levels during a 2-day simulation in Detroit, base air quality, August 1-2, 2007. Red bars indicate exposure results when considering effect of averting. 5-46
Figure 5-13. Comparison of APEX exposure results generated for three study areas (Atlanta, Detroit, and Philadelphia) using three different 2005 air quality input data sets: AQS, VNA, and eVNA. 5-47
Figure 5-14. Comparison of exposure results generated by APEX using two different air quality adjustment approaches to just meet the existing standard in Atlanta: quadratic rollback (left panel) and HDDM (right panel). 5-48
Figure 5-15. Distribution of daily average O_3 exposures (top panels) and daily afternoon outdoor time (bottom panels) and for DEARS study participants (left panels) and APEX simulated individuals (right panels) in Wayne County, MI, July-August 2006 5-50
Figure 5-16. Means (and range) of 6-day average personal O_3 exposures, measured and modeled (APEX), Upland Ca. Obtained from Figure 8-22 of US EPA (2007b). 5-51
Figure 5-17. Comparison of body mass normalized mean daily ventilation rates estimated by APEX (closed symbols) and by Brochu et al., 2006 (open symbols). 5-52
Figure 5-18. Comparison of body mass normalized mean daily ventilation rates in male and female school-age children (5-18) when correcting Brochu et al. (2006a) results with child appropriate VQ estimates. 5-53
Figure 5-19. Comparison of body mass normalized daily mean ventilation rates in school-age children (5-18) estimated using APEX and literature reported values. 5-55
Figure 5-20. Incremental increases in percent of all school-age children exposed to O_3 at or above 60 ppb-8hr for each study area, year 2006-2010 air quality. Average

percent (left panels), maximum percent (right panels), at least one exposure (top panels), at least two exposures (bottom panels) per year..... 5-69
Figure 6-1. Two-Compartment Model..... 6-1
Figure 6-2. Distribution of Responses (Lung Function Decrements in FEV1) Predicted by the MSS Model for 20-Year Old Individuals. Exposure to 100 ppb O_3 at Moderate Exercise (40 L/min, BSA=2 m2) Under the Conditions of a Typical 6.6-hour Clinical Study..... 6-13
Figure 6-3. Median Response (Lung Function Decrements in FEV1) Predicted by the MSS Model for 20-Year Old Individuals. Exposure to 100 ppb O_3 at Moderate Exercise (40 L/min, BSA=2 m2) Under the Conditions of a Typical 6.6-hour Clinical Study. 6-14
Figure 6-4. Median Response (FEV1 Decrements) Predicted by the MSS Threshold and Non-Threshold Models for 20-Year Old Individuals, Constant 100 ppb O_3 Exposure, 2 Hours Heavy Exercise (30 L/min-m2 BSA)..... 6-15
Figure 6-5. Probability of Response ≥ 10% Predicted by the MSS Threshold and Non-Threshold Models for 20-Year Old Individuals, Constant 100 ppb O_3 Exposure, 2 Hours Heavy Exercise (30 L/min-m2 BSA)..... 6-15
Figure 6-6. Probabilistic Exposure-Response Relationships for FEV1 Decrements ≥ 10% for 8-Hour Exposures At Moderate Exertion, Ages 18-35. Values associated with data points are the number of subject-exposures at each exposure concentration 6-20
Figure 6-7. Risk results for all school-aged children with ≥ 1 occurrences of FEV1 decrements ≥ 10, 15, 20% for all cities, year, and scenarios (y-axis is percent of children affected). 6-23
Figure 6-8. Risk results for all school-aged children with ≥ 1 occurrences of FEV1 decrements ≥ 10% under the 0.07 ppm alternative standard showing variability across cities (horizontally) and years (vertically)..... 6-24
Figure 6-9. Distribution of Daily FEV1 Decrements ≥ 10% Across Ranges of 8-hour Average Ambient O_3 Concentrations (Los Angeles, 2006 recent air quality). 6-27
Figure 6-10. Comparison of E-R and MSS Model (restricted to 8-hour average EVR ≥ 13) Response Functions (Atlanta 2006 base case, ages 18-35)..... 6-33

Figure 6-11. Distribution of Daily Maximum 8-hour Average EVR For Values of EVR ≥ 13 (L/min-m2) (midpoints on vertical axis) (Atlanta 2006 base case, ages 18-35)..... 6-34
Figure 6-12. Sensitivity (Percent Change) of Population With One or More FEV1 Decrements ≥ 10% to a 5% Increase in Individual MSS Model Parameter Estimates. 6-40
Figure 6-13. Lung Function Risk Results, Incremental Increases In Risk For Increasing Standard Levels: Percent of All School-aged Children With FEV1 Decrement ≥ 10%, Highest Value For Each Study area Over Years 6-48
Figure 6-14. Lung Function Risk Results, Incremental Increases In Risk For Increasing Standard Levels: Percent of All School-aged Children With FEV1 Decrement ≥ 10%, Mean Value For Each Study Area Over Years 6-49
Figure 7-1. Flow Diagram of Risk Assessment for Short-term Exposure Studies 7-10
Figure 7-2. Heat Maps for Short Term O_3-attributable Mortality (Just meeting existing standard and risk reductions from just meeting alternative standards) (2007) (Smith et al., 2009 C-R functions) 7-55
Figure 7-3. Heat Maps for Short Term O_3-attributable Mortality (Just meeting existing standard and risk reductions from just meeting alternative standards) (2009) (Smith et al., 2009 C-R functions) (see Key at bottom of figure) 7-56
Figure 7-4. Plots of Short-Term O_3-attributable All-Cause Mortality for Meeting Existing standard and Alternative Standards (Smith et al., 2009) (Simulation year 2007 and 2009) 7-57
Figure 7-5. Plots of Short-Term O_3-attributable Respiratory HA for Meeting Existing standard and Alternative Standards (Medina-Ramon, et al., 2006) (Simulation year 2007 and 2009) 7-61
Figure 7-6. Plots of Long-Term O_3-attributable Respiratory Mortality for Meeting Existing standard and Alternative Standards (Jerrett et al., 2009) (Simulation year 2007 and 2009) 7-64
Figure 7-7. Sensitivity Analysis: Short-Term O_3-attributable Mortality (air quality-related factors including study area size and method used to simulate attainment of existing and alternative standard levels) (2009) SA1-smaller (Smith-based) study area, SA2-alternative method for simulating standards. 7-79

Figure 7-8. Sensitivity Analysis: Short-Term O_3-attributable Mortality (C-R function specification) (2009) SA1-regional Bayes-based adjustment; SA2-copollutant model (PM_{10}); SA3-Zanobetti and Schwartz-based effect estimates 7-80
Figure 8-1. Conceptual Diagram for National-scale Mortality Risk Assessment 8-2
Figure 8-2. Estimated annual non-accidental premature deaths (individuals) in 2007 associated with average 2006-2008 May-September average 8-hr daily maximum O_3 levels by county using Smith et al. (2009) effect estimates. 8-10
Figure 8-3. Estimated annual all-cause premature deaths (individuals) in 2007 associated with average 2006-2008 June-August average 8-hr daily mean (10am-6pm) O_3 levels by county using Zanobetti and Schwartz (2008) effect estimates. 8-10
Figure 8-4. Estimated annual adult (age 30+) respiratory premature deaths (individuals) in 2007 associated with average 2006-2008 April-September average 1-hr daily max O_3 levels by county using Jerrett et al. (2009) effect estimates 8-11
Figure 8-5. Estimated percentage of May-September total non-accidental mortality (all ages) attributable to 2006-2008 average O_3 levels by county using Smith et al. (2009) effect estimates..... 8-11
Figure 8-6. Estimated percentage of June-August total all-cause mortality (all ages) attributable to 2006-2008 average O_3 levels by county using Zanobetti and Schwartz (2008) effect estimates..... 8-12
Figure 8-7. Estimated percentage of April-September respiratory mortality among adults age 30+ attributable to 2006-2008 average O_3 levels by county using Jerrett et al. (2009) effect estimates..... 8-12
Figure 8-8. Cumulative distribution of county-level percentage of all-cause, all-year, and all-age mortality attributable to 2006-2008 average O_3 for the U.S. 8-13
Figure 8-9. Cumulative percentage of total O_3 deaths by baseline O_3 concentration. O_3 concentrations are reported as May-September average 8-hr daily maximum for results based on Smith et al. (2009) effect estimates, June-August average 8-hr mean (10am to 6pm) for results based on Zanobetti and Schwartz (2008) effect estimates, and April-September average 1-hr daily maximum for results based on Jerrett et al. (2009) effect estimates. 8-15

Figure 8-10.Regions used in the sensitivity analysis based on the Smith et al. (2009) regional-prior Bayes-shrunken city-specific and regional average effect estimates (Source: Samet et al. 2000).. 8-18
Figure 8-11. O_3-attributable premature deaths by region as calculated by applying Smith et al. (2009) regional prior Bayes-shrunken and regional average effect estimates, as compared with the national prior Bayes-shrunken and national average effect estimates as in the main results. 8-19
Figure 8-12. Comparison of county-level populations of urban case study area counties to the frequency distribution of population in 3,143 U.S. counties. 8-32
Figure 8-13. Comparison of county-level seasonal mean 8-hr daily maximum O_3 concentrations in urban case study area counties to the frequency distribution of seasonal mean 8-hr daily maximum O_3 concentrations in 671 U.S. counties with O_3 monitors..... 8-33
Figure 8-14. Comparison of 2007 county-level 4th high 8-hr daily maximum O_3 concentrations in urban case study area counties to the frequency distribution of 2007 4th high 8-hr daily maximum O_3 concentrations in 725 U.S. counties with O_3 monitors. 8-34
Figure 8-15. Comparison of county-level all-cause mortality in urban case study area counties to the frequency distribution of all-cause mortality in 3,137 U.S. counties 8-35
Figure 8-16. Comparison of county-level non-accidental mortality in urban case study area counties to the frequency distribution of non-accidental mortality in 3,135 U.S. counties 8-36
Figure 8-17. Comparison of city-level all-cause mortality risk coefficients from Zanobetti and Schwartz (2008) in urban case study areas to the frequency distribution of all-cause mortality risk coefficients from Zanobetti and Schwartz (2008) in 48 U.S. cities 8-37
Figure 8-18. Comparison of city-level national prior Bayes-shrunken non-accidental mortality risk coefficients from Smith et al. (2009) in urban case study areas to the frequency distribution of national prior Bayes-shrunken non-accidental mortality risk coefficients from Smith et al. (2009) in 98 U.S. cities 8-38

Figure 8-19. Comparison of county-level percent of population 0 to 14 years old in urban case study area counties to the frequency distribution of percent of population 0 to 14 years old in 3,141 U.S. counties 8-39
Figure 8-20. Comparison of county-level percent of population 0 to 14 years old in urban case study area counties to the frequency distribution of percent of population 0 to 14 years old in 3,141 U.S. counties 8-40
Figure 8-21. Comparison of county-level income per capita in urban case study areas to the frequency distribution of income per capita in 3,141 U.S. counties..... 8-41
Figure 8-22. Comparison of county-level July temperature in urban case study area counties to the frequency distribution of July temperature in all U.S. counties..... 8-42
Figure 8-23. Comparison of city-level asthma prevalence in urban case study areas to the frequency distribution of asthma prevalence in 184 U.S. cities 8-43
Figure 8-24. Comparison of city-level air conditioning prevalence in urban case study areas to the frequency distribution of air conditioning prevalence in 76 U.S. cities 8-44
Figure 8-25. Cumulative distribution of county-level percentage of May-September non-accidental mortality for all ages attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Smith et al. (2009) effect estimates..... 8-47
Figure 8-26. Cumulative distribution of county-level percentage of June-August all-cause mortality for all ages attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Zanobetti and Schwartz (2008) effect estimates 8-48
Figure 8-27. Cumulative distribution of county-level percentage of April-September adult (age 30+) respiratory mortality attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Jerrett et al. (2009) effect estimates and city definitions from Zanobetti and Schwartz (2008)..... 8-49
Figure 8-28. Change in 50th percentile summer season (April-October) daily 8-hr maximum O_3 concentrations between 2001-2003 and 2008-2010. 8-52
Figure 8-29. Change in 95th percentile summer season (April-October) daily 8-hr maximum O_3 concentrations between 2001-2003 and 2008-2010 8-53

Figure 8-30. Population density at each O_3 monitor 8-54
Figure 8-31. Distributions of O_3 concentrations for high population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011. 8-56
Figure 8-32. Distributions of O_3 concentrations for medium population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011 8-57
Figure 8-33. Distributions of O_3 concentrations for low population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011 8-58
Figure 8-34. Map of O_3 trends at specific monitors in the New York area. All upward and downward facing triangles represent statistically significant trends from 1998-2011 ($p < 0.05$), circles represent locations with no significant trends. Sites used in Smith et al (2009) and the Zanobetti and Schwartz (2008) epidemiology studies are represented by colored dots. Only monitors with at least seven years of data are displayed. The pink star indicates the site with the higher design value in 2011. The MSA border as defined by the U.S. census bureau is delineated by the light blue line. Left panel shows trends in annual 4th highest 8-hr daily maximum O_3 values, center panel shows trends in annual mean 8-hr daily maximum O_3 values, and right panel shows trends in annual median 8-hr daily maximum O_3 values.. 8-59
Figure 8-35. Map of O_3 trends at specific monitors in the Chicago area. All upward and downward facing triangles represent statistically significant trends from 1998-2011 ($p < 0.05$), circles represent locations with no significant trends. Sites used in Smith et al (2009) and the Zanobetti and Schwartz (2008) epidemiology studies are represented by colored dots. Only monitors with at least seven years of data are displayed. The pink star indicates the site with the higher design value in

2011. The MSA border as defined by the U.S. census bureau is delineated by the light blue line. Left panel shows trends in annual 4th highest 8-hr daily maximum O_3 values, center panel shows trends in annual mean 8-hr daily maximum O_3 values, and right panel shows trends in annual median 8-hr daily maximum O_3 values.. 8-60
Figure 8-36. Ratio of April-October seasonal average O_3 concentrations in the brute force 50% NOx emissions reduction CMAQ simulations to April-October seasonal average O_3 concentrations in the 2007 base CMAQ simulation 8-64
Figure 8-37. Ratio of April-October seasonal average O_3 concentrations in the brute force 90% NOx emissions reduction CMAQ simulations to April-October seasonal average O_3 concentrations in the 2007 base CMAQ simulation 8-65
Figure 8-38. Ratio of January monthly average O_3 concentrations in brute force 50% NOx emissions reduction CMAQ simulations to January monthly average O_3 concentrations in the 2007 base CMAQ simulation 8-66
Figure 8-39. Ratio of January monthly average O_3 concentrations in brute force 90% NOx emissions reduction CMAQ simulations to January monthly average O_3 concentrations in the 2007 base CMAQ simulation 8-67
Figure 8-40. Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent change in mean O_3 (ppb) from the 2007 base CMAQ simulation to the 50% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the case study areas. Top plots show changes in January monthly mean O_3, middle plots show changes in seasonal mean June-August O_3, and bottom plots show changes in seasonal mean April-October O_3 8-69
Figure 8-41. Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent change in mean O_3 (ppb) from the 2007 base CMAQ simulation to the 90% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1

to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the case study areas. Top plots show changes in January monthly mean O_3, middle plots show changes in seasonal mean June-August O_3, and bottom plots show changes in seasonal mean April-October O_3 8-70

Figure 8-42. Population (as % of total case-study area population) living in locations of increasing April-October seasonal mean O_3 in the 50% NOx reduction CMAQ simulation..... 8-71

Figure 8-43. Population (as % of total case-study area population) living in locations of increasing April-October seasonal mean O_3 in the 90% NOx reduction CMAQ simulation..... 8-72

Figure 8-44. Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent the change in seasonal mean (April-October) O_3 from the 2007 base CMAQ simulation to the 50% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the urban case study areas. Bottom plots show histograms for low-mid population density areas while top plots show histograms for high population density areas..... 8-74

Figure 8-45. Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent the change in seasonal mean (April-October) O_3 from the 2007 base CMAQ simulation to the 90% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the urban case study areas. Bottom plots show histograms for

low-mid population density areas while top plots show histograms for high population density areas 8-75
Figure 9-1. Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard 9-7
Figure 9-2. Effects of just meeting existing (columns 1 and 2) and alternative (columns 3 through 8) standards on percent of children (ages 5-18) with at least one O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010 9-11
Figure 9-3. Effects of just meeting existing (75 ppb) and alternative standards on percent of children (ages 5-18) exceeding 60 ppb exposure benchmark, highest value across years for each urban case study area, 2006-2010 9-13
Figure 9-4. Effects of just meeting existing (column 1) and alternative (columns 2-4) standards on percent of children (ages 5-18) with FEV1 decrement > 10, 15, and 20%, years 2006-2010 9-16
Figure 9-5. Impact of just meeting existing (75 ppb) and alternative standards on percent of children (ages 5-18) with FEV_1 decrement > 10%, highest value for each urban case study area, 2006-2010 9-18
Figure 9-6. Impacts of just meeting existing (75 ppb) and alternative standard levels on mortality risk per 100,000 population for 2007 and 2009 9-21
Figure 9-7. Impacts of just meeting existing and alternative standard levels on adult (ages 65 and older) respiratory hospital admissions risk per 100,000 population for 2007 and 2009 9-22
Figure 9-8. Comparison of Exposure (Row 1) Lung Function Risk (Row 2) and Epidemiology-Based Risk (Rows 3 and 4) Metrics after Just Meeting the Existing 75 ppb Standard 9-30
Figure 9-9. Comparison of Percent Reduction in Key Risk Metrics for Alternative Standard Levels Relative to Just Meeting the Existing 75 ppb Standard 9-33

TABLE OF TABLES

Table 3-1. Short-term O_3 Exposure Health Endpoints Evaluated in Urban Case Study Areas 3-20

Table 4-1. Monitor and Area Information for the 15 Urban Case Study Areas in the Exposure Modeling and Clinical Study Based Risk Assessment 4-6

Table 4-2. Monitor and Area Information for the 12 Urban Case Study Areas in the Epidemiology Based Risk Assessment 4-9

Table 4-3. Summary Statistics Based on the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells 4-36

Table 4-4. Correlation Coefficients Between the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells 4-36

Table 4-5. Correlation Coefficients and Ratios of the 2006-2008 O_3 Design Values to the 2006-2008 Fused Seasonal Average O_3 Levels for the CMAQ 12km Grid Cells Containing O_3 Monitors 4-38

Table 4-6. Summary of Qualitative Uncertainty Analysis of Key Air Quality Elements in the O_3 NAAQS Risk Assessment 4-42

Table 5-1. General Characteristics of the Population Exposure Modeling Domain Comprising Each Study Area 5-9

Table 5-2. Asthma Prevalence for Children and Adults Estimated by APEX in Each Simulated Study Area 5-12

Table 5-3. Consolidated Human Activity Database (CHAD) Study Information and Diary-days Used by APEX 5-14

Table 5-4. Ventilation equation coefficient estimates (b_i) and residuals distributions (e_i) 5-16

Table 5-5. Microenvironments Modeled, Calculation Method Used, and Variables Included 5-17

Table 5-6. Characterization of Key Uncertainties in Historical and Current APEX Exposure Assessments 5-58

Table 5-7. Mean and Maximum Percent of all School-age Children Estimated to Experience at Least One Daily Maximum 8-hr Average Exposure to O_3 at or Above Selected Health Benchmark Levels 5-68

Table 5-8. Mean and Maximum Percent of All School-age Children Estimated to Experience at Least Two Daily Maximum 8-hr Average Exposures to O_3 At or Above Selected Health Benchmark Levels 5-70
Table 6-1. Estimated Parameters in the MSS Models 6-11
Table 6-2. Age Term Parameters for Application of the 2012 MSS Threshold Model to All Ages 6-12
Table 6-3. Study-specific Ozone Exposure-response Data for Lung Function Decrements Based on Correcting Individual Responses for the Effect on Lung Function of Exercise in Clean Air, Ages 18-35 6-16
Table 6-4. Ranges of percents of population experiencing one or more days during the ozone season with lung function decrement (ΔFEV_1) more than 10%. The numbers in this table are the minimum and maximum percents estimated over all cities and years 6-25
Table 6-5. Ranges of percents of population experiencing one or more days during the ozone season with lung function decrement (ΔFEV_1) more than 15%. The numbers in this table are the minimum and maximum percents estimated over all cities and years. 6-26
Table 6-6. Percents of the General Population and Outdoor Workers (ages 19-35) Experiencing 1 or More and 6 or More FEV_1 Decrements $\geq$ 15% (based on Atlanta 2006 APEX simulations) 6-28
Table 6-7. Ranges of percents of school-aged children experiencing one or more days during the ozone season with lung function decrement (ΔFEV_1) more than 10 and 15%. The numbers in this table are the minimum and maximum percents estimated over all cities and years 6-29
Table 6-8. Comparison of responses from the MSS model with responses from the population exposure-response (E-R) method. 2006 existing standard, ages 5 to 18 6-30
Table 6-9. Comparison of MSS Model and E-R Model of Previous Reviews for Atlanta, Mar 1-Oct 30, 2006, ages 18-35 6-32
Table 6-10. Comparison of MSS Model and E-R Model of Previous Reviews for Los Angeles, Jan 1-Dec 31, 2006, ages 18-35 6-32

Table 6-11. Comparison of Responses from the MSS 2010 Model with Responses from McDonnell et al. (1985) 6-36
Table 6-12. Percents of the population by age group with one or more days during the ozone season with lung function (FEV_1) decrements more than 10, 15, and 20% (Atlanta 2006 base case). MSS Threshold model, monitors air quality 6-37
Table 6-13. Percents of the population by age group with one or more days during the ozone season with lung function (FEV_1) decrements more than 10, 15, and 20% (Atlanta 2006 base case). MSS No-Threshold model, monitors air quality 6-37
Table 6-14. MSS threshold model estimated parameters with confidence intervals 6-39
Table 6-15. Convergence results for the Atlanta 2006 base case with 200,000 simulated individuals. Percents of the population by age group with one or more days (and six or more days) during the ozone season with lung function (FEV_1) decrements more than 10, 15, and 20%. Minimum and maximum values and ranges over 40 APEX runs 6-42
Table 6-16. Summary of Qualitative Uncertainties of Key Modeling Elements in the O_3 Lung Function Risk Assessment 6-44
Table 7-1. Information on the 12 Urban Case Study Areas in the Risk Assessment 7-14
Table 7-2. Overview of Epidemiological Studies Used in Specifying C-R Functions 7-23
Table 7-3. CBSA-based Study Areas with Multiple Effect Estimates from the Smith et al., 2009 Study 7-27
Table 7-4. Summary of Qualitative Uncertainty Analysis of Key Modeling Elements in the O_3 NAAQS Risk Assessment 7-41
Table 7-5. Specification of the Core and Sensitivity Analyses (air quality simulation) 7-46
Table 7-6. Specification of the Core and Sensitivity Analyses (alternative C-R function specification) 7-47
Table 7-7. Short-Term O_3-attributable All Cause Mortality Incidence (2007 and 2009) (Smith et al., 2009 C-R Functions) 7-53
Table 7-8. Percent of Total All-Cause Mortality Attributable to O_3 and Percent Change in O_3-Attributable Risk (2007 and 2009) (Smith et al., 2009 C-R functions) 7-54

Table 7-9. Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in O_3-attributable Risk – Respiratory-Related Hospital Admissions (2007 and 2009) 7-58
Table 7-10. Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in Ozone-attributable Risk – Emergency Room Visits (2007 and 2009) 7-59
Table 7-11. Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in Ozone-attributable Risk – Asthma Exacerbations (2007 and 2009) 7-60
Table 7-12. Long-Term O_3-attributable Respiratory Mortality Incidence (2007 and 2009) (Jerrett et al., 2009 C-R Functions) 7-62
Table 7-13. Long-Term O_3-attributable Respiratory Mortality Percent of Baseline Incidence and Percent Reduction in O_3-attributable Risk (simulation years 2007 and 2009) (Jerrett et al., 2009 C-R Functions) 7-63
Table 7-14. Sensitivity Analysis for Long-Term O_3-attributable Respiratory Mortality – Alternative C-R Function Specification (regional effect estimates) % of baseline all-cause mortality and change in O_3-attribuable risk (2009) (Smith et al., 2009, O_3 season) 7-81
Table 7-15. Sensitivity Analysis for Long-Term O_3-attributable Respiratory Mortality – Alternative C-R Function Specification (national O_3-only effect estimates) % of baseline all-cause mortality and change in O_3-attribuable risk (2009) (Smith et al., 2009, O_3 season) 7-82
Table 8-1. Estimated annual O_3-related premature mortality in 2007 associated with 2006-2008 average O_3 concentrations (95th percentile confidence interval) 8-7
Table 8-2. Mean, median, 2.5 percentile, and 97.5 percentile of the estimated percentage of mortality attributable to ambient O_3 for all 3087 counties in the continental U.S.... 8-9
Table 8-3. Sensitivity of estimated O_3-attributable premature deaths to the application of the 5th lowest and 5th highest city-specific risk estimates found by Smith et al. (2009) and Zanobetti and Schwartz (2008) to the gridcells between the cities included in those studies 8-16
Table 8-4. Sensitivity of estimated O_3-attributable premature deaths to the application of Smith et al. (2009) regional prior Bayes-shrunken city-specific and regional average effect

estimates, as compared with the national prior Bayes-shrunken city-specific and national average effect estimates as in the main results 8-18
Table 8-5. Data Sources for O_3 risk-related Attributes 8-25
Table 8-6. Summary Statistics for Selected O_3 Risk-related Attributes 8-29
Table 8-7. Broad Regional Annual Trends of Concurrent O_3 Concentrations and Emissions of NOx and VOCs over the 2000-2011 Time Period 8-61
Table 9-1. Area and Monitoring Information for the 15 Case Study Areas 9-3
Table 9-2. General Patterns in Seasonal (May-Sept) Mean of Daily Maximum 8-hour O_3 Concentrations after Adjusting to Meet Existing and Alternative Standards 9-5

LIST OF ACRONYMS/ABBREVIATIONS

AER	air exchange rate
AHRQ	Agency for Healthcare Research and Quality
APEX	Air Pollution Exposure Model
AQI	Air Quality Index
AQS	Air Quality System
ATUS	American Time Use Survey
BenMAP	Benefits Mapping and Analysis Program
BRFSS	Behavioral Risk Factor Surveillance System
BSA	body surface area
CAA	Clean Air Act
CASAC	Clean Air Science Advisory Committee
CDC	Center for Disease Control and Prevention
CDF	cumulative distribution functions
CH_4	methane
CHAD	Consolidated Human Activity Database
CI	confidence interval
CMAQ	Community Multi-scale Air Quality
CO_2	carbon dioxide
C-R	Concentration Response (function)
ED	emergency department
EGU	electric generating unit
EPA	U.S. Environmental Protection Agency
ER	emergency room
eVNA	enhanced Voronoi Neighbor Averaging
EVR	equivalent ventilation rate
FEM	Federal Equivalent Method
FEV1	one-second forced expiratory volume
FRM	Federal Reference Method

FVC	forced vital capacity
HA	hospital admissions
HDDM	Higher-order Decoupled Direct Method
HNO_3	nitric acid
HO_2	hydro-peroxy radical
HUCP	Healthcare Cost and Utilization Program
IPCC	Intergovernmental Panel on Climate Change
IRP	Integrated Review Plan
ISA	Integrated Science Assessment
LML	lowest measured level
MATS	Modeled Attainment Test Software
METs	metabolic equivalents of work
MSA	Metropolitan Statistical Area
MT	metric ton
NAAQS	National Ambient Air Quality Standards
NCDC	National Climatic Data Center
NEI	National Emissions Inventory
NO	nitric oxide
NO_2	nitrite
NO_x	nitrogen oxides
O_3	Ozone
OAQPS	Office of Air Quality Planning and Standards
OH	hydroxyl radical
PA	Policy Assessment
PDI	pain on deep inspiration
PI	posterior interval
PM	particulate matter
ppb	parts per billion
ppm	parts per million
PRB	Policy Relevant Background
REA	Risk and Exposure Assessment

RR	relative risk
SAB	Science Advisory Board
SEDD	State Emergency Department Databases
SES	socioeconomic status
SID	State Inpatient Databases
SO2	sulfur dioxide
STE	stratosphere-troposphere exchange
TRIM Expo	Total Risk Integrated Methodology Inhalation Exposure
VE	ventilation rate
VNA	Voronoi Neighbor Averaging
VOC	volatile organic carbon
WHO	World Health Organization

1 INTRODUCTION

The U.S. Environmental Protection Agency (EPA) is presently conducting a review of the national ambient air quality standards (NAAQS) for ozone (O_3), and related photochemical oxidants. The NAAQS review process includes four key phases: planning, science assessment, risk/exposure assessment, and policy assessment/rulemaking.[1] This process and the overall plan for this review of the O_3 NAAQS is presented in the *Integrated Review Plan for the Ozone National Ambient Air Quality Standards* (IRP, U.S. EPA, 2011a). The IRP additionally presents the schedule for the review; identifies key policy-relevant issues; and discusses the key scientific, technical, and policy documents. These documents include an Integrated Science Assessment (ISA), Risk and Exposure Assessments (REAs), and a Policy Assessment (PA). This draft Health REA is one of the two quantitative REAs developed for the review by the EPA's Office of Air Quality Planning and Standards (OAQPS); the second is a Welfare REA. This draft Health REA focuses on assessments to inform consideration of the review of the primary (health-based) NAAQS for O_3.

The existing primary (health-based) NAAQS for O_3 is set at a level of 75 ppb (0.075 ppm), based on the annual fourth-highest daily maximum 8-hour average concentration, averaged over three years, and the secondary standard is identical to the primary standard (73 FR 16436). The EPA initiated the current review of the O_3 NAAQS on September 29, 2008, with an announcement of the development of an O_3 ISA and a public workshop to discuss policy-relevant science to inform EPA's integrated plan for the review of the O_3 NAAQS (73 FR 56581). Discussions at the workshop, held on October 29-30, 2008, informed identification of key policy issues and questions to frame the review of the O_3 NAAQS. Drawing from the workshop discussions, the EPA developed a draft and then final IRP (U.S. EPA, 2011).[2] In early 2013, the EPA completed the *Integrated Science Assessment for Ozone and Related Photochemical Oxidants* (U.S. EPA, 2013). The ISA provides a concise review, synthesis and evaluation of the most policy-relevant science to serve as a scientific foundation for the review of the NAAQS. The scientific and technical information in the ISA, including that newly

[1] For more information on the NAAQS review process see http://www.epa.gov/ttn/naaqs/review.html.

[2] On March 30, 2009, EPA held a public consultation with the CASAC Ozone Panel on the draft IRP. The final IRP took into consideration comments received from CASAC and the public on the draft plan as well as input from senior Agency managers.

available since the previous review on the health effects of O_3 includes information on exposure, physiological mechanisms by which O_3 might adversely impact human health, an evaluation of the toxicological and controlled human exposure study evidence, and an evaluation of the epidemiological evidence, including information on reported concentration-response (C-R) relationships for O_3-related morbidity and mortality associations, and also includes information on potentially at-risk populations and life-stages.[3]

This REA is a concise presentation of the conceptual model, scope, methods, key results, observations, and related uncertainties associated with the quantitative analyses performed. This REA builds upon the health effects evidence presented and assessed in the ISA, as well as CASAC advice (Samet, 2011), and public comments on a scope and methods planning document for the REA (here after, "Scope and Methods Plan," U.S. EPA, 2011). Preparation of this second draft REA draws upon the final ISA and reflects consideration of CASAC and public comments on the first draft REA (Frey and Samet, 2012). This second draft health REA is being released, concurrently with the second draft welfare REA and second draft PA for review by the CASAC O_3 Panel at a public meeting scheduled for March 25-27, 2014, and for public comment.

The second draft PA presents a staff evaluation and preliminary staff conclusions of the policy implications of the key scientific and technical information in the ISA, and second draft REAs. When final, the PA is intended to help "bridge the gap" between the Agency's scientific assessments presented in the ISA and REAs, and the judgments required of the EPA Administrator in determining whether it is appropriate to retain or revise the NAAQS. The PA integrates and interprets the information from the ISA and REAs to frame policy options for consideration by the Administrator. In so doing, the PA recognizes that the selection of a specific approach to reaching final decisions on primary and secondary NAAQS will reflect the judgments of the Administrator. The development of the various scientific, technical and policy documents and their roles in informing this NAAQS review are described in more detail in the second draft PA.

[3] The ISA also evaluates scientific evidence for the effects of O_3 on public welfare which EPA will consider in its review of the secondary O_3 NAAQS. Building upon the effects evidence presented in the ISA, OAQPS has also developed a second draft of a second REA titled *Ozone Welfare Effects Risk and Exposure Assessment (U.S. EPA, 2013).*

1.1 HISTORY

As part of the last O_3 NAAQS review completed in 2008, EPA's OAQPS conducted quantitative risk and exposure assessments to estimate exposures above health benchmarks and risks of various health effects associated with exposure to ambient O_3 in a number of urban study areas, selected to illustrate the public health impacts of this pollutant (U.S. EPA 2007a, U.S. EPA, 2007b). The assessment scope and methodology were developed with considerable input from CASAC and the public, with CASAC generally concluding that the exposure assessment reflected generally accepted modeling approaches, and that the risk assessments were well done, balanced and reasonably communicated (Henderson, 2006a). The final quantitative risk and exposure assessments took into consideration CASAC advice (Henderson, 2006a; Henderson, 2006b), and public comments on two drafts of the risk and exposure assessments.

The exposure and health risk assessment conducted in the last review developed exposure and health risk estimates for 12 urban areas across the U.S., based on 2002 to 2004 air quality data. That assessment provided annual or O_3 season-specific exposure and risk estimates for these years of air quality and for air quality scenarios, simulating just meeting the then-existing 8-hour O_3 standard set in 1997 at a level of 0.08 ppm and several alternative 8-hour standards. The strengths and limitations in the assessment were characterized, and analyses of key uncertainties were presented.

Exposure estimates from the last assessment were used as an input to the risk assessment for lung function responses (a health endpoint for which exposure-response functions were available from controlled human exposure studies). Exposure estimates were developed for the general population and population groups including school age children with asthma as well as all school age children. The exposure estimates also provided information on exposures to ambient O_3 concentrations at and above specified benchmark levels (referred to as "exposures of concern"), to provide some perspective on the public health impacts of health effects associated with O_3 exposures in controlled human exposure studies that could not be evaluated in the quantitative risk assessment (e.g., lung inflammation, increased airway responsiveness, and decreased resistance to infection). For several other health endpoints, O_3-related risk estimates were generated using concentration-response relationships reported in epidemiological or field studies, together with ambient air quality concentrations, baseline health incidence rates, and population data for the various locations included in the assessment. Health endpoints included

in the assessment based on epidemiological or field studies included: hospital admissions for respiratory illness in four urban areas, premature mortality in 12 urban areas, and respiratory symptoms in asthmatic children in 1 urban area.

Based on the 2006 *Air Quality Criteria for Ozone* (U.S. EPA, 2006), the Staff Paper (U.S. EPA, 2007), and related technical support documents (including the REAs), the proposed decision was published in the Federal Register on July 11, 2007 (72 FR 37818). The EPA proposed to revise the level of the primary standard to a level within the range of 0.075 to 0.070 ppm. Two options were proposed for the secondary standard: (1) replacing the current standard with a cumulative seasonal standard, expressed as an index of the annual sum of weighted hourly concentrations cumulated over 12 daylight hours during the consecutive 3-month period within the O_3 season with the maximum index value (W126), set at a level within the range of 7 to 21 ppm-hours, and (2) setting the secondary standard identical to the revised primary standard. The EPA completed the review with publication of a final decision on March 27, 2008 (73 FR 16436), revising the level of the 8-hour primary O_3 standard from 0.08 ppm to 0.075 ppm, as the 3-year average of the fourth highest daily maximum 8-hour average concentration, and revising the secondary standard to be identical to the revised primary standard.

Following promulgation of the revised O_3 standard in March 2008, state, public health, environmental, and industry petitioners filed suit against EPA regarding that final decision. At EPA's request, the consolidated cases were held in abeyance pending EPA's reconsideration of the 2008 decision. A notice of proposed rulemaking to reconsider the 2008 final decision was issued by the Administrator on January 6, 2010. Three public hearings were held. The Agency solicited CASAC review of the proposed rule on January 25, 2010, and additional CASAC advice on January 26, 2011. On September 2, 2011, the Office of Management and Budget returned the draft final rule on reconsideration to EPA for further consideration. EPA decided to coordinate further proceedings on its voluntary rulemaking on reconsideration with this ongoing periodic review, by deferring the completion of its voluntary rulemaking on reconsideration until it completes its statutorily-required periodic review. In light of that, the litigation on the 2008 final decision proceeded. On July 23, 2013, the Court ruled on the litigation of the 2008 decision, denying the petitioners suit except with respect to the secondary standard, which was remanded to the Agency for reconsideration. The second draft PA provides additional description of the court

ruling with regard to the secondary standard.

1.2 CURRENT RISK AND EXPOSURE ASSESSMENT: GOALS AND PLANNED APPROACH

The goals of the current quantitative exposure and health risk assessments are to provide information relevant to answering questions regarding the adequacy of the existing O_3 standard and the potential improvements in public health from meeting alternative standards. To meet these goals, this assessment provides results from several analyses, including (1) estimates of the number of people in the general population and in at-risk populations and lifestages with O_3 exposures above benchmark levels, while at moderate or greater exertion levels; (2) estimates of the number of people in the general population and in at-risk populations and lifestages with impaired lung function resulting from exposures to O_3; and (3) estimates of the potential magnitude of premature mortality and selected morbidity health effects in the population, including at-risk populations and lifestages, where data are available to assess these groups. For each of the analyses, we provide estimates for recent ambient levels of O_3 and for air quality conditions simulated to just meet the existing O_3 standard and alternative standards.

In presenting these results, we evaluate the influence of various inputs and assumptions on the exposure and risk estimates to more clearly differentiate alternative standards that might be considered, including potential impacts on various at-risk populations and lifestages. We also evaluate the distribution of risks and patterns of risk reduction and uncertainties in those risk estimates. In addition, we have conducted an assessment to provide nationwide estimates of the potential magnitude of premature mortality associated with recent ambient O_3 concentrations, to more broadly characterize this risk on a national scale. This assessment includes an evaluation of the distribution of risk across the U.S., to assess the extent to which we have captured the upper end of the risk distribution with our urban study area analyses.

This current quantitative risk and exposure assessment builds on the approach used and lessons learned in the last O_3 risk and exposure assessment, and focuses on improving the characterization of the overall confidence in the exposure and risk estimates, including related uncertainties, by incorporating a number of enhancements, in terms of both the methods and data used in the analyses. This risk assessment considers a variety of health endpoints for which, in staff's judgment, there is adequate information to develop quantitative risk estimates that can meaningfully inform the review of the primary O_3 NAAQS.

1.3 ORGANIZATION OF DOCUMENT

The remainder of this document is organized as follows. Chapter 2 provides a conceptual framework for the risk and exposure assessment, including discussions of O_3 chemistry, sources of O_3 precursors, exposure pathways and microenvironments where O_3 exposure can be high, at-risk populations and lifestages, and health endpoints associated with O_3. This conceptual framework sets the stage for the scope of the risk and exposure assessments. Chapter 3 provides an overview of the scope of the quantitative risk and exposure assessments, including a summary of the previous risk and exposure assessments, and an overview of the current risk and exposure assessments. Chapter 4 discusses air quality considerations relevant to the exposure and risk assessments, including available O_3 monitoring data, and important inputs to the risk and exposure assessments. Chapter 5 describes the inputs, models, and results for the human exposure assessment, and discusses the literature on exposure to O_3, exposure modeling approaches using the Air Pollution Exposure Model (APEX), the scope of the exposure assessment, inputs to the exposure modeling, sensitivity and uncertainty evaluations, and estimation of results. Chapter 6 describes the estimation of health risks based on application of the results of controlled human exposure studies, including discussions of health endpoint selection, approaches to calculating risk, and results. Chapter 7 describes the estimation of health risks in selected urban areas based on application of the results of observational epidemiology studies, including discussions of air quality characterizations, model inputs, variability and uncertainty, and results. Chapter 8 describes the national scale risk characterization and urban area representativeness analysis. Chapter 9 provides an integrative discussion of the exposure and risk estimates generated in the analyses drawing on the results of the analyses based on both clinical and epidemiology studies, and incorporating considerations from the national scale risk characterization.

2 OVERVIEW OF EXPOSURE AND RISK ASSESSMENT DESIGN

In this chapter, we summarize our framework for assessing exposures to O_3 and the associated risks to human populations. Figure 2-1 provides an overview of the general design of this exposure and risk assessment, which includes air quality characterization, review of relevant scientific evidence on health effects, modeling of exposure, modeling of risk, and risk characterization. Each element identified in the diagram is described in a specific, identified chapter of this exposure and risk assessment.

In this O_3 exposure and risk assessment, modeling of personal exposure and estimation of risks which rely on personal exposure estimates, are implemented using the Air Pollution Exposure model (APEX)[1] (U.S. EPA, 2012 a, b). Modeling of population level risks for endpoints based on application of results of epidemiological studies, is implemented using the environmental Benefits Mapping and Analysis Program (BenMAP),[2] a peer reviewed software tool for estimating risks and impacts associated with changes in ambient air quality (U.S. EPA, 2013). The overall characterization of risk draws from the results of the exposure assessment and both types of risk assessment.

The remainder of this chapter includes summary discussions of each of the main elements of Figure 2-1, including policy-relevant exposure and risk questions (Section 2.1), characterization of ambient O_3, including important sources of O_3 precursors, and its relation to population exposures, as well as simulation of just meeting existing and potential alternative O_3 standards (Section 2.2), review of health evidence identified in the literature describing associations with ambient O_3 (Section 2.3), key components of exposure modeling (Section 2.4), key components of risk modeling (Section 2.5), and risk characterization (Section 2.6).

Specific details related to the scope of the exposure and risk assessments and how each element will be addressed in the quantitative exposure and risk analysis are provided in Chapter 3.

[1] APEX is available for download at http://www.epa.gov/ttn/fera/human_apex.html
[2] BenMAP is available for download at http://www.epa.gov/air/benmap/

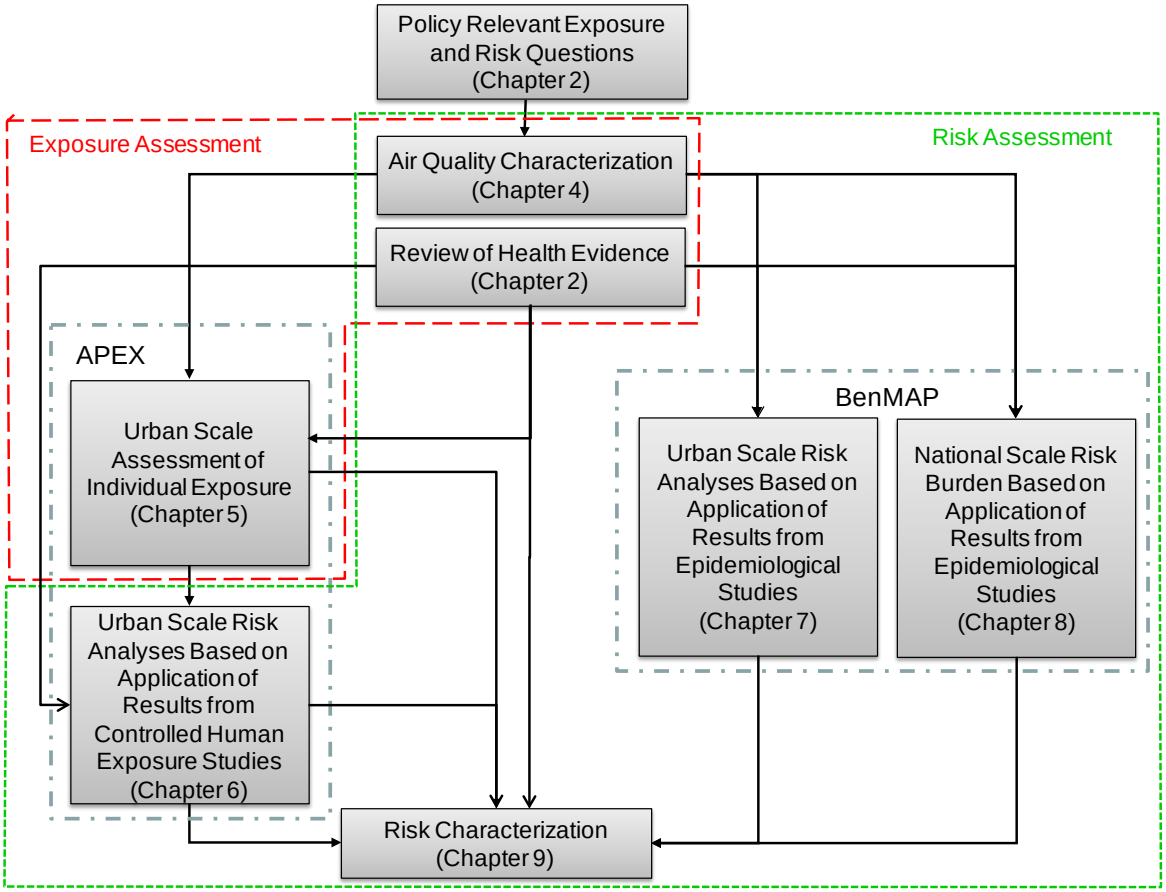


Figure 2-1 Overview of Exposure and Risk Assessment Design

2.1 POLICY-RELEVANT EXPOSURE AND RISK QUESTIONS

The first step in the design is to determine the set of policy-relevant exposure and risk questions that will be informed by the assessment. Consistent with recommendations from the recent National Academy of Sciences report "Science and Decisions: Advancing Risk Assessment" (NAS, 2009), these exposure and risk assessments have been designed to address the risk questions identified in the Integrated Review Plan for the Ozone National Ambient Air Quality Standards (U.S. EPA, 2011). We have focused on designing the exposure and risk assessments to inform consideration of those risk-related policy-relevant questions in the separately developed O_3 NAAQS Policy Assessment. The risk-related policy-relevant questions identified in the Integrated Review Plan are related to two main activities, evaluation of the adequacy of the existing standards and, if appropriate, evaluation of potential alternative standards (U.S. EPA, 2011). With regard to evaluation of the adequacy of the existing standards, the risk-related policy-relevant questions are:

"To what extent do risk and/or exposure analyses suggest that exposures of concern for O_3-related health effects are likely to occur with existing ambient levels of O_3 or with levels that just meet the O_3 standard? Are these risks/exposures of sufficient magnitude such that the health effects might reasonably be judged to be important from a public health perspective? What are the important uncertainties associated with these risk/exposure estimates?"

With regards to evaluation of potential alternative standards, the risk-related policy-relevant questions are:

"To what extent do alternative standards, taking together levels, averaging times and forms, reduce estimated exposures and risks of concern attributable to O3 and other photochemical oxidants, and what are the uncertainties associated with the estimated exposure and risk reductions? What conclusions can be drawn regarding the health protection afforded at-risk populations?"

This risk and exposure assessment is designed to inform consideration of these questions through application of exposure and risk modeling for a set of urban case study areas. Exposure and risk estimates will be generated for recent O3 concentrations, O3 concentrations after simulating just meeting the existing standards, and O3 concentrations after simulating just meeting potential alternative standards. Careful consideration will be given to addressing variability and uncertainty in the estimates, and to the degree to which at-risk populations experience exposures and risks. Exposure modeling is discussed in Chapter 5 (Urban-Scale Assessment of Individual Exposure), while risk modeling is discussed in Chapter 6 (Characterization of Health Risks Based on Clinical Studies) and Chapter 7 (Characterization of Health Risks Based on Epidemiological Studies). Chapter 8 (National-Scale Risk Assessment and Representativeness Analysis) provides a national-scale assessment of risks under recent O3 concentrations to provide context for the urban-scale analyses and to help characterize the representativeness of the urban-scale analyses.

In order to inform consideration of the risk-related policy-relevant questions, the first step for all of the exposure and risk analyses is simulation of meeting the existing and alternative standards. To do this, recent air quality measurements of O_3 are adjusted such that they mimic a realistic and general atmospheric response to changes in precursor emissions for the specific urban area and so that they just meet the existing and alternative standard levels. Conceptually, there is an almost infinite set of combinations of precursor emissions reductions that will result in just meeting the existing or alternative standards. The specific combinations of reductions that might actually be implemented are not relevant for the exposure and risk analyses, as those will result from the implementation processes which follow the establishment of a standard.

However, it is appropriate to ask the question of how the patterns of ambient O_3 on multiple temporal scales (hourly, daily, monthly, seasonally) and across each urban area, may respond to precursor emissions reductions that result in meeting the existing and potential alternative standards, and how these different patterns of O_3 could affect the exposure and risk results. The answers to these questions are critical inputs to the exposure and risk analyses. Consideration of the available methods for simulating just meeting existing and alternative standards is discussed in Chapter 4 (Air Quality Characterization).

Analyses presented in this document to inform the policy-relevant risk questions regarding potential alternative standards, are focused on alternative levels for an 8-hour standard. Other elements of the standard (indicator, averaging time, and form),[3] are addressed in the Policy Assessment as part of the overall evaluation of the health protection afforded by the primary O_3 standards.

With regard to potential alternative levels for an 8-hour O_3 standard, the quantitative risk assessment evaluates the range of levels in 5 ppb increments from 60 to 70 ppb. These levels were selected based on the evaluations of the evidence provided in the first draft PA, which received support from the CASAC in their advisory letter on the first draft PA (Frey and Samet, 2012). For a subset of urban areas, we also evaluated a standard level of 55 ppb, consistent with recommendations from CASAC to also give consideration to evaluating a level somewhat below 60 ppb. Thus, for most areas, we evaluate exposures and risks for potential alternative standard levels of 70, 65, and 60 ppb. Some additional analyses were also included for evaluation of exposures and risks for a potential alternative 8-hour standard level of 55 ppb.

2.2 AIR QUALITY CHARACTERIZATION

In order to address the policy-relevant questions discussed in Section 2.1, the first step is characterizing O_3 concentrations relevant to estimation of exposure and risk. This requires characterization of recent O_3 concentrations, O_3 concentrations after simulating just meeting the existing standards, and O_3 concentrations after simulating just meeting potential alternative standards. This section provides conceptual information on O_3 formation and responsiveness of O_3 to changes in precursor emissions, that inform the simulations of just meeting existing and alternative standards.

[3] The "form" of a standard defines the air quality statistic that is compared to the level of the standard in determining whether an area attains the standard. The existing form of the 8-hour O_3 standard is the 4th highest daily maximum 8-hour average O_3, averaged over 3 years. The "indicator" of a standard defines the chemical species or mixture that is to be measured in determining whether an area attains the standard.

2.2.1 O_3 chemistry and response to changes in precursor emissions

O_3 occurs naturally in the stratosphere where it provides protection against harmful solar ultraviolet radiation, and it is formed closer to the surface in the troposphere from precursor emissions from both natural and anthropogenic sources. O_3 is created when its two primary precursors, volatile organic compounds (VOC) and oxides of nitrogen (NO_x), combine in the presence of sunlight. VOC and NO_x are, for the most part, emitted directly into the atmosphere. Carbon monoxide (CO) and methane (CH_4) can also be important for O_3 formation (U.S. EPA, 2013, section 3.2.2).

Rather than varying directly with emissions of its precursors, O_3 changes in a nonlinear fashion with the concentrations of its precursors. NO_x emissions lead to both the formation and destruction of O_3, depending on the local concentrations of NO_x, VOC, and radicals such as the hydroxyl (OH) and hydroperoxy (HO_2) radicals. In areas dominated by fresh emissions of NO_x, these radicals are removed via the production of nitric acid (HNO_3), which lowers the O_3 formation rate. In addition, the depletion of O_3 by reaction with NO is called "titration" and is often found in downtown metropolitan areas, especially near busy streets and roads, and in power plant plumes. This "titration" results in O_3 concentrations that can be much lower than in surrounding areas. Titration is usually confined to areas close to strong NO_x sources, and the NO_2 formed can lead to O_3 formation later and further downwind. Consequently, O_3 response to reductions in NO_x emissions is complex and may include O_3 decreases at some times and locations and increases of O_3 in other times and locations. In areas with low NO_x concentrations, such as those found in remote continental areas and rural and suburban areas downwind of urban centers, the net production of O_3 typically varies directly with NO_x concentrations, and increases with increasing NO_x emissions.

In general, the rate of O_3 production is limited by either the concentration of VOCs or NO_x, and O_3 formation, using these two precursors relies on the relative sources of OH and NO_x. When OH radicals are abundant and are not depleted by reaction with NO_x and/or other species, O_3 production is referred to as being "NO_x-limited" (U.S. EPA, 2013, section 3.2.4). In this situation, O_3 concentrations are most effectively reduced by lowering NO_x emissions, rather than lowering emissions of VOCs. When the abundance of OH and other radicals is limited either through low production or reactions with NO_x and other species, O_3 production is sometimes called "VOC-limited" or "radical limited" or "NO_x-saturated" (Jaegle et al., 2001), and O_3 is most effectively reduced by lowering VOCs. However, even in NO_x-saturated conditions, very large decreases in NO_x emissions can cause the O_3 formation regime to become NO_x-limited. Consequently, reductions in NO_x emissions (when large), can make further emissions reductions more effective at reducing O_3. Between the NO_x-limited and NO_x-saturated extremes there is a transitional region, where O_3 is less sensitive to marginal changes in either NO_x or VOCs. In

rural areas and downwind of urban areas, O_3 production is generally NO_x-limited. However, across urban areas with high populations, conditions may vary. For contrast, while data from monitors in Nashville, TN, suggest NO_x-limited conditions exist there, data from monitors in Los Angeles suggest NO_x-saturated conditions (U.S. EPA, 2013, Figure 3-3).

2.2.2 Sources of O_3 and O_3 Precursors

O_3 precursor emissions can be divided into anthropogenic and natural source categories, with natural sources further divided into biogenic emissions (from vegetation, microbes, and animals), and abiotic emissions (from biomass burning, lightning, and geogenic sources). The anthropogenic precursors of O_3 originate from a wide variety of stationary and mobile sources.

In urban areas, both biogenic and anthropogenic VOCs, as well as CO, are important for O_3 formation. Hundreds of VOCs are emitted by evaporation and combustion processes from a large number of anthropogenic sources. Based on the 2005 national emissions inventory (NEI), solvent use and highway vehicles are the two main anthropogenic sources of VOCs, with roughly equal contributions to total emissions (U.S. EPA, 2013, Figure 3-2). The emissions inventory categories of "miscellaneous" (which includes agriculture and forestry, wildfires, prescribed burns, and structural fires), and off-highway mobile sources are the next two largest contributing emissions categories with a combined total of over 5.5 million metric tons a year (MT/year).

On the U.S. and global scales, emissions of VOCs from vegetation are much larger than those from anthropogenic sources. Emissions of VOCs from anthropogenic sources in the 2005 NEI were ~17 MT/year (wildfires constitute ~1/6 of that total), compared to emissions from biogenic sources of 29 MT/year. Vegetation emits substantial quantities of VOCs, such as isoprene and other terpenoid and sesqui-terpenoid compounds. Most biogenic emissions occur during the summer because of their dependence on temperature and incident sunlight. Biogenic emissions are also higher in southern and eastern states than in northern and western states for these reasons and because of species variations.

Anthropogenic NO_x emissions are associated with combustion processes. Based on the 2005 NEI, the three largest sources of NO_x are on-road and off-road mobile sources (e.g., construction and agricultural equipment), and electric power generation plants (EGUs) (U.S. EPA, 2013, Figure 3-2). Emissions of NO_x therefore are highest in areas having a high density of power plants and in urban areas having high traffic density. However, it is not possible to make an overall statement about their relative impacts on O_3 in all local areas because EGUs are sparser than mobile sources, particularly in the west and south and because of the nonlinear nature of O_3 chemistry discussed in Section 2.2.1.

Major natural sources of NO_x in the U.S. include lightning, soils, and wildfires. Biogenic NO_x emissions are generally highest during the summer and occur across the entire country, including areas where anthropogenic emissions are low. It should be noted that uncertainties in estimating natural NO_x emissions are much larger than for anthropogenic NO_x emissions.

O_3 concentrations in a region are maintained by a balance between photochemical production and transport of O_3 into the region; and loss of O_3 by chemical reactions, deposition to the surface and transport out of the region. O_3 transport occurs on many spatial scales including local transport between cities, regional transport over large regions of the U.S. and international/long-range transport. In addition, O_3 is also transfered into the troposphere from the stratosphere, which is rich in O_3, through stratosphere-troposphere exchange (STE). STE occurs in tropopause "foldings" that occur behind cold fronts, bringing stratospheric air with them (U.S. EPA, 2013, section 3.4.1.1). Contributions to O_3 concentrations in an area from STE are defined as being part of background O_3 (U.S. EPA, 2013, section 3.4).

2.2.3 Simulation of Meeting Existing and Alternative Standards

Conceptually, simulation of meeting existing and alternative standards should reflect the physical and chemical processes of O_3 formation in the atmosphere and estimate how hourly values of O_3 at each monitor in an urban area would change in response to reductions in precursor emissions, allowing for nonlinearities in response to emissions reductions and allowing for nonlinear interactions between reductions in NO_x and VOC emissions. For this assessment, we have employed sophisticated air quality models to conduct simulations of hourly O_3 responses to reductions in precursor emissions. This modeling incorporates all known emissions, including emissions from both natural and anthropogenic sources within and outside of the U.S. By using the model-adjustment methodology we are able to more realistically simulate the temporal and spatial patterns of O_3 response to precursor emissions. We chose to simulate just meeting the existing and alternative standards, by applying equal proportional decreases in U.S. anthropogenic emissions of NOx and VOC, in order to avoid any suggestion that we are approximating a specific emissions control strategy that a state or urban area might adopt to meet a standard. These analyses allow us to apply an adjustment to ambient O_3 measurements in the urban case study areas, to better represent how air quality concentrations at each monitor would change to meet the existing and alternative standard levels. The details of the specific approach used to simulating attainment for the existing and alternative standards, are discussed in greater detail in Chapter 4 and in the Chapter 4 appendices.

It is fundamentally a policy decision, as to which sources of precursor emissions are most appropriate to decrease to simulate just meeting existing and alternative O_3 standards. In addressing the policy-relevant questions regarding the evaluation of alternative standards,

consistent with previous reviews of the O_3 standards, this analysis is focused on simulating reductions in risk associated with precursor emissions originating from anthropogenic sources within the U.S. In doing so, we recognize that the CAA provides mechanisms primarily for reducing emissions from U.S. emissions sources. As such, we estimate changes in exposure and risks likely to result from just meeting alternative standards relative to just meeting the existing standards, by simulating changes in atmospheric concentrations that represent atmospheric response to reductions in U.S. anthropogenic emissions. However, we recognize that, in this approach, we are simulating attainment of existing and alternative standard levels, based on recent air quality concentrations and the chemical environment and emissions in those years. We have not mimicked the future-year atmospheric conditions and emissions inventory as would be done for the implementation process.

In addition, while it is possible to decrease O_3 concentrations using decreases in either NOx or VOC or both NO_x and VOC, the specific combination of the reductions in those emissions is a policy decision, with recognition that atmospheric chemistry considerations will make NO_x and VOC decreases more or less effective in specific urban areas, depending on the degree to which O_3 formation is NO_x or VOC limited. As discussed above, in most locations, decreases in NO_x are the most effective means to decrease ambient O_3 concentrations. However, in some downtown urban areas, O_3 formation is VOC-limited, and therefore smaller decreases in NO_x will not decrease O_3.

2.2.4 Consideration of Health Evidence

A critical input for both the exposure and risk assessments is the health evidence summarized in the Integrated Science Assessment (ISA) (U.S. EPA, 2013). This health evidence provides the basis for evaluating the significance of exposures to O_3, by informing health benchmarks for estimating exposures of concern. The evidence also provides the basis for selecting health endpoints that will be modeled in the risk assessment. This evidence includes controlled human exposure studies and observational epidemiology studies. The health evidence is also the source of the specific studies that are used to develop exposure-response (E-R) and concentration-response (C-R) functions, used in the risk assessment. Finally, the health evidence provides information on at-risk populations to guide the selections of study populations used in the exposure and risk assessments. The following subsections summarize key conceptual aspects regarding exposures of concern, health endpoints, E-R and C-R functions, and at-risk populations.

2.2.5 Exposures of Concern

The O_3 ISA identifies health effects associated with exposures to varying concentrations of O_3. However, not all of the evidence is suitable for evaluation in a quantitative risk assessment. Estimating exposures to ambient O_3 concentrations at and above benchmark levels where health effects have been observed in studies provides a perspective on the public health impacts of O_3-related health effects that have been demonstrated in human clinical and toxicological studies but cannot currently be evaluated in quantitative risk assessments, such as lung inflammation, increased airway responsiveness, and decreased resistance to infection.

To inform the selection of benchmark levels for O_3 exposure, it is appropriate to consider the evidence from clinical studies which have evaluated individual controlled levels of O_3 exposure. There is substantial clinical evidence demonstrating a range of O_3-related effects including lung inflammation and airway responsiveness in healthy individuals at an exposure level of 0.080 ppm. There is additional evidence that asthmatics have larger and more serious effects than healthy people at 0.070 ppm, as well as a substantial body of epidemiological evidence of associations with O_3 levels that extend well below 0.080 ppm. There is a more limited set of evidence based on clinical studies of healthy individuals exposed at 0.060 ppm in which O_3-related effects have been observed. This is the lowest level at which any O_3-related effects have been observed in clinical studies of healthy individuals (U.S. EPA, 2013, section 6.2.1).

Thus, benchmark levels of 0.060 ppm, 0.070 ppm, and 0.080 ppm are used in this assessment to characterize exposures of concern for a range of potential health effects in healthy and at-risk populations exposed to O_3.

2.2.6 Health Endpoints

The O_3 ISA identifies a wide range of health outcomes associated with short-term exposure to ambient O_3, including an array of morbidity effects as well as premature mortality. The ISA also identifies several morbidity effects and some evidence for premature mortality associated with longer-term exposures to O_3. In identifying health endpoints for risk assessment, we have focused on endpoints that pertain to at-risk populations, have public health significance, and for which information is sufficient to support a quantitative concentration-response

relationship, in the case of epidemiological studies, or exposure-response relationship, in the case of controlled human exposure studies.[4]

In considering such endpoints for O3, we draw from two types of studies: controlled human exposure and epidemiological studies. Each study type informs our characterization of O3 risk and can do so in different ways. Estimates of risk based on results of controlled human exposure studies are valuable because they provide clear evidence of the detrimental effects of controlled (and measured) exposures to O_3 over multiple hours on lung function at moderate levels of exertion. Results of these studies can be applied to modeled estimates of population exposure to provide insights into population exposure characteristics, including types of activity patterns and microenvironments, which are associated with high levels of risk. Controlled human exposure studies, however, cannot directly provide relationships for endpoints such as premature death or hospitalizations, focusing more on intermediate biological endpoints including inflammatory, blood, neurological, cardiovascular, and respiratory biomarkers or symptoms. Estimates of risk based on concentration-response functions from observational epidemiology studies can provide insights on risk for more serious or chronic health endpoints. For example, epidemiological studies of O_3 described in the ISA have evaluated associations between O_3 and various endpoints including respiratory symptoms, respiratory-related hospitalizations and emergency department (ED) visits, and premature mortality (U.S. EPA, 2013, sections 6.2.9 and 6.3.4). Epidemiological studies also generally focus on a population residing in specific area, which may reflect a broad range of susceptibilities and sensitivities. Controlled human exposure studies typically involve a smaller number of individuals over a more limited range of health status, in some cases focused on at-risk populations, such as asthmatics and individuals with COPD. Lastly, while controlled human exposure studies directly measure the exposures eliciting the recorded effects, epidemiology studies have not traditionally been based on observations of personal exposure to ambient O_3, relying instead on surrogate measures of population exposure. Such surrogates are often based on simple averages of ambient O_3 monitor observations. Thus, with attention to their differing strengths and limitations, risk analyses based on each type of study can inform the risk characterization.

The O_3 ISA makes overall causal determinations based on the full range of evidence including epidemiological, controlled human exposure, and

[4] The distinction between concentration-response and exposure-response functions reflects the typical use of ambient concentrations as measured at monitor locations as surrogates for population exposures in observational epidemiology studies, as compared to the personal exposures to controlled concentrations of O_3 that are typically used in controlled human exposure studies. Both types of studies are intended to produce an exposure-responserelationship, however, the epidemiology studies are actually providing a concentration-response relationship, which captures the exposure-response relationship with errors in exposure measurement.

toxicological studies. Figure 2-1 shows the O_3 health effects which have been categorized by strength of evidence for causality in the O_3 ISA (U.S. EPA, 2013, chapter 2). The ISA determined there to be causal relationships between short-term exposure to ambient O_3 and respiratory effects, including respiratory-related morbidity and mortality and a likely causal relationship with all-cause total mortality and with cardiovascular effects; the evidence was concluded to be suggestive of a causal relationship between short-term exposure to ambient O_3 and central nervous system effects. The ISA determined to also be a likely causal relationship between long-term O_3 exposures and respiratory effects (including respiratory symptoms, new-onset asthma, and respiratory mortality), and determined the evidence to be suggestive of causal relationships between long-term O_3 exposures and total mortality as well as cardiovascular, reproductive and developmental, and central nervous system effects.

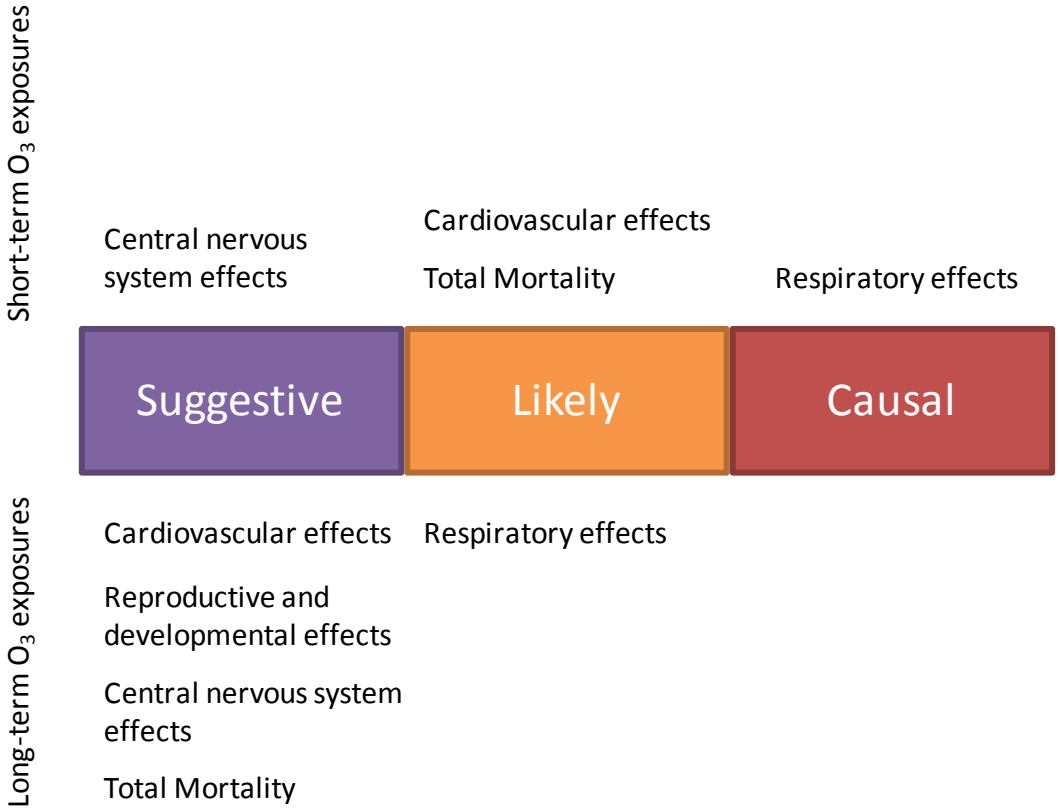


Figure 2-2 Causal Determinations for O_3 Health Effects

The ISA identifies several specific respiratory responses to short-term O_3 exposure that have been evaluated in controlled human exposure studies (U.S. EPA, 2013, section 6.2.1). These include decreased inspiratory capacity, decreased forced vital capacity (FVC) and forced expiratory volume in 1 second (FEV1); mild bronchoconstriction; rapid, shallow breathing patterns during exercise; symptoms of cough and pain on deep inspiration (PDI); and pulmonary inflammation. While such studies document quantitative relationships between short-term O_3 exposure and an array of respiratory-related effects, exposure-response data across a range of concentrations sufficient for developing quantitative risk estimates are only available for O_3-related decrements in FEV1 (U.S. EPA, 2013, section 6.2.1).

Within the broad category of respiratory morbidity effects, the epidemiology literature has provided effect estimates for a wide range of health endpoints associated with short-term O_3 exposures which we have considered for risk assessment. These health endpoints include lung function, respiratory symptoms and medication use, respiratory-related hospital admissions, and emergency department visits. In the case of respiratory symptoms, the evidence is most consistently supportive of the relationship between short-term ambient O_3 metrics and respiratory symptoms and asthma medication use in children with asthma, but not for a relationship between O_3 and respiratory symptoms in children without asthma (U.S. EPA, 2013, section 6.2.9). In the case of hospital admissions, there is evidence of associations between short-term ambient O_3 metrics and general respiratory-related hospital admissions as well as more specific asthma-related hospital admissions (U.S. EPA, 2013, section 6.2.7.2).

With regard to mortality, studies have evaluated associations between short-term ambient O_3 metrics and all-cause, non-accidental, and cause-specific (usually respiratory or cardiovascular) mortality. The evidence from respiratory-related morbidity studies provides strong support for respiratory-related mortality for which a causal determination has been made (U.S. EPA, 2013, Table 2-3). There are also a number of large studies that have found associations between O_3 and all-cause and all non-accidental mortality for which a likely causal determination has been made (U.S. EPA, 2013, Table 2-3). Thus, it is appropriate to assess risks for respiratory-related mortality as well as for all-cause total mortality associated with O_3 exposure. The ISA also reports a likely causal determination for short-term O_3 and cardiovascular effects, including cardiovascular mortality (U.S. EPA, 2013, Table 2-3). This determination is supported by studies relating total and cardiovascular mortality, coupled with evidence from animal toxicological studies and controlled human exposure studies which find effects of O_3 on systemic inflammation and oxidative stress. Cardiovascular mortality effects are covered through the estimation of risks associated with total mortality, which is dominated by cardiovascular mortality. There are not sufficient epidemiological studies of cardiovascular

morbidity showing consistent associations to justify inclusion of any cardiovascular morbidity endpoints in the quantitative risk assessment.

With regard to effects associated with long-term O_3 exposures, the ISA states that the relationship between O_3 and respiratory-related effects, including respiratory symptoms, new-onset asthma, and respiratory mortality is likely causal (U.S. EPA, 2013, Table 2-3). This suggests that for long-term exposures, when comparing the evidence for respiratory-related mortality and total mortality, the evidence is strongest for respiratory-related mortality, which is supported by the strong evidence for respiratory morbidity. As a result, it is appropriate to include respiratory mortality rather than total mortality in the risk assessment and to give consideration to inclusion of additional respiratory-related health endpoints.

2.2.7 Exposure and Concentration-response Functions for Health Endpoints

Estimation of risk requires characterization of the E-R and C-R functions along the full range of potential exposures. For E-R functions, the evidence from individual controlled human exposure studies provides responses for exposures at and above 60 ppb. McDonnell et al. (2012) develop an integrated model of FEV1 response that is fit to the results from controlled human exposure studies and find that a model with a threshold provides the best fit to the data. In addition, the ISA notes that it is difficult to characterize the E-R relationship at and below 40 ppb due to the dearth of data at these lower concentrations (U.S. EPA, 2013, section 2.5.4.4). Thus, for the portion of the risk assessment based on application of results of controlled human exposure studies, the threshold model is applied.

The evidence for a threshold in the C-R functions for mortality and morbidity outcomes derived from the epidemiological literature is limited. In general, the epidemiological evidence suggests a generally linear C-R function with no indication of a threshold. However, evaluation of evidence for a threshold in the C-R function is complicated by the high degree of heterogeneity between cities in the C-R functions and by the sparse data available at lower ambient O3 concentrations (U.S. EPA, 2013, sections 2.5.4.4 and 2.5.4.5).

The ISA also evaluated whether the magnitude of the relationship between short-term exposures to O_3 and mortality changes at lower concentrations (e.g., whether the C-R function is non-linear). The ISA concludes that epidemiologic studies that examined the shape of the C-R curve and the potential presence of a threshold have indicated a generally linear C-R function with no indication of a threshold in analyses that have examined 8-h max and 24-h avg O3 concentrations, and that the evidence supports less certainty in the shape of the C-R function at the lower end of the distribution of O_3 concentrations, e.g., 24-hour average O_3 below 20 ppb, due to the low density of data in this range (U.S. EPA, 2013, section 2.5.4.4). In the absence of information in the scientific literature on alternative forms of C-R functions at low O_3

concentrations, the best estimate of the C-R function is a linear, no-threshold function. The scientific literature does not provide sufficient information with which to quantitatively characterize any potential additional uncertainty in the C-R functions at lower O_3 concentrations for use in the quantitative risk assessment.

Multiple exposures to elevated O_3 levels over the course of an O_3 season may result in adaptation within exposed population. Evidence suggests that repeated or chronic exposures to elevated O_3 can result in morphologic and biochemical adaptation which reduces the impacts of subsequent O_3 exposures (U.S. EPA, 2013, section 6.2.1.1). This has implications for exposure modeling, in that the effects of modeled repeat exposures on risk may be attenuated relative to the effects of the initial exposures. The ISA notes that "neither tolerance nor attenuation should be presumed to imply complete protection from the biological effects of inhaled O3, because continuing injury still occurs despite the desensitization to some responses (U.S. EPA, 2013, section 6.2.1.1)." The ISA reports that there are limited epidemiological studies evaluating adaptation to the mortality effects of O_3, although the limited evidence does suggest that mortality effects are decreased in later months during the O_3 season relative to earlier months (U.S. EPA, 2013, section 6.3.3). The impact of this phenomenon on risks based on application of results from epidemiological studies is likely to be small, because the relative risk estimates from those studies already incorporate any adaptive phenomenon.

2.2.8 At-risk Populations

The O_3 ISA refers to "at-risk" populations as an all-encompassing term used for groups with specific factors that increase the risk of an air pollutant- (e.g., O_3) related health effect in a population group (U.S. EPA, 2013, chapter 8). Populations or lifestages can experience elevated risks from O_3 exposure for a number of reasons. These include high levels of exposure due to activity patterns which include a high duration of time in high-O_3 locations, e.g., outdoor recreation or work, high levels of activity which increase the dose of O_3, e.g., high levels of exercise, genetic or other biological factors, e.g., life stage, which predispose an individual to sensitivity to a given dose of O_3, pre-existing diseases, e.g., asthma or COPD, and socioeconomic factors which may result in more severe health outcomes, e.g., low access to primary care that can lead to increased emergency department visits or hospital admissions. To consider risks to these populations, modeling of exposures to O_3 needs to incorporate information on time spent by potentially at-risk populations in high O_3 locations. This requires identification of populations with the identified exposure-related risk factors, e.g. children or adults engaging in activities involving moderate to high levels of outdoor exertion, especially on a repeated basis typical of student athletes or outdoor workers, as well as identifying populations with high sensitivity to O_3, e.g. asthmatic children. It also requires that information on O_3

concentrations be mapped to locations where at-risk populations are likely to be exposed, e.g. near roadways where running may occur, or at schools or parks where children are likely to be engaged in outdoor activities.

In addition to consideration of factors that lead to increased exposure to O_3, modeling of risk from O_3 exposures should incorporate additional information on factors that can lead to increased dose of O_3 for a given exposure, e.g., increased breathing rates during periods of exertion. These factors are especially important for risk estimates based on application of the results of controlled human exposure studies. For risk modeling based on application of observational epidemiology results, it is also important to understand characteristics of study populations that can impact observed relationships between ambient O_3 and population health responses.

The O_3 ISA identifies a number of factors which have been associated with modifications of the effect of ambient O_3 on health outcomes. Building on the causal framework used throughout the O_3 ISA, conclusions are made regarding the strength of evidence for each factor that may contribute to increased risk of an O_3-related health effect based on the evaluation and synthesis of evidence across scientific disciplines. The O_3 ISA categorizes potential risk modifying factors by the degree of available evidence. These categories include "adequate evidence," "suggestive evidence," "inadequate evidence," and "evidence of no effect." See Table 8-1 of the O_3 ISA for a discussion of these categories (U.S. EPA, 2013, chapter 8).

Factors categorized as having adequate evidence include asthma, lifestage (children less than 18 years of age, adults older than 65 years of age), diets with nutritional deficiencies, and working outdoors. For example, children are the group considered to be at greatest risk because they breathe more air per unit of body weight, are more likely to be active outdoors when O_3 levels are high, are more likely than adults to have asthma, and are in a critical time period of rapid lung growth and organ development. Factors categorized as having suggestive evidence include genetic markers, sex (some studies have shown that females are at greater risk of mortality from O_3 compared to males), low socioeconomic status, and obesity. Factors characterized as having inadequate evidence include influenza and other respiratory infections, COPD, cardiovascular disease, diabetes, hyperthyroidism, race, and smoking (U.S. EPA, 2013, section 8.5, Table 8-6).

2.3 URBAN-SCALE MODELING OF INDIVIDUAL EXPOSURE

Estimates of human exposure to O_3 provide important information to inform consideration of policy-relevant questions identified in Section 2.2 regarding the occurrence of exposures of concern under air quality conditions that meet existing and potential alternative standards, and also to provide inputs to the portion of the risk assessment based on application of

results of controlled human exposure studies. Studies that measure human exposure to O_3 are limited. More commonly, human exposure is estimated using sophisticated models which combine information on ambient O_3 concentrations in various microenvironments, e.g. near roads, in schools, etc., with information on activity patterns for individuals sampled from the general population or specific subpopulations, e.g. children with asthma.

O_3 exposure is highly dependent on the ambient O_3 concentrations in an urban area. Given that these concentrations are variable from year to year, it is important to model multiple years representing the range of variability on O_3 concentrations to provide a better characterization of potential exposures of concern. In addition, other important sources of variability and uncertainty affecting the exposure estimates should be characterized, including uncertainty and variability in the data on time-activity patterns, O_3 concentrations, and population inputs. This can be accomplished in part by modeling exposure for multiple urban areas selected to represent variability in these underlying sources of variability.

This section briefly describes the conceptual foundation for key components of exposure modeling, characterization of microenvironmental O_3 concentrations, and characterization of human activity patterns, including behaviors intended to avert exposures to O_3. In addition, a brief discussion of key factors to consider in selecting urban case study areas for the exposure analysis is provided. The specific exposure model used in this assessment, APEX, is described more fully in Chapters 3 and 5. Characterization of ambient O_3 concentrations is discussed earlier in this chapter and in greater detail in Chapter 4.

2.3.1 Microenvironmental O_3 Concentrations

Human exposure to O_3 involves the contact (via inhalation) between a person and the pollutant in the various locations (or microenvironments) in which people spend their time. O_3 concentrations in some indoor microenvironments, such as within homes or offices, are considerably lower than O_3 concentrations in similarly located outdoor microenvironments, primarily due to deposition processes and the transformation of O_3 into other chemical compounds within those indoor microenvironments. Concentrations of O_3 may also be quite different in roadway environments, such as might occur while an individual is in a vehicle.

Thus, three important classes of microenvironments that should be considered when evaluating population exposures to ambient O_3 are indoors, outdoors, and in-vehicle. Within each of these broad classes of microenvironments, there are many subcategories, reflecting types of buildings, types of vehicles, etc. The O_3 ISA evaluated the literature on indoor-outdoor O_3 concentration relationships and found that studies consistently show that indoor concentrations of O_3 are often substantially lower than outdoor concentrations unless indoor sources are present. This relationship is greatly affected by the air exchange rate, which can be affected by open

windows, use of air conditioning, and other factors. Ratios of indoor to outdoor O_3 concentrations generally range from about 0.1 to 0.4 (U.S. EPA, 2013, section 4.3.2). In some indoor locations, such as schools, there can be large temporal variability in the indoor-outdoor ratios because of differences in air exchange rates over the day. For example, during the school day, there is an increase in open doors and windows, so the indoor-outdoor ratio is higher during the school day compared with an overall average across all hours and days. In-vehicle concentrations are also likely to be lower than ambient concentrations, although the literature providing quantitative estimates is smaller. Studies of personal exposure to O_3 have identified that O_3 exposures are highest when individuals are in outdoor microenvironments, such as walking outdoors midday, moderate when in vehicle microenvironments, and lowest in residential indoor microenvironments (U.S. EPA, 2013, section 4.3.3). Thus the time spent indoors, outdoors, and in vehicles is likely to be a critical component in estimating O_3 exposures.

Because of localized chemistry, O_3 concentrations on or near roadways can be much lower than away from roadways. This is due to the high levels of NO_X emissions from motor vehicles, which can lead to NOx titration of O_3, reducing O_3 levels during times of peak traffic. The ISA reports evidence that concentrations of NO, NO_2, and NOx are negatively correlated with concentrations of O_3 near busy roadways. Because few monitors are located in direct proximity to roadways, it is important to account for differences between near-road O_3 concentrations and ambient O_3 measurements in modeling exposure.

2.3.2 Human Activity Patterns

Human exposure can be measured using several metrics. Exposure to ambient concentrations is one such metric. It is also possible to model dose, which combines exposure information with physiological parameters related to activity levels. In order to model exposure to ambient concentrations, detailed information on the patterns of time spent in different microenvironments is critical. In order to model O_3 dose, additional information on the activities conducted while in those microenvironments is needed, along with data on physiological parameters associated with different activities.

Several large-scale databases of human time-activity-location patterns have been compiled. The most comprehensive of these databases in the Consolidated Human Activity Database (CHAD), which has been the basis of several previous exposure analyses for previous NAAQS reviews. These databases compile large numbers of diaries of time spent at different activities in different locations collected as part of smaller studies. The ISA notes the high degree of variability in activity patterns across the population, as well as the variability in time spent in different microenvironments. Time-activity-location patterns vary by age group, as well as by

region of the U.S. Children generally spend more time in outdoor locations and also generally have higher activity levels in those environments.

The dose of O3 received for any given exposure in a microenvironment depends not only on the activity levels and O_3 concentrations in the microenvironment, but also on ventilation rates, which are related to age, body weight, and other physiological parameters. Children generally have lower ventilation rates than adults when considering the volume of air breathed per unit time; however, they tend to have a greater oral breathing contribution than adults, and due to smaller lung volumes and generally greater breathing frequencies, children breathe at higher body mass or surface area normalized minute ventilation rates, relative to their lung volumes. Both of these factors tend to increase their applied or intake dose normalized to lung surface area. For example, when comparing daily body mass normalized ventilation rates, children can have up to a factor of two greater ventilation rates when compared to that of adults. During periods of high activity, ventilation rates for children and young adults can be nearly double those during moderate activity. Thus, it is important to model levels of activity and associated ventilation rate as well as time spent in different microenvironments.

In addition to modeling daily exposures, it may also be important to understand the patterns of exposure over an O_3 season, including multiple repeated exposures for the same individuals. Some individuals or subpopulations may exhibit multiple high daily exposures due to persistent patterns of high activity in microenvironments with high O_3 concentrations. For example, children engaged in numerous outdoor sports over a summer O_3 season may have multiple exposures to elevated O_3 levels.

Another important issue in characterizing exposure involves consideration of the extent to which people in relevant population groups modify their behavior for the purpose of decreasing their personal exposure to O_3 based on information about predicted air quality levels made public through the Air Quality Index (AQI). The AQI is the primary tool EPA has used to communicate information on predicted occurrences of high levels of O_3 and other pollutants. The AQI provides both the predicted level of air quality in an area along with a set of potential actions that individuals and communities can take to reduce exposure to air pollution and thus reduce the risk of health effects associated with breathing ambient air pollution. There are several studies, discussed in the O_3 ISA, that have evaluated the degree to which populations are aware of the AQI and what actions individuals and communities take in response to AQI values in the unhealthy range. These studies suggest that at-risk populations, such as children, older adults, and asthmatics, modify their behavior in response to days with bad air quality, most commonly by reducing their time spent outdoors or limiting their outdoor activity exertion level. A challenge remains in how to consider existing averting behaviors within the assessment tools we use and how best to use improved knowledge of participation rates, the varying types of

actions performed particularly by potentially at-risk individuals, and the duration of these averting behaviors to quantify the impact on estimated exposures and health risks.

2.3.3 Modeling of Exposures Associated with Simulating Just Meeting O_3 Standards

In order to address policy-relevant questions regarding changes in exposure associated with potential alternative standards, the exposure assessment evaluates changes in the O_3 concentrations, and the resulting changes in exposure, associated with simulating just meeting alternative standards relative to just meeting the existing standards. The new, model-adjustment methodology being implemented in this risk and exposure assessment provides for more realistic responses of hourly O_3 concentrations to changes in the precursor emissions that lead to O_3 formation. Characterization of exposure and changes in exposure when simulating just meeting the alternative standards are discussed in greater detail in Chapter 5.

2.3.4 Considerations in Selecting Urban Case Study Areas for the Exposure Analysis

The goal of the urban area exposure analysis is to characterize the variability in exposures for different locations, taking into account variability in essential factors that affect exposures. Important factors identified earlier that may influence exposure include time activity patterns, especially activities occurring in outdoor environments; demographics of the exposed population, e.g., age and income level; and O_3 concentrations. In addition to these factors, the selection of urban areas to include in the exposure analysis takes into consideration the location of O_3 epidemiological studies (for comparability with the risk assessments), the availability of ambient O_3 data and specific exposure information (e.g., air conditioning prevalence), and the desire to represent a range of geographic areas. To make the exposure analysis most useful in addressing the key policy-relevant questions, urban case study areas were also chosen such that most of them exceeded the existing 8-hr O_3 standards and potential alternative standards during the time period of interest.

2.4 RISK ASSESSMENT

Assessment of risk entails joint consideration of the exposure to a hazard, frequency of adverse outcomes given exposure, and severity of resulting adverse outcomes. A risk assessment for O_3 requires characterization of exposures to ambient O_3 for relevant populations, identification of appropriate dose-response or concentration-response functions linking O_3 with adverse health outcomes, and characterizing risks for individuals and populations.

As discussed above, there are two classes of studies that have provided information to inform the risk modeling: controlled human exposure studies and observational epidemiology studies. The conceptual approach to risk assessment varies based on which type of study result is being applied. This section briefly describes the conceptual foundation for several aspects of risk

modeling, including the concept of attributable risk, modeling of total risk and incremental risk reductions, development of risk estimates based on controlled human exposure studies, and development of risk estimates based on results of observational epidemiology studies.

This section briefly describes the conceptual foundation for key elements of risk modeling, including a discussion of the concept of attributable risk, modeling of risk for total O_3 exposure and the distribution of risk over O_3 concentrations, modeling of risk reductions associated with alternative standards, and key factors to consider in selecting urban case study areas for the risk analysis. Characterization of ambient O_3 concentrations is discussed earlier in this chapter and in greater detail in Chapter 4. The specific risk models used in the urban case study area risk analyses, APEX for analyses based on application of controlled human exposure studies and BenMAP for analyses based on application of observational epidemiology studies, are described more fully in Chapters 6 and 7, respectively. Chapter 8 provides an additional national-scale assessment of mortality risk associated with recent O_3 concentrations, to provide context for evaluating the magnitude of health risks in the urban case study areas and to evaluate the representativeness of the urban case study areas in estimating O_3 risks.

2.4.1 Attributable Risk

This risk and exposure assessment relies on the concept of attributable risk in evaluating both total risk and incremental changes in risk associated with just meeting existing and potential alternative O_3 standards. Attributable risk is defined as the difference in incidence of an adverse effect between an exposed and unexposed population for a specific stressor. Attributable risk is an important concept when addressing risks that are associated with multiple causes, such as mortality and respiratory hospital admissions.

Estimates of attributable risk require either an exposure-response (E-R) function (for analyses based on results of controlled human exposure studies) or a concentration-response (C-R) function (for analyses based on results of epidemiology studies).

E-R functions require estimates of exposure, in this case supplied by the APEX modeling described above. In the case of the lung function endpoint evaluated in this risk analysis, the E-R function also requires information on age and exertion levels to predict the impact of O_3 exposure on decrements in lung function. E-R functions may provide estimates of the incidence of an endpoint or the probability of exceeding benchmark decrement levels.

C-R functions derived from relative risk estimates reported in the epidemiological literature generally require estimates of ambient O_3 concentrations, baseline incidence rates, and estimates of exposed populations. Ambient O_3 concentrations should generally be constructed to match the spatial and temporal averaging used in the underlying epidemiology study; e.g., a

study may have used a spatial average over a metropolitan statistical area of the 8-hour daily maximum.

As with exposure, attributable risk is highly dependent on the ambient O_3 concentrations in an urban area. Given that these concentrations are variable from year to year, it is important to model multiple years representing the range of variability of O_3 concentrations to provide a better characterization of risk. In addition, other important sources of variability and uncertainty affecting the risk estimates should be characterized, including uncertainty and variability in the C-R and E-R functions, O_3 concentrations and O_3 exposure, and population inputs. This can be accomplished in part by modeling risk for multiple urban areas selected to represent variability in these underlying risk drivers.

2.4.2 Modeling of Risk for Total Exposure to O_3

As discussed earlier in this chapter, ambient O_3 is contributed to by emissions from a variety of sources, including natural, U.S. anthropogenic, and non-U.S. anthropogenic sources. Once in the atmosphere, O_3 molecules created from these different sources of emissions are not distinguishable. Individuals and populations are exposed to total O_3 from all sources, and risks associated with O_3 exposure are due to total O_3 exposure and do not vary for O_3 exposure associated with any specific source. Given the absence of a detectable threshold in the available C-R functions, total risk attributable to O_3 will thus be the risk associated with total exposure to O_3, with no threshold or cutpoint applied. To address certain policy-related questions, it is possible to approximately attribute risk to specific sources through the use of air quality modeling techniques, and this is explored in the Policy Assessment. However, these techniques are based on applying model results to total O_3 risk, rather than on directly modeling risk attributable to specific sources.

As discussed earlier in this chapter, a critical policy-relevant risk question is the O_3 attributable risk remaining after just meeting the existing O_3 standards. This risk includes risks associated with O_3 from all sources after we have simulated just meeting the existing daily 8-hour maximum standard level of 75 ppb. The estimates of total risk remaining after meeting the existing standard form the reference values for evaluating reductions in risk associated with just meeting alternative levels of the standard.

In addition to providing risk estimates for urban case study areas, it is also useful to evaluate O_3 risks across the entire U.S., both to better understand the total magnitude of the health burden associated with O_3 and to evaluate the representativeness of selected urban case study areas in characterizing the range and variability in risks across the U.S. The national-scale risk assessment presented in Chapter 8 is focused on estimating risk associated with recent O_3 concentrations, rather than on risk after just meeting existing or alternative standards. This is the

appropriate focus for the national analysis, because the techniques used to simulate just meeting existing and alternative standards in urban case study areas are less certain in a national context due to concerns about interdependence between air quality responses in different urban areas; e.g., just meeting a standard in one urban area would likely have impacts on O_3 air quality in surrounding urban areas. It is beyond the scope of this REA to attempt to simulate control strategies that would result in national attainment of existing or alternative primary health standards.

2.4.3 Distributions of Risk Across O_3 concentrations

Total O_3 risk for the O_3 season is calculated by summing daily risks across all days in the O_3 season. Because of the high degree of variability in daily O_3 concentrations across an O_3 season, total O_3 risk will include risks calculated for some days with high O_3 concentrations as well as for some days with very low O_3 concentrations. Therefore it is appropriate to provide the distribution of total risk over the range of daily O_3 concentrations to allow for an understanding of how O_3 concentrations on different days are contributing to the estimates of total risk. In addition, as noted in the ISA and discussed above, because of the relatively lower density of data on days with low concentrations of O_3, there is decreased confidence in the shape of the C-R function at lower O_3 concentrations, and therefore lower confidence in risk estimates for days with lower O_3 concentrations, especially in the range below 20 ppb. As a result, it is appropriate to provide the distribution of total risk over the range of daily O_3 concentrations to allow for better characterization of confidence in the estimates of total risk.

2.5 MODELING OF RISKS ASSOCIATED WITH SIMULATING JUST MEETING O_3 STANDARDS

In order to address policy-relevant questions regarding changes in risk associated with potential alternative standards, the risk assessment evaluates changes in the distribution of O_3 concentrations, and the resulting changes in risk, associated with simulating just meeting alternative standards relative to just meeting the existing standards. The new, model-adjustment methodology being implemented in this risk and exposure assessment provides for more realistic responses of hourly O_3 concentrations to changes in the precursor emissions that lead to O_3 formation. As noted earlier there are multiple combinations of reductions in precursor emissions that can result in just meeting alternative standards. As a result, there is variability in the potential changes in the distribution of O_3 concentrations and risk that would result from just meeting existing and alternative standards. Characterization of this variability, as well as uncertainties in the simulation of just meeting the standards, will be included in Chapters 6 and 7.

2.6 CONSIDERATIONS IN SELECTING URBAN CASE STUDY AREAS FOR THE RISK ANALYSIS

The goal of the urban area risk analysis is to characterize the magnitude of risk and the impact on risk of meeting existing and potential alternative standards. The selection of specific urban case study areas is based on a set of factors reflecting both variability in factors that affect risk and availability of high quality input data, to provide risk estimates that have higher overall confidence. Important factors identified earlier that may influence risk include O_3 concentrations, demographics, exposure factors, and magnitude of the effect estimate in the C-R function. In addition to consideration of variability in these factors, urban areas are preferentially selected if they have O_3 concentrations that are above the existing standards and potential alternative standards, if they have suitable epidemiological studies to provide C-R functions for mortality or morbidity, if they have adequate monitoring data available to characterize population exposures, and if they have appropriate baseline health incidence data available.

2.7 RISK CHARACTERIZATION

Risk characterization is the process of communicating the results of risk (and exposure) modeling in metrics that have meaning to decision makers. In the specific context of this review, this translates into providing metrics that are most useful in the Policy Assessment to assess the adequacy of the existing O_3 standards in protecting public health with an adequate margin of safety and to evaluate the additional protection provided by potential alternative standards.

Risk characterization requires careful translation of very complex outputs of exposure and risk models into simpler metrics, for example, translating hourly O_3 exposures in various microenvironments into estimates of population exposures above alternative exposure benchmarks. Risk characterization also requires the condensation of a large number of analytical steps and results to (a) summarize the results of the risk analysis, usually taking detailed results and condensing them into a more aggregate interpretation while still providing information about heterogeneity across space and time; (b) communicate the sensitivity of results to different modeling assumptions; and (c) characterize the qualitative and quantitative uncertainty in results.

As described more fully in Chapter 5 and in the Policy Assessment, EPA has selected, based on providing a reasonable measure of exposures of concern for at-risk populations and lifestages, aggregate exposure metrics including the number and percent of certain highly vulnerable populations exposed to levels of O_3 above exposure levels that have been identified in the scientific literature as associated with adverse respiratory responses. As noted in section 2.3.1, these benchmark exposure levels are 0.060 ppm, 0.070 ppm, and 0.080 ppm. Highly vulnerable populations include active children, older adults, and outdoor workers.

As described more fully in Chapters 6 and 7 and in the Policy Assessment, EPA has selected, based on providing characterization of risks to the public including at-risk populations and lifestages, aggregate risk metrics including the number and percent of vulnerable populations experiencing adverse respiratory responses based on application of results of controlled human exposure studies and the attributable incidence and percent of baseline incidence of mortality and morbidity endpoints based on application of results of epidemiology studies.

For all three types of metrics (exposure, risk based on controlled human exposure studies, and risk based on epidemiology studies) and for the purpose of evaluating the adequacy of the existing standards, the focus is on the exposure and risk remaining upon just meeting the existing standards. For the purpose of evaluating alternative standards, the focus in on the changes in exposure and risk after simulating just meeting the alternative standards, compared to exposures and risk after simulating just meeting the existing standards.

As detailed in Chapter 3, quantitative sensitivity analyses are provided to evaluate the impacts of critical inputs to the exposure and risk modeling. Limited quantitative uncertainty analyses are also included, along with a comprehensive qualitative uncertainty assessment. The overall treatment of uncertainty is guided by the WHO guidelines for uncertainty assessment (World Health Organization, 2008). These guidelines recommend a tiered approach in which progressively more sophisticated methods are used to evaluate and characterize sources of uncertainty depending on the overall complexity of the risk assessment.

In order to inform considerations of overall confidence in the risk estimates derived from application of C-R functions derived from the epidemiological literature, we provide the distributions of total risk across the entire range of daily 8-hour maximum O_3 concentrations. In addition, we provide an assessment of the representativeness of the urban areas selected for the risk and exposure analysis in characterizing the overall distribution of risk across the U.S. This assessment evaluates how well the selected urban areas capture important characteristics that are associated with risk, including demographics, air quality levels, and factors affecting exposure such as air conditioning prevalence.

2.8 REFERENCES

Frey, C. and J. Samet. 2012. "CASAC Review of the EPA's Policy Assessment for the Review of the Ozone National Ambient Air Quality Standards (First External Review Draft – August 2012)." U.S. Environmental Protection Agency Science Advisory Board, EPA-CASAC-13-003.

U.S. Environmental Protection Agency. 2012a. Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume I: User's Guide. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001a). <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2012 b. "Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume II: Technical Support Document." Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001b).

< http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2013. Integrated Science Assessment of Ozone and Related Photochemical Oxidants (Final Report). Washington, DC: EPA Office of Air and Radiation. (EPA document number EPA/600/R-10/076F).

U.S. EPA. 2013. "Environmental Benefits Mapping Analysis Program Community Edition (BenMAP-CE v1.0)," posted on December 03, 2013, <http://www.epa.gov/air/benmap/>.

World Health Organization. 2008. "Part 1: Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment, Harmonization Project Document No. 6." Published under joint sponsorship of the World Health Organization, the International Labor Organization and the United Nations Environment Programme. WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 2476).

3 SCOPE

This chapter provides an overview of the scope and key design elements of this quantitative exposure and health risk assessment. The design of this assessment began with a review of the exposure and risk assessments completed during the last O_3 NAAQS review (U.S. EPA, 2007a,b), with an emphasis on considering key limitations and sources of uncertainty recognized in that analysis.

As an initial step in the current O_3 NAAQS review in October 2009, EPA invited outside experts, representing a broad range of expertise (e.g., epidemiology, human and animal toxicology, statistics, risk/exposure analysis, atmospheric science), to participate in a workshop with EPA staff to help inform EPA's plan for the review. The participants discussed key policy-relevant issues that would frame the review and the most relevant new science that would be available to inform our understanding of these issues. One workshop session focused on planning for quantitative risk and exposure assessments, taking into consideration what new research and/or improved methodologies would be available to inform the design of quantitative exposure and health risk assessment. Based in part on the workshop discussions, EPA developed a draft IRP (U.S. EPA, 2009) outlining the schedule, process, and key policy-relevant questions that would frame this review. On November 13, 2009, EPA held a consultation with CASAC on the draft IRP (74 FR 54562, October 22, 2009), which included opportunity for public comment. The final IRP incorporated comments from CASAC (Samet, 2009) and the public on the draft plan, as well as input from senior Agency managers. The final IRP included initial plans for quantitative risk and exposure assessments for both human health and welfare (U.S. EPA, 2011a, chapters 5 and 6).

As a next step in the design of these quantitative assessments, OAQPS staff developed more detailed planning documents, the O_3 *National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment* (Health Scope and Methods Plan, U.S. EPA, 2011b) and the O_3 *National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment* (Welfare Scope and Methods Plan, U.S. EPA, 2011c). These Scope and Methods Plans was the subject of a consultation with CASAC on May 19-20, 2011 (76 FR 23809, April 28, 2011). Based on consideration of CASAC (Samet, 2011) and public comments on the Scope and Methods Plans, and information in the second draft ISA, we modified the scope and design of the quantitative risk assessment and provided a memo with updates to information presented in the Scope and Methods Plans (Wegman, 2012). The Scope and Methods Plans together with the update memo provide the basis for the discussion of the scope of this exposure and risk assessment provided in this chapter. This chapter also reflects

comments received from CASAC based on their review of the first draft Risk and Exposure Assessment on September 11-12, 2012 (Frey and Samet, 2012).

In presenting the scope and key design elements of the current risk assessment, this chapter first provides a brief overview of the quantitative exposure and risk assessment completed for the previous O_3 NAAQS review in section 3.1, including key limitations and uncertainties associated with that analysis. The remaining sections describe the current exposure and risk assessment, following the general conceptual framework described in Chapter 2. Section 3.2 provides a summary of the design of the urban-scale exposure assessment. Section 3.3 provides a summary of the design of the urban-scale risk assessment based on application of results of human clinical studies. Section 3.4 provides a summary of the design of the urban-scale risk assessment based on application of results of epidemiology studies. Section 3.5 provides a summary of the design of the national-scale risk burden assessment and representativeness analysis.

3.1 OVERVIEW OF EXPOSURE AND RISK ASSESSMENTS FROM LAST REVIEW

The exposure and health risk assessment conducted in the review, completed in March 2008, developed exposure and health risk estimates for 12 urban areas across the U.S. which were chosen based on the location of O_3 epidemiological studies and availability of ambient O_3 data and to represent a range of geographic areas, population demographics, and O_3 climatology. That analysis was in part based upon the exposure and health risk assessments included in the review completed in 1997.[1] The exposure and risk assessment incorporated air quality data (i.e., 2002 through 2004), and provided annual or O_3 season-specific exposure and risk estimates for these recent years of air quality and for air quality scenarios simulating just meeting the existing 8-hour O_3 standard and several alternative 8-hour O_3 standards.

3.1.1 Overview of exposure assessment from last review

Exposure estimates were used as an input to the risk assessment for lung function responses (a health endpoint for which exposure-response functions were available from controlled human exposure studies). Exposure estimates were developed for the general population and population groups including school-age children with asthma as well as all school-age children. The exposure estimates also provided information on population exposures

[1] In the 1994-1997 O_3 NAAQS review, EPA conducted exposure analyses for the general population, children who spent more time outdoors, and outdoor workers. Exposure estimates were generated for 9 urban areas for as is air quality and for just meeting the existing 1-hour standard and several alternative 8-hour standards. Several reports that describe these analyses can be found at: http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_pr.html.

exceeding potential health effect benchmark levels that were identified based on the observed occurrence of health endpoints not explicitly modeled in the health risk assessment (e.g., lung inflammation, increased airway responsiveness, and decreased resistance to infection) associated with 6-8 hour exposures to O_3 in controlled human exposure studies.

The exposure analysis took into account several important factors including the magnitude and duration of exposures, frequency of repeated high exposures, and breathing rate of individuals at the time of exposure. Estimates were developed for several indicators of exposure to various levels of O_3 air quality, including counts of people exposed one or more times to a given O_3 concentration while at a specified breathing rate and counts of person-occurrences (which accumulate occurrences of specific exposure conditions over all people in the population groups of interest over an O_3 season).

As discussed in the 2007 Staff Paper (U.S. EPA, 2007c) and in Section II a of the O_3 Final Rule (73 FR 16440 to 16442, March 27, 2008), the most important uncertainties affecting the exposure estimates were related to modeling human activity patterns over an O_3 season, modeling of variations in ambient concentrations near roadways, and modeling of air exchange rates that affect the amount of O_3 that penetrates indoors. Another important uncertainty, discussed in more detail in the Staff Paper (U.S. EPA, 2007c, section 4.3.4.7), was the uncertainty in energy expenditure values which directly affected the modeled breathing rates. These were important since they were used to classify exposures occurring when children were engaged in moderate or greater exertion. Health effects observed in the controlled human exposure studies generally occurred under these exertion levels for 6 to 8-hour exposures to O_3 concentrations at or near 0.08 ppm. Reports that describe these analyses (U.S. EPA, 2007a, c; Langstaff, 2007) can be found at: http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_index.html.

3.1.2 Overview of risk assessment from last review

The human health risk assessment presented in the review completed in March 2008 was designed to estimate population risks in a number of urban areas across the U.S., consistent with the scope of the exposure analysis described above (U.S. EPA, 2007b, c). The risk assessment included risk estimates based on both controlled human exposure studies and epidemiological and field studies. O_3-related risk estimates for lung function decrements were generated using probabilistic exposure-response relationships based on data from controlled human exposure studies, together with probabilistic exposure estimates from the exposure analysis. For several other health endpoints, O_3-related risk estimates were generated using concentration-response relationships reported in epidemiological or field studies, together with ambient air quality concentrations, baseline health incidence rates, and population data for the various locations included in the assessment. Health endpoints included in the assessment based on

epidemiological or field studies included hospital admissions for respiratory illness in four urban areas, premature mortality in 12 urban areas, and respiratory symptoms in asthmatic children in 1 urban area.

In the health risk assessment conducted in the previous review, EPA recognized that there were many sources of uncertainty and variability in the inputs to the assessment and that there was significant uncertainty in the resulting risk estimates. The statistical uncertainty surrounding the estimated O_3 coefficients in epidemiology-based concentration-response functions as well as the shape of the exposure-response relationship chosen for the lung function risk assessment were addressed quantitatively. Additional uncertainties were addressed through sensitivity analyses and/or qualitatively. The risk assessment conducted for the previous O_3 NAAQS review incorporated some of the variability in key inputs to the assessment by using location-specific inputs (e.g., location-specific concentration-response functions, baseline incidence rates and population data, and air quality data for epidemiological-based endpoints, location specific air quality data and exposure estimates for the lung function risk assessment). In that review, several urban areas were included in the health risk assessment to provide some sense of the variability in the risk estimates across the U.S.

Key observations and insights from the O_3 risk assessment, in addition to important caveats and limitations, were addressed in Section II.B of the Final Rule notice (73 FR 16440 to 14 16443, March 27, 2008). In general, estimated risk reductions associated with going from then-current O_3 levels to just meeting the then-existing and alternative 8-hour standards showed patterns of decreasing estimated risk associated with just meeting the lower alternative 8-hour standards considered. Furthermore, the estimated percentage reductions in risk were strongly influenced by the baseline air quality year used in the analysis, which was due to significant year-to-year variability in O_3 concentrations. There was also noticeable city-to-city variability in the estimated O_3-related incidence of morbidity and mortality across the 12 urban areas. Uncertainties associated with estimated policy-relevant background (PRB) concentrations[2] were also addressed and revealed differential impacts on the risk estimates depending on the health effect considered as well as the location. EPA also acknowledged that at the time of the previous review there were considerable uncertainties surrounding estimates of O_3 C-R coefficients and the shape of concentration-response relationships and whether or not a population threshold or non-linear relationship exists within the range of concentrations examined in the epidemiological studies.

[2]Policy-relevant background (PRB) O_3 has been defined in previous reviews as the distribution of O_3 concentrations that would be observed in the U.S. in the absence of anthropogenic (man-made) emissions of O_3 precursor emissions (e.g., VOC, CO, NOx) in the U.S., Canada, and Mexico.

3.2 PLAN FOR THE CURRENT EXPOSURE AND RISK ASSESSMENTS

The Scope and Methods Plan, including updates (U.S. EPA, 2011b; Wegman, 2012), outlined a planned approach for conducting the current quantitative O_3 exposure and risk assessments, including broad design issues as well as more detailed aspects of the analyses. A critical step in designing the quantitative risk and exposure assessments is to clearly identify the goals for the analysis based on the policy-relevant questions identified in Chapter 2. We have identified the following goals for the urban area exposure and risk assessments: (1) to provide estimates of the percent of people in the general population and in sensitive populations with O_3 exposures above health-based benchmark levels; (2) to provide estimates of the percentage of the general population and in sensitive populations with impaired lung function (defined based on decrements in FEV_1) resulting from exposures to O_3; (3) to provide estimates of the potential magnitude of premature mortality associated with both short-term and long-term O_3 exposures, and selected morbidity health effects associated with short-term O_3 exposures; (4) to evaluate the influence of various inputs and assumptions on risk estimates to the extent possible given available methods and data; (5) to gain insights into the spatial and temporal distribution of risks and patterns of risk reduction and uncertainties in those risk estimates. For the exposure and risk analyses, we will estimate exposures and risks for recent ambient levels of O_3 and for O_3 concentrations after simulating just meeting the existing O_3 standard and potential alternative standards.

With regard to selecting alternative levels for the 8-hour O_3 standards for evaluation in the quantitative risk assessment, we base the range of levels on the evaluations of the evidence provided in the first draft PA, which received support from the CASAC in their advisory letter on the first draft PA. The first draft PA recommended evaluation of 8-hour maximum concentrations in the range of 60 to 70 ppb, with possible consideration of levels somewhat below 60 ppb. The upper end of this range is supported by the clear evidence from both clinical and epidemiological studies of effects at exposures of 70 ppb reported in the ISA and summarized in the first draft PA. The lower end of this range is based on considerations of evidence from clinical studies that have shown lung function decrements in healthy adult populations at 60 ppb O_3 exposures, and that 10 percent of healthy adults exposed to 60 ppb O_3 experienced lung function decrements that could be adverse to asthmatics. The evidence showing effects in healthy adults at exposures of 60 ppb supports the consideration of risks to sensitive populations at exposure levels below 60 ppb, although specific exposure levels below 60 ppb at which risks may be occurring are not supported by the evidence. An important distinction is that the evidence from controlled human exposure studies is based on exposures, while the standard

addresses ambient concentrations. Typically, exposures are lower than ambient concentrations because people spend a large fraction of their time indoors where O_3 concentrations are lower.[3]

Because of the year-to-year variability in O_3 concentrations that results from temporal variability in meteorology and emissions that drive O_3 formation, the exposure and risk assessments evaluate scenarios for meeting the existing and alternative standards based on multiple years of O_3 data. O_3 concentrations from 2006-2010 are used in estimating exposure and risk. This range of years captures a high degree of variability in meteorological conditions, as well as reflecting years with higher and lower emissions of O_3 precursors.

In order to provide greater confidence in the exposure and risk estimates, this REA uses an urban case study approach for assessing both exposure and risk. This approach provides greater confidence in estimates by allowing us to make use of air quality data, population information, health data, and epidemiology results that are well matched, and it does not require extrapolation of results to locations without these data. In addition, the urban case study approach allows us to simulate just meeting existing and alternative O_3 standards for each urban area, which is not currently feasible for health risk assessment at the national scale.[4] Specific selection criteria for case study urban areas included in the exposure and risk assessments are described in the following sections. In order to gain an understanding of how well the urban case study areas represent risks at a national level and to provide context for the urban case study results, we also include two national level analyses, 1) estimation of the national mortality burden associated with recent ambient O_3 and 2) characterization of how well the risk estimates for the set of urban areas modeled reflect the national distribution of mortality risk.

Throughout the exposure and risk analyses, we recognize that there are many sources of variability and uncertainty. Each analysis considers carefully the potential sources and significance of variability and uncertainties and, where data are available, provides quantitative assessment of variability and uncertainties, either through probabilistic analyses or through sensitivity or scenario analyses. In general the analyses follow the WHO guidelines for uncertainty assessment (World Health Organization, 2008), which recommend a tiered approach

[3] While almost all people spend a large fraction of their time indoors, there is high variability in this fraction between children and adults, and between outdoor workers and indoor workers. The ratio of exposures to ambient concentrations will likely be higher for children than adults, and for outdoor workers compared to indoor workers.

[4] In order to simulate just meeting alternative standards everywhere nationwide using the model-based adjustment approach employed in this REA, some areas would see O_3 design values decreased below the targeted standard level due to O_3 transport between locations. We were not able to devise an approach that would just meet the standard in every location simultaneously. Using the urban case study approach, we can, acknowledging the counterfactual nature of the analysis, assume independence of attainment for each urban case study area, which allows us to simulate just meeting the standards in each urban case study area.

in which progressively more sophisticated methods can be used to evaluate and characterize sources of uncertainty depending on the overall complexity, end use of the assessment, and resources and data available to conduct particular uncertainty characterizations.

The planned approaches for conducting the exposure and risk analyses are briefly summarized below. We begin with a general discussion of how uncertainty and variability are addressed in the different elements of the exposure and risk assessment. This is followed by a discussion of the air quality data that will be used in both the exposure and risk assessments and then discussions of each component of the exposure and risk assessments.

3.3 CHARACTERIZATION OF UNCERTAINTY AND VARIABILITY IN THE CONTEXT OF THE O_3 EXPOSURE AND RISK ASSESSMENT

An important component of this population exposure and health risk assessment is the characterization of both uncertainty and variability. Variability refers to the heterogeneity of a variable of interest within a population or across different populations. For example, populations in different regions of the country may have different behavior and activity patterns (e.g., air conditioning use and time spent indoors) that affect their exposure to ambient O_3 and thus the population health response. The composition of populations in different regions of the country may vary in ways that can affect the population response to exposure to O_3 – e.g., two populations exposed to the same levels of O_3 might respond differently if one population is older than the other. Variability is inherent and cannot be reduced through further research. Refinements in the design of a population risk assessment are often focused on more completely characterizing variability in key factors affecting population risk – e.g., factors affecting population exposure or response – in order to produce risk estimates whose distribution adequately characterizes the distribution in the underlying population(s).

Uncertainty refers to the lack of knowledge regarding the actual values of inputs to an analysis. Models are typically used in analyses, and there is uncertainty about the true values of the parameters of the model (parameter uncertainty) – e.g., the value of the coefficient for O_3 in a C-R function. There is also uncertainty about the extent to which the model is an accurate representation of the underlying physical systems or relationships being modeled (model uncertainty) – e.g., the shapes of C-R functions. In addition, there may be some uncertainty surrounding other inputs to an analysis due to possible measurement error—e.g., the values of daily O_3 concentrations in a risk assessment location or the value of the baseline incidence rate for a health effect in a population.[5]

[5] It is also important to point out that failure to characterize variability in an input used in modeling can also introduce uncertainty into the analysis. This reflects the important link between uncertainty and variability

In any risk assessment, uncertainty is, ideally, reduced to the maximum extent possible through improved measurement of key variables and ongoing model refinement. However, significant uncertainty often remains, and emphasis is then placed on characterizing the nature of that uncertainty and its impact on risk estimates. The characterization of uncertainty can be both qualitative and, if a sufficient knowledge base is available, quantitative.

The characterization of uncertainty associated with risk assessment is ideally addressed in the regulatory context using a tiered approach in which progressively more sophisticated methods are used to evaluate and characterize sources of uncertainty depending on the overall complexity and intended use of the risk assessment (WHO, 2008). Guidance documents developed by EPA for assessing air toxics-related risk and Superfund Site risks as well as recent guidance from the World Health Organization specify multitier approaches for addressing uncertainty.

Following the approach used for previous NAAQS risk and exposure assessments (U.S. EPA, 2008c, 2009b, 2010a, b), for the O_3 risk assessment, we are using a tiered framework developed by WHO to guide the characterization of uncertainty. The WHO guidance presents a four-tiered approach, where the decision to proceed to the next tier is based on the outcome of the previous tier's assessment. The four tiers described in the WHO guidance include:

Tier 0: recommended for routine screening assessments, uses default uncertainty factors (rather than developing site-specific uncertainty characterizations);

Tier 1: the lowest level of site-specific uncertainty characterization, involves qualitative characterization of sources of uncertainty (e.g., a qualitative assessment of the general magnitude and direction of the effect on risk results);

Tier 2: site-specific deterministic quantitative analysis involving sensitivity analysis, interval-based assessment, and possibly probability bounded (high-and low-end) assessment; and

Tier 3: uses probabilistic methods to characterize the effects on risk estimates of sources of uncertainty, individually and combined.

With this four-tiered approach, the WHO framework provides a means for systematically linking the characterization of uncertainty to the sophistication of the underlying risk assessment. Ultimately, the decision as to which tier of uncertainty characterization to include in a risk

---

with the effort to accurately characterize variability in key model inputs actually reflecting an effort to reduce uncertainty.

assessment will depend both on the overall sophistication of the risk assessment and the availability of information for characterizing the various sources of uncertainty.

This risk and exposure assessment for the O_3 NAAQS review is relatively complex, possibly warranting consideration of a full probabilistic (WHO Tier 3) uncertainty analysis. For the exposure assessment, we include probabilistic representations of important sources of variability; however, due to lack of information regarding reasonable alternative parameter settings for model input variable distributions, we are not able to include a complete probabilistic analysis incorporating both variability and uncertainty. Instead, we provide sensitivity analyses to explore the impact of specific model assumptions, and we include a comprehensive qualitative discussion of uncertainty regarding the model inputs and outputs.

While a full probabilistic uncertainty analysis is not undertaken for the epidemiology-based risk assessment due to limits in available information on distributions of model inputs, we provide a limited assessment using the confidence intervals associated with effects estimates (obtained from epidemiological studies) to incorporate statistical uncertainty associated with sample size considerations in the presentation of risk estimates. Technically, this type of probabilistic simulation represents a Tier 3 uncertainty analysis, although as noted here, it will be limited and only address uncertainty related to the fit of the C-R functions. Incorporation of additional sources of uncertainty related to key elements of C-R functions (e.g., competing lag structures, alternative functional forms, etc.) into a full probabilistic WHO Tier 3 analysis would require that probabilities be assigned to each competing specification of a given model element (with each probability reflecting a subjective assessment of the probability that the given specification is the correct description of reality). However, for most model elements there is insufficient information on which to base these probabilities. One approach that has been taken in such cases is expert elicitation; however, this approach is resource- and time-intensive, and, consequently, it is not feasible to use this technique in support of this O_3 risk assessment.

For most elements of the quantitative risk assessments, rather than conducting a full probabilistic uncertainty analysis, we include a qualitative discussion of the potential impact of uncertainty on risk results (WHO Tier 2). For some critical elements of the epidemiology-based risk assessment, e.g., the effect-estimate in the C-R function, we include sensitivity analyses to explore the potential impact of our assumptions. This falls under the WHO Tier 2 classification, although we are not able to assign probabilities to the sensitivity analyses. For these sensitivity analyses, we will include only those alternative specifications for input parameters or modeling approaches that are deemed to have scientific support in the literature (and so represent alternative reasonable input parameter values or modeling options). This means that the array of risk estimates presented in this assessment is expected to represent reasonable risk estimates that

can be used to provide some information regarding the potential impacts of uncertainty in the model elements.

3.4 AIR QUALITY CHARACTERIZATION

Figure 3-1 diagrams the basic information used in developing the air quality inputs for the REA. Air quality inputs to the urban area exposure and risk assessments include (1) recent air quality data developed from O_3 ambient monitors in each selected urban study area and (2) simulated air quality that reflects changes in the distribution of O_3 air quality estimated to occur when the urban area just meets the existing or alternative O_3 standards under consideration. In addition, O_3 air quality surfaces for recent years covering the entire continental U.S. were generated for use in the national-scale assessment. Details of the air quality data used in the REA are discussed in Chapter 4.

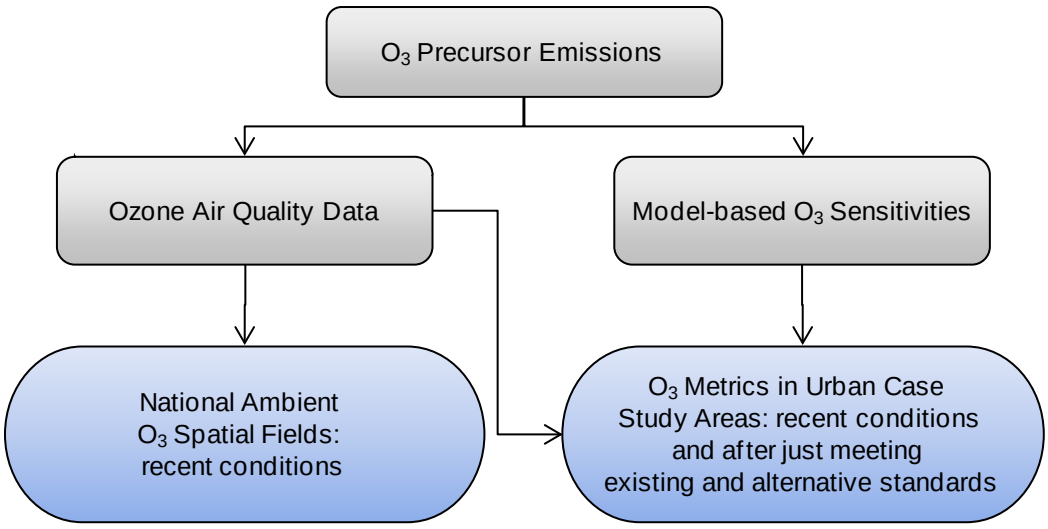


Figure 3-1 Conceptual Diagram for Air Quality Characterization in the Health REA

The urban case study area exposure and risk analyses are based on five recent years of air quality data, 2006-2010. We are including 5 years to reflect the considerable variability in meteorological conditions and the variation in O_3 precursor emissions that have occurred in recent years. The analyses focus on the O_3 season, which ranges from April to October in much of the nation but is longer in some warmer areas such as Los Angeles and Houston. The required O_3 monitoring seasons for the urban case study areas are described in more detail in Chapter 4.

In developing the O_3 air quality surfaces for the national-scale analysis, a combination of monitoring data and modeled O_3 concentrations are used to provide greater coverage across the

U.S. The procedure for fusing O_3 monitor data with modeling results is described further in Chapter 4.

Several O_3 metrics are generated for use in the urban area exposure and risk analyses. The exposure analyses use hourly O_3 concentrations, while the risk analyses use several different averaging times. The specific metrics used in each analysis are discussed further in following chapters. For the exposure analysis, hourly O_3 concentrations are interpolated to census tracts using Voronoi neighbor averaging (VNA), a distance weighted interpolation method (Gold, 1997; Chen et al., 2004). For the epidemiology-based risk analysis, we developed a composite of all monitors in the urban area for application with the epidemiology studies. We also evaluated several different definitions of the spatial boundaries of the urban areas that determined the monitors included in the spatial average. Some of the epidemiological studies specify a relatively narrow set of counties within an urban area, while others use a broader definition, such as all counties in a core based statistical area (CBSA) as defined by the Census Bureau. For those epidemiological studies that used a relatively narrow set of counties, most were based on counties in the center of the urban area. In most of these areas, the non-attaining O_3 monitors are not located in the center of the urban area, but instead in the surrounding areas, reflecting the transport and atmospheric chemistry governing O_3 formation. As a result, using a monitor set that exactly reflects the specific counties used in the epidemiology studies can exclude counties in an urban area that would realize the most risk reduction resulting from just meeting the O_3 standard. To better represent the changes in risk that could be experienced in the urban areas, the core risk estimates for all endpoints will be based on the CBSA definition. Sensitivity analyses are included to evaluate the effect of using only the counties in each urban area that specifically match the county set used in the epidemiology studies.

Simulation of just meeting the existing and alternative O_3 standards is accomplished by adjusting hourly O_3 concentrations measured over the O_3 season using a model-based adjustment methodology that estimates O_3 sensitivities to precursor emissions changes.[6] These sensitivities, which estimate the response of O_3 concentrations to reductions in anthropogenic NOx and VOC emissions, are developed using the Higher-order Decoupled Direct Method (HDDM) capabilities in the Community Multi-scale Air Quality (CMAQ) model. This modeling approach incorporates all known emissions, including sources of natural and anthropogenic emissions in and outside of the U.S. By using the model-based adjustment methodology we are able to more realistically simulate the temporal and spatial patterns of O_3 response to precursor emissions. We chose to

[6] In the first draft of this REA, we used a statistical quadratic rollback approach to simulate just meeting the existing O_3 standards. In that draft, we proposed using the model based approach that is being used in this draft, and received support for the model based approach from CASAC.

simulate just meeting the existing and alternative standards in the urban cast study areas by decreasing U.S. anthropogenic emissions of NOx and VOC throughout the U.S using equal proportional decreases in emissions throughout the U.S., in order to avoid any suggestion that we are approximating a specific emissions control strategy that a state or urban area might choose to meet a standard. More details on the HDDM-adjustment approach are presented in Chapter 4 of this REA and in Simon et al. (2013).

In the previous review, background O_3 (referred to in that review as policy relevant background, or PRB) was incorporated into the REA by calculating risk only in excess of PRB. CASAC members recommended that EPA move away from using PRB in calculating risks (Henderson, 2007). In addition, comments received from CASAC, based on their review of the first draft Risk and Exposure Assessment on September 11-12, 2012 (Frey and Samet, 2012), agreed with the development of risk estimates with reference to zero O_3 concentration. Based on these recommendations and comments, the second draft REA includes risks associated with O_3 from all sources after we have simulated just meeting the existing standard and estimates of total risk remaining after meeting alternative levels of the standards. EPA believes that presenting total risk is most relevant given that individuals and populations are exposed to total O_3 from all sources, and risks associated with O_3 exposure are due to total O_3 exposure and do not vary for O_3 exposure associated with any specific source. In addition, background O_3 is fully represented in estimates of total risk given that the measured and adjusted air quality concentrations being used in the risk and exposure analyses include O_3 produced from precursor emissions from both anthropogenic and background sources. The evidence and information on background O_3 that is assessed in the Integrated Science Assessment (ISA) is considered in the Policy Assessment (PA) in conjunction with the total risk estimates provided in this second draft REA. With regard to background O_3 concentrations, the PA will consider available information on ambient O_3 concentrations resulting from natural sources, anthropogenic sources outside the U.S., and anthropogenic sources outside of North America.

In providing a broader national characterization of O_3 air quality in the U.S., this REA draws upon air quality data analyzed in the O_3 ISA as well as national O_3 databases and modeling of O_3 using the Community Multiscale Air Quality (CMAQ) model. This information, along with additional analyses, is used to develop a broad characterization of recent air quality across the nation. This characterization includes O_3 levels in the urban case study areas for the time periods relevant to the risk analysis and information on the spatial and temporal characterization of O_3 across the national monitoring network. This information is then used to place the relative comparative attributes of the selected study areas into a broader national comparative context to help judge the overall representativeness of the selected study areas in characterizing O_3 risk for the nation. In addition, to better characterize the spatial patterns of

responses of the distribution of O_3 to just meeting existing and alternative O_3 standards, we also provide assessments of the historical patterns of responses of O_3 to emissions changes over time and an assessment of national patterns of responses to emissions changes relative to the spatial distribution of populations. These analyses are presented in more detail in Chapter 8 and Chapter 8 appendices.

3.5 EXPOSURE ASSESSMENT

Figure 3-2 diagrams the basic structure of the population exposure assessment. Basic inputs to the exposure assessment include the following: (1) recent measurements of O_3 concentrations from monitors in each selected urban study area; (2) O_3 concentrations that reflect changes in the distribution of O_3 air quality estimated to occur when an area just meets the existing or alternative O_3 standards under consideration; (3) population and demographic information, e.g., age, gender, etc.; (4) time-location activity pattern data; and (5) physiological data, e.g., body mass index, ventilation rates, life-stage development, etc. Basic outputs include numbers and percent of persons with O_3 exposures exceeding health-based benchmark levels and time-series of O_3 exposures and ventilation rates for individuals (for use in the lung function risk analysis). Details of the exposure modeling are discussed in Chapter 5.

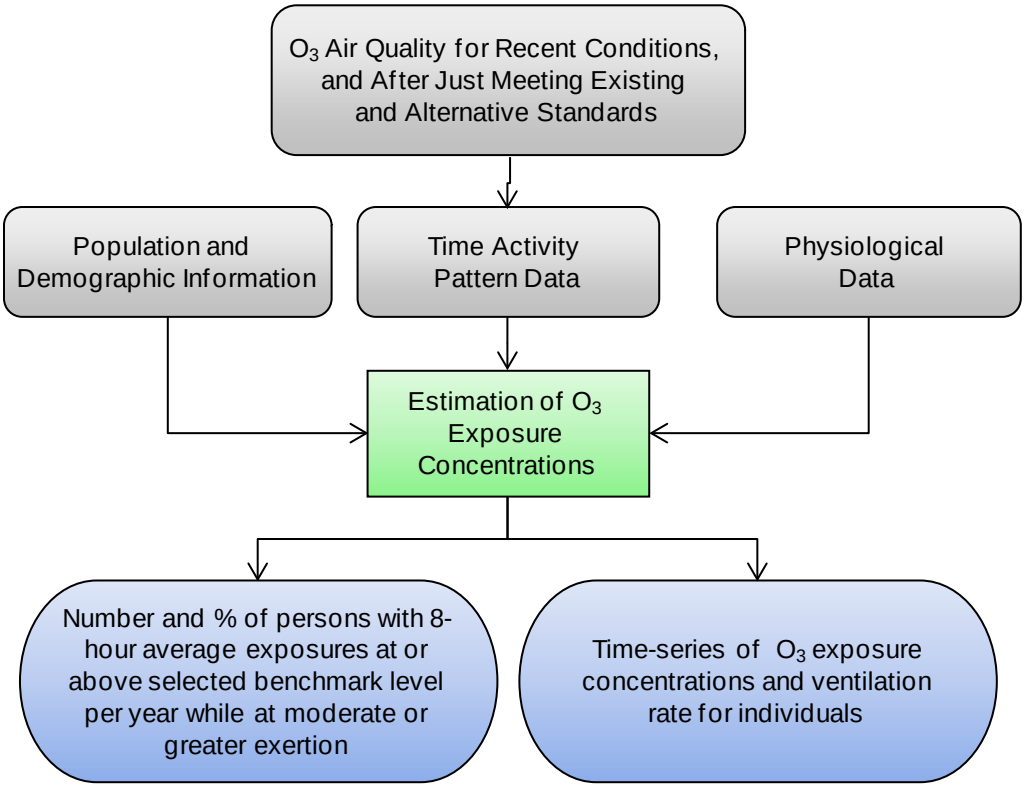


Figure 3-2 Conceptual Diagram for Population Exposure Assessment

The scope of the exposure assessment includes 15 urban case study areas.[7] These areas were selected to be generally representative of U.S. populations, geographic areas, climates, and different O_3 and co-pollutant levels, and they include all of the urban case study areas used in the epidemiology-based risk analysis (see Chapter 7). Three additional cities are included in the exposure modeling beyond those included in the epidemiology-based risk analysis. These cities are included to provide additional information on heterogeneity in exposure but could not be included in the epidemiology-based risk analysis because those analyses require additional information not available in the three additional cities. In addition to providing population exposures for estimation of lung function effects, the exposure modeling provides a characterization of urban air pollution exposure environments and activities resulting in the highest exposures.

Population exposure to ambient O_3 levels is evaluated using version 4.5 of the APEX model. The model and updated documentation are available at http://www.epa.gov/ttn/fera/apex_download.html. Exposures are estimated using recent ambient O_3 concentrations, based on 2006-2010 air quality data, and for O_3 concentrations resulting from simulations of just meeting the existing 8-hour O_3 standard and alternative O_3 standards, based on adjusting 2006-2010 air quality data. Because the O_3 standard is based on the 3-year average of the 4th highest daily 8-hour maximum, we simulate just meeting the standard for two periods, 2006-2008 and 2008-2010. Exposures are estimated for school-age children (ages 5 to 18), asthmatic school-age children, asthmatic adults (ages 19-95), and older persons (ages 65-95). This choice of population groups includes a strong emphasis on children, asthmatics, and persons $\geq$ 65 years old and reflects the finding of the last O_3 NAAQS review (EPA, 2007a) and the ISA (EPA, 2013, Chapter 8) that these are important at-risk groups.

In addition to estimating exposures exceeding health-based exposure benchmarks, the exposure estimates are used as an input to the portion of the health risk assessment that is based on exposure-response relationships derived from controlled human exposure studies. The exposure analysis also provides a characterization of populations with high exposures in terms of exposure environments and activities. In addition, the exposure analysis offers key observations based on the results of the APEX modeling, viewed in the context of factors such as averting behavior and key uncertainties and limitations of the model.

[7] These cities are Atlanta, GA; Baltimore, MD; Boston, MA; Chicago, IL; Cleveland, OH; Dallas, TX; Denver, CO; Detroit, MI; Houston, TX; Los Angeles, CA; New York, NY; Philadelphia, PA; Sacramento, CA; St. Louis, MO; and Washington, D.C. We also considered included Seattle; however, the available monitoring data was not sufficient to accurately characterize O_3 exposures for most populations in the Seattle area.

3.6 URBAN-SCALE LUNG FUNCTION RISK ANALYSES BASED ON APPLICATION OF RESULTS FROM CONTROLLED HUMAN EXPOSURE STUDIES

The major components in the lung function risk assessment are shown in Figure 3-3. Basic inputs to the analysis include 1) personal exposure to ambient O_3 derived from the exposure modeling described in Section 3.2.3., 2) data from controlled human exposure studies, used to construct exposure-response functions, 3) physiological data, including body mass index, age, etc., and 4) exercise levels, which determine breathing rates and affect dose. Basic outputs include the percentage of total population and sub-populations, e.g., children with asthma, with predicted lung function decrements (measured as decrements in forced expiratory volume in one second, or FEV_1), greater than or equal to 10, 15, and 20 percent, for recent O_3 levels and for O_3 levels after just meeting existing and alternative standards.

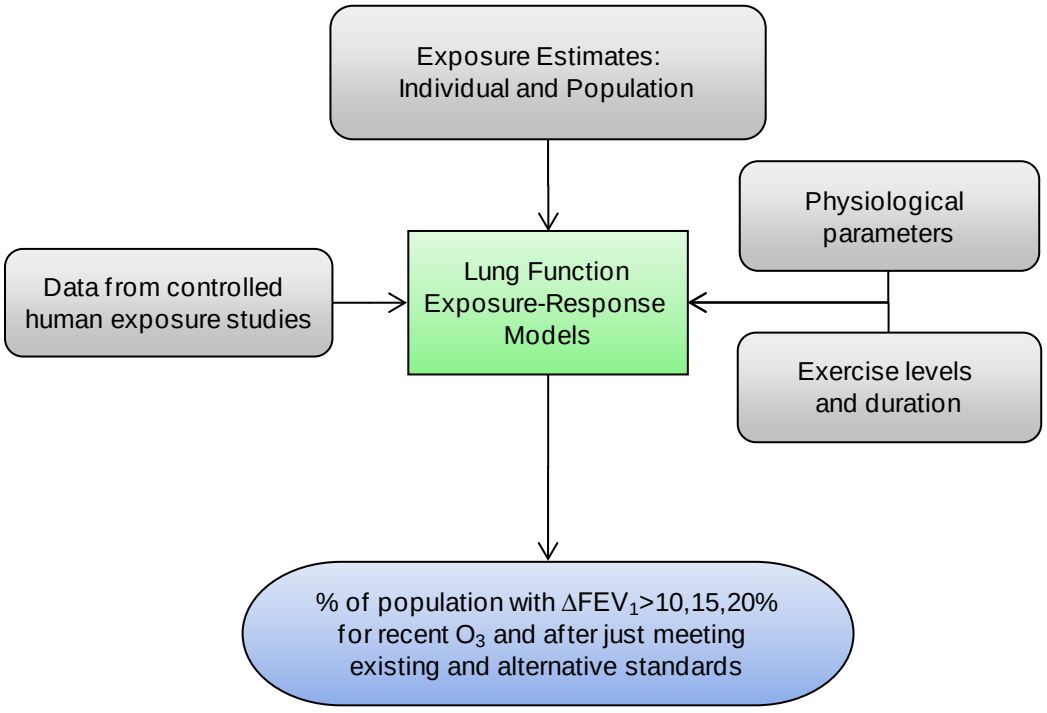


Figure 3-3 Conceptual Diagram of O_3 Lung Function Health Risk Assessment Based on Controlled Human Exposure Studies

Prior EPA risk assessments for O_3 have included risk estimates for lung function decrements and respiratory symptoms based on analysis of individual data from controlled human exposure studies. The current assessment applies probabilistic exposure-response relationships which are based on analyses of individual data that describe the relationship between a measure of personal exposure to O_3 and the measure(s) of lung function recorded in

the study. The current quantitative lung function risk assessment presents only a partial picture of the risks to public health associated with short-term O_3 exposures, as there are additional controlled human exposure studies that have evaluated cardiovascular and neurological outcomes due to O_3 exposure. However, these studies do not provide sufficient information with which to generate exposure-response functions and therefore are not suitable for quantitative risk assessment.

Modeling of risks of lung function decrements is based on application of results from controlled human exposure studies. These studies involve volunteer subjects who are exposed while engaged in different exercise regimens to specified levels of O_3 under controlled conditions for specified amounts of time. The responses measured in such studies have included measures of lung function, such as forced expiratory volume in one second (FEV_1), respiratory symptoms, airway hyper-responsiveness, and inflammation. The lung function risk assessment includes lung function decrement risk estimates, using FEV_1, for the adult population, school-age children (ages 5-18), and asthmatic school-age children (ages 5-18).

In addition to estimating lung function decrements for healthy adults that were the study groups in the controlled human exposure studies, this lung function risk assessment estimates lung function decrements (≥ 10, ≥ 15, and $\geq 20\%$ changes in FEV_1) in children 5 to <18 years old. The lung function estimates for children are based on applying data from young adult subjects (18-35 years old) to children aged 5-18. This is based on findings from other chamber studies and summer camp field studies documented in the 1996 O_3 Staff Paper (U.S. EPA, 1996a) and 1996 O_3 Criteria Document (U.S. EPA, 1996b), that lung function changes in healthy children are similar to those observed in healthy young adults exposed to O_3 under controlled chamber conditions.

Risk metrics estimated for lung function risk include the numbers of school-age children and other population groups experiencing one or more occurrences of a lung function decrement ≥ 10, ≥ 15, and $\geq 20\%$ in an O_3 season and the total number of occurrences of these lung function decrements in school-age children and active school-age children.

The risk assessment includes two different modeling approaches. The first approach employs a model that estimates FEV_1 responses for individuals associated with short-term exposures to O_3 (McDonnell et al., 2012). This model is based on the data from controlled human exposure studies included in the prior lung function risk assessment as well as additional data sets for different averaging times and breathing rates. These data were from 23 controlled human O_3 exposure studies that included exposure of 742 volunteers aged 18–35 years (see McDonnell et al., 2007 and McDonnell et al., 2012, for a description of these data). Outputs from this model include FEV_1 decrements for each simulated individual for each day, which can be used to calculate the population distribution of FEV_1 decrements, and the percent of the

population with FEV_1 decrements ≥ 10, ≥ 15, and ≥ 20% after just meeting existing and alternative standards.

In addition, we are applying the approach used in the last review and in the first draft of the REA, which employs a probabilistic population-level exposure-response function derived from the results of a number of controlled human exposure studies.

This modeling approach uses a smaller set of controlled human exposure studies and the population distribution of O_3 exposures to directly estimate the percent of the population with moderate levels of exertion with lung function decrements ≥ 10, ≥ 15, and ≥ 20%.

Controlled human exposure studies, carried out in laboratory settings, are generally not specific to any particular real-world location. A controlled human exposure studies-based risk assessment can therefore appropriately be carried out for any locations for which there are adequate air quality data on which to base the modeling of personal exposures. For this assessment, we have selected 15 urban case study areas (matching the areas used in the exposure analysis), representing a range of geographic areas, population demographics, and O_3 climatology. These 15 areas also include the 12 urban case study areas evaluated in the risk analyses based on concentration-response relationships developed from epidemiological or field studies.

In the controlled human exposure study based risk assessment, there are two broad sources of uncertainty to the risk estimates. One of the important sources of uncertainty is the estimation of the population distribution of individual time series of O_3 exposures and ventilation rates; these uncertainties are addressed as part of the exposure assessment. The second broad source of uncertainty in the risk calculation results from uncertainties in the lung function risk model. Sensitivity analyses are conducted to inform a qualitative discussion of these uncertainties.

3.7 URBAN CASE STUDY AREA EPIDEMIOLOGY-BASED RISK ASSESSMENT

The major components of the portion of the urban case study area health risk assessment based on data from epidemiological studies are illustrated in Figure 3-4. Basic inputs to this analysis include 1) measured O_3 concentrations for recent conditions and adjusted air quality representing O_3 concentrations after just meeting existing and alternative standards, 2) C-R functions derived from epidemiological studies evaluating associations between O_3 concentrations and mortality and morbidity endpoints and 3) population counts and baseline incidence rates for mortality and morbidity endpoints. Basic outputs for each urban area include estimates of O_3-attributable incidence and percent O_3-attributable incidence for selected mortality and morbidity endpoints and changes and percent changes in O_3-attributable incidence.

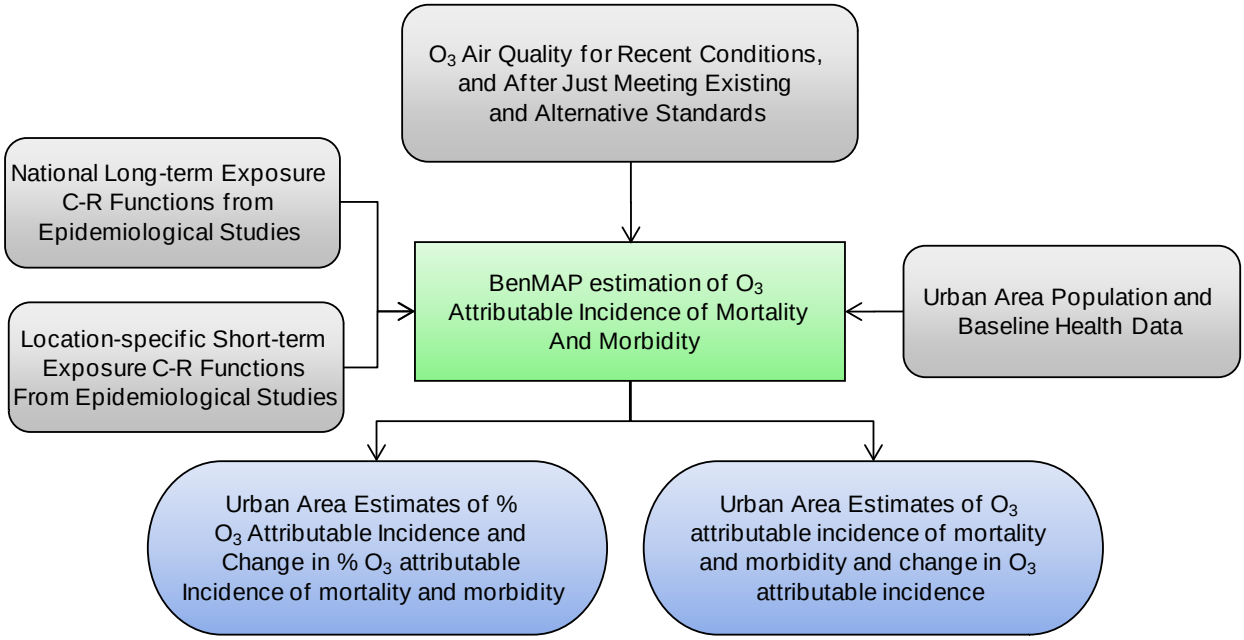


Figure 3-4 Conceptual Diagram of Urban Case Study Area Health Risk Assessment Based on Results of Epidemiology Studies

Epidemiological and field studies provide estimated concentration-response relationships based on data collected in real--world settings. Ambient O_3 concentrations used in these studies are typically spatial averages of monitor-specific measurements, using population-oriented monitors. Population health responses for O_3 have included population counts of school absences, emergency room visits, hospital admissions for respiratory and cardiac illness, respiratory symptoms, and premature mortality. Risk assessment based on epidemiological studies typically requires baseline incidence rates and population data for the risk assessment locations. To minimize uncertainties introduced by extrapolation, a risk assessment based on epidemiological studies can be performed for the locations in which the studies were carried out, rather than extrapolating results to urban areas where studies for a particular health endpoint have not been conducted.

The set of urban case study areas included in this portion of the risk assessment was chosen in order to provide population coverage and to capture the observed heterogeneity in O_3-related risk across selected urban study areas. In addition, locations had to have at least one epidemiological study conducted in order for the location to be included for a specific endpoint. This assessment also evaluates the mortality risk results for the selected urban areas within a broader national context to better characterize the nature, magnitude, extent, variability, and uncertainty of the public health impacts associated with O_3 exposures. This national-scale assessment is discussed in the next section.

We selected 2007 and 2009 as analysis years for the urban case study area risk analysis. These two years are the midpoint years in the two three-year periods 2006-2008 and 2008-2010. 2007 represents a year with generally higher O_3 concentrations, and 2009 represents a year with generally lower O_3 concentrations. Analyses for these two years will provide a good representation of the effects of baseline O_3 concentrations on the risk estimates.

This risk assessment is focused on health effect endpoints for which the weight of the evidence as assessed in the O_3 ISA supports the causal determination that a likely causal or causal relationship exits between a specific health effect category to be due to exposure to $O_{3.}$. The analysis includes estimates of mortality risk associated with short-term 8-hour maximum or 8-hour mean O_3 concentrations in all 12 urban case study areas, as well as risk of hospitalization for chronic obstructive pulmonary disease and pneumonia. In addition, the analysis includes analysis of hospitalizations for additional respiratory diseases in Los Angeles, New York City, and Detroit, due to limited availability of epidemiological studies covering these endpoints across the 12 urban areas. The analysis also evaluates risks of respiratory related emergency department visits in Atlanta and New York City and risks of respiratory symptoms in Boston, again based on availability of epidemiological studies in these locations. Table 3-1 summarizes the endpoints evaluated for each of the 12 urban case study areas.

Table 3-1 Short-term O_3 Exposure Health Endpoints Evaluated in Urban Case Study Areas

Urban Case Study Area	Mortality	COPD and Pneumonia hospitalizations	Other respiratory hospitalizations	Respiratory Related ED visits	Respiratory Symptoms
Atlanta, GA	X	X		X	
Baltimore, MD	X	X			
Boston, MA	X	X			X
Cleveland, OH	X	X			
Denver, CO	X	X			
Detroit, MI	X	X	X		
Houston, TX	X	X			
Los Angeles, CA	X	X	X		
New York, NY	X	X	X	X	
Philadelphia, PA	X	X			
Sacramento, CA	X	X			
St. Louis, MO	X	X			

This analysis will also estimate the respiratory mortality risks associated with longer-term exposures to O_3. This is supported by the O_3 ISA, which concluded that the evidence for long-term exposures to O_3 as likely to be causally related to respiratory effects, including respiratory mortality and morbidity, indicates causal relationship with. There is one national study of long-term exposures and respiratory mortality which provides a C-R function for use in the risk assessment. Several other studies have examined long-term exposures and cardiopulmonary mortality, but consistent with the ISA, we focused on respiratory mortality because of the additional supporting evidence related to long-term exposure and morbidity. Because the long-term exposure C-R function is based on comparing O_3 and mortality across urban areas, the same C-R function is applied in each of the 12 urban case study areas. The available epidemiological studies evaluating long term O_3 exposures and morbidity endpoints do not provide information that can be used to develop suitable C-R functions. As a result, we are not including quantitative risk estimates for morbidity associated with long-term exposures.

We have identified multiple options for specifying the concentration-response functions for particular health endpoints. This risk assessment provides an array of reasonable estimates for

each endpoint based on the available epidemiological evidence. This array of results provides a limited degree of information on the variability and uncertainty in risk due to differences in study designs, model specification, and analysis years, amongst other differences.

As part of the risk assessment, we address both uncertainty and variability. We provide a limited probabilistic characterization of uncertainty in the national-scale mortality risk estimates using the confidence intervals associated with effects estimates (obtained from epidemiological studies). However, this addresses only one source of uncertainty. For other sources of uncertainty, we include a number of sensitivity analyses to evaluate the impact of alternative approaches to simulating just meeting existing and alternative standards, alternative C-R functions, definitions of O_3 seasons to which C-R functions are applied, and definitions of urban areas to which the C-R functions are applied. In addition, we evaluate the impact in a subset of locations of using co-pollutant C-R functions. In the case of variability, we identify key sources of variability associated with O_3 risk (for both short-term and long-term exposure-related endpoints included in the risk assessment) and discuss the degree to which these sources of variability are reflected in the design of the risk assessment. Finally, we also include a comprehensive qualitative assessment of uncertainty and variability.

We also provide a representativeness analysis (see Chapter 8) designed to support the interpretation of risk estimates generated for the set of urban study areas included in the risk assessment. The representativeness analysis focuses on comparing the urban study areas to national-scale distributions for key O_3-risk related attributes (e.g., demographics including socioeconomic status, air-conditioning use, baseline incidence rates and ambient O_3 levels). The goal of these comparisons is to assess the degree to which the urban study areas provide coverage for different regions of the country as well as for areas likely to experience elevated O_3-related risk due to their specific mix of O_3-risk related attributes.

The risk assessment based on application of results of epidemiological studies is implemented using the environmental Benefits Mapping and Analysis Program Community Edition (BenMAP-CE) (U.S. EPA, 2013), EPA's GIS-based computer program for the estimation of health impacts associated with air pollution. BenMAP-CE draws upon a database of population, baseline incidence and effect estimates (regression coefficients) to automate the calculation of health impacts. EPA has traditionally relied upon the BenMAP program to estimate the health impacts avoided and economic benefits associated with adopting new air quality rules. It is also suitable for estimating risks associated with ambient concentrations of O_3 and changes in risk resulting from just meeting existing and alternative O_3 standards.

3.8 NATIONAL-SCALE MORTALITY RISK ASSESSMENT

The major components of the national-scale mortality risk assessment are shown in Figure 3-5. Basic inputs to this analysis are similar to those for the urban case study area epidemiology--based assessment and include 1) gridded O_3 concentrations over the continental U.S. for recent conditions, 2) C-R functions relating long-term and short-term exposures to O_3 to mortality, and 3) population and baseline mortality rates. Basic outputs include county and national estimates of incidence and percent of mortality attributable to O_3.

The national-scale mortality risk assessment serves two primary purposes. First, it serves as part of the representativeness analysis discussed above, providing an assessment of the degree to which the urban study areas included in the risk assessment provide coverage for areas of the country expected to experience elevated mortality rates due to O_3-exposure. Second, it provides a broader perspective on the distribution of risks associated with recent O_3 concentrations throughout the U.S., and provides a more complete understanding of the overall public health burden associated with O_3.[8] We note that a national-scale assessment such as this was completed for the risk assessment supporting the latest PM NAAQS review (US EPA, 2010) with the results of the analysis being used to support an assessment of the representativeness of the urban study areas assessed in the PM NAAQS risk assessment, as described here for O_3.

[8] In the previous O_3 NAAQS review, CASAC commented that "There is an underestimation of the affected population when one considers only twelve urban "Metropolitan Statistical Areas" (MSAs). The CASAC acknowledges that EPA may have intended to illustrate a range of impacts rather than be comprehensive in their analyses. However, it must be recognized that O_3 is a regional pollutant that will affect people living outside these 12 MSAs, as well as inside and outside other urban areas." Inclusion of the national-scale mortality risk assessment partially addresses this concern by providing a broader characterization of risk for an important O_3 health endpoint.

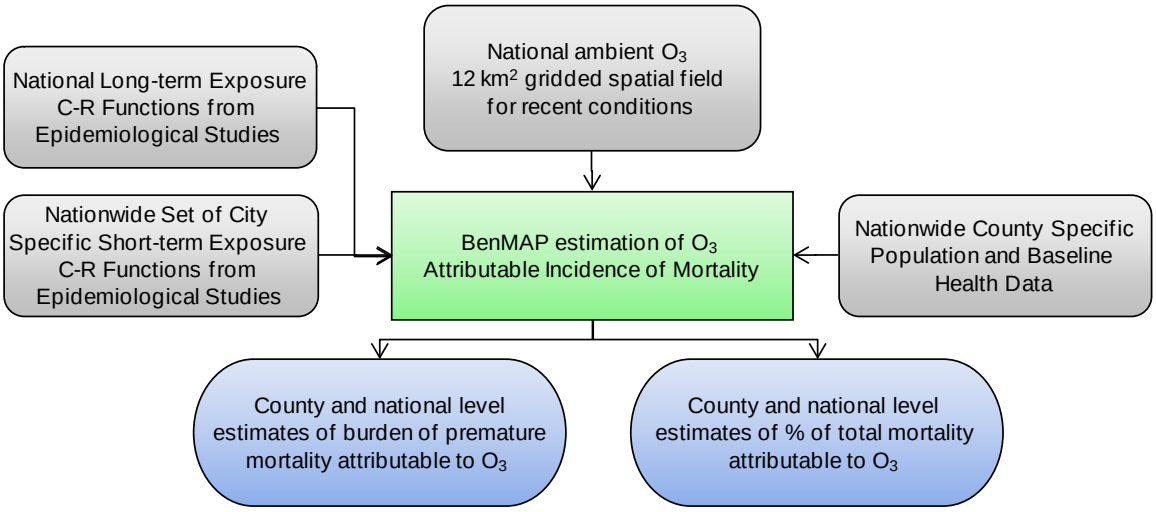


Figure 3-5 Conceptual Diagram of National O_3 Mortality Risk Assessment Based on Results of Epidemiology Studies

The national-scale risk assessment focuses on mortality only, due to the availability of large multi-city epidemiology studies for short-term mortality and the availability of a long-term mortality study which provides information to develop a suitable C-R function. As noted in the discussion of the urban case study area analyses, the available epidemiological studies evaluating long-term O_3 exposures and morbidity endpoints do not provide information that can be used to develop suitable C-R functions. In the case of short-term morbidity endpoints, the available epidemiological studies are generally located in only a few urban areas and, even in the case of the multi-city hospitalization studies, cover only a small fraction of the urban areas in the U.S. In addition, baseline mortality rates are available for every county in the U.S., while baseline hospitalization rates are available in only a small subset of counties. For these reasons, the national-scale risk assessment includes only mortality associated with short- and long-term O_3 exposures.

We provide a limited probabilistic characterization of uncertainty in the national-scale mortality risk estimates using the confidence intervals associated with effects estimates (obtained from epidemiological studies). However, this addresses only one source of uncertainty. To address some other key potential sources of uncertainty in the national assessment, we conduct sensitivity analyses. Risk estimates are provided for two alternative C-R functions for short-term exposure, reflecting two multi-city epidemiological studies. For short-term exposure-related mortality, the assessment provides several estimates of national mortality risk, including a full national-scale estimate including all counties in the continental U.S., and an analysis restricted to the set of urban areas included in the time-series studies that provide the effect estimates. We

have greater confidence in the analysis based on the large urban areas included in the epidemiological studies, but the information from the full analysis of all counties is useful to gain understanding of the potential magnitude of risk in less urbanized areas. In addition, the national-scale mortality risk assessment evaluates the sensitivity of the nationwide estimates to assumptions about the transferability of effect estimates from the cities included in the underlying epidemiological studies to other cities in the U.S. Finally, the assessment includes a sensitivity analysis evaluating the use of regional priors city--rather than using a national prior in developing the city specific Bayesian adjusted effect estimates.[9] These sensitivity analyses are described in detail in Chapter 8.

The national-scale risk assessment is conducted only for recent O_3 conditions. We do not attempt to simulate nationwide O_3 concentrations that would result from just meeting the existing or alternative O_3 standards everywhere in the U.S. Such a simulation would require detailed modeling of attainment strategies in all potential non-attainment areas and would need to take into account the interdependence of O_3 concentrations across urban areas. This type of analysis is beyond the scope of this risk assessment. Analyses of nationwide attainment are included as part of the Regulatory Impact Analyses that accompany proposed and final rulemaking packages and will likely be included in the rulemaking portion of this review.

3.9 PRESENTATION OF EXPOSURE AND RISK ESTIMATES TO INFORM THE O_3 NAAQS POLICY ASSESSMENT

We present exposure estimates in three ways: person-occurrences, number, and percent of persons in different populations (e.g., adults, all school-age children, asthmatic school-age children, outdoor workers) with at least one 8-hour average exposure at or above benchmark levels of 60 ppb, 70 ppb, and 80 ppb. In addition, the same types of results are shown for persons with multiple exposures at or above the benchmark levels. The results are presented in summary tables and graphics, while detailed tables of results are provided in an appendix. The focus in the presentation of results is on exposures occurring after simulating just meeting the existing standard and on the change in number and percent of exposures between meeting the existing standard and meeting alternative standards. Results are presented for the five modeled years, for all 15 urban case study areas.

Quantitative risk estimates from the analyses based on application of controlled human exposure studies are presented for the two different risk models. For each model, we provide

[9] In multi-city Bayesian analyses, it is necessary to specify initial values or "priors" which are then "updated" using information from the individual city specific estimates. These priors are generally a mean value across all of the cities, in this case, cities in regions or cities across the nation.

estimates of the percent of different populations (adults, all children, children with asthma) with lung function decrements greater than or equal to 10, 15, and 20 percent. As with exposure, the focus in the presentation of results is on risk occurring after simulating just meeting the existing standards and on the change in risk occurring between meeting the existing standard and meeting alternative standards.

Results from the epidemiology-based risk assessment are presented in two ways: (1) total (absolute) health effects incidence for recent air quality and simulations of air quality just meeting the existing and alternative standards under consideration and (2) risk reduction estimates, reflecting the change in the distribution of O_3 between scenarios of just meeting the existing standard and just meeting alternative standards. In addition, risks are presented as the percent of baseline incidence, and risks per 100,000 population, to allow for comparisons between urban areas with very different population sizes. We include risk modeled across the full distribution of O_3 concentrations, as well as core risk estimates for O_3 concentrations down to 0 ppb.

We present an array of risk estimates in order to provide additional context for understanding the potential impact of uncertainty on the risk estimates. For core estimates and sensitivity analyses, we provide the statistical confidence intervals, demonstrating the relative precision of estimates. The graphical presentation of sensitivity analyses focuses on the differences from the core estimates in terms of risk per 100,000 population.

The results of the representativeness analysis are presented using cumulative probability plots (for the national-level distribution of O_3 risk-related parameters) with the locations where the individual urban study areas fall within those distributions noted in the plots using vertical lines. Similar types of plots are used to present the distribution of national-scale mortality estimates based on the national-scale risk assessment, showing the location of the urban case study areas within the overall national distribution.

Chapter 9 of this risk and exposure assessment provides a synthesis of the results from the four assessments (urban case study area exposure, urban case study area lung function risk, urban case study area epidemiology-based risk, and national mortality risk). Chapter 9 focuses on comparing patterns of results across locations, years, and alternative standards. Chapter 9 also provides perspective on the overall degree of confidence of the analyses and the representativeness of the set of results in characterizing patterns of exposure and risk and patterns of changes in exposure and risk from just meeting alternative standards relative to just meeting the existing standards.

3.10 REFERENCES

Chen, J., R. Zhao and Z. Li. 2004. Voronoi-based k-order neighbor relations for spatial analysis. ISPRS J Photogrammetry Remote Sensing, 59(1-2), 60-72.

Folinsbee, L.J.; W.F. McDonnell and D.H. Horstman. 1988. Pulmonary function and symptom responses after 6.6-hour exposure to 0.12 ppm O_3 with moderate exercise. *JAPCA.* 38: 28-35.

Gold, C. 1997. Voronoi methods in GIS. In: Algorithmic Foundation of Geographic Information Systems (va Kereveld M.; J. Nievergelt; T. Roos; P. Widmayer, eds). Lecture Notes in Computer Science, Vol 1340. Berlin: Springer-Verlag, 21-35.

Henderson, R. 2007. Clean Air Scientific Advisory Committee's (CASAC) Review of the Agency's Final O_3 Staff Paper. EPA-CASAC-07-002. March 26.

Horstman, D.H.; L.J. Folinsbee; P.J. Ives; S. Abdul-Salaam and W.F. McDonnell. 1990. O_3 concentration and pulmonary response relationships for 6.6-hr exposures with five hrs of moderate exercise to 0.08, 0.10, and 0.12 ppm. *Am Rev Respir Dis*. 142:1158-1163.

Kim, Chong S.; N.E. Alexis; A.G. Rappold; H. Kehrl; J. Milan; J.C. Hazucha; J.C. Lay; M.T. Schmitt; C. Martin; R.B. Devlin; D.B. Peden and D. Diaz-Sanchez. 2011. "Lung Function and Inflammatory Responses in Healthy Young Adults Exposed to 0.06 ppm O_3 for 6.6 Hours." American Journal of Respiratory and Critical Care Medicine 183, no. 9 (2011): 1215-1221.

Langstaff, J.E. 2007. OAQPS Staff Memorandum to O_3 NAAQS Review Docket (EPA HQ-OAR-2005-0172). Subject: Analysis of Uncertainty in O_3 Population Exposure Modeling. [January 31, 2007]. <http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_cr_td.html>.

McDonnell, W.F., et al. 1991. "Respiratory response of humans exposed to low levels of O_3 for 6.6 hours." *American Review of Respiratory Disease*, 147:804-810.

McDonnell, W.F.; P.W. Stewart and M.V. Smith. 2007. "The temporal dynamics of O_3 –induced FEV_1 changes in humans: an exposure-response model." *Inhalation Toxicology*, 19:483–494.

McDonnell W.F.; P.W. Stewart and M.V. Smith. 2010. "Prediction of O_3 -induced lung function responses in humans." *Inhalation Toxicology*, 22(2):160-8.

McDonnell, W.F.; P.W. Stewart; M.V. Smith; C.S. Kim and E.S. Schelegle. 2012. Prediction of lung function response for populations exposed to a wide range of ozone conditions. Inhalation Toxicology 24: 619-633.

Samet, J. 2009. Consultation on EPA's Draft Integrated Review Plan for the National Ambient Air Quality Standards for Particulate Matter. (EPA document number EPA-CASAC-10-004. December 3).

Samet, J. 2011. Consultation on EPA's O_3 National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment (April 2011) and O_3 National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment (April 2011). EPA-CASAC-11-008. June 21.

Schelegle, E.S.; C.A. Morales; W.F. Walby; S. Marion and R.P. Allen. 2009. “6.6-Hour Inhalation of O_3 Concentrations from 60 to 87 Parts per Billion in Healthy Humans.” American Journal of Respiratory and Critical Care Medicine 180 (2009): 265-272.

Simon H.; K.R. Baker; F. Akhtar; S.L. Napelenok; N. Possiel; B. Wells and B. Timin. 2013. A Direct Sensitivity Approach to Predict Hourly Ozone Resulting from Compliance with the National Ambient Air Quality Standard. Environmental Science and Technology. 2013 Feb. 12. [Epub ahead of print]

U.S. Environmental Protection Agency. 1996a. "Review of National Ambient Air Quality Standards for O_3 : Assessment of Scientific and Technical Information - OAQPS Staff Paper." (EPA document number EPA/452/R-96-007). Research Triangle Park, NC: Office of Air Quality Planning and Standards. Available from: NTIS, Springfield, VA; PB96–203435. < http://www.epa.gov/ttn/naaqs/standards/O_3 /s_O_3_pr.html>.

U.S. EPA. 1996b. "Air Quality Criteria for O_3 and Related Photochemical Oxidants." (EPA document number EPA/600/P-93/004aF-cF). Research Triangle Park, NC: Office of Research and Development, National Center for Environmental Assessment.

< http://cfpub.epa.gov/ ncea/cfm/recordisplay.cfm?deid=2831>.

U.S. EPA. 2007a. O_3 *Population Exposure Analysis for Selected Urban Areas*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-07-010, July).

U.S. EPA. 2007b. O_3 *Health Risk Assessment for Selected Urban Areas*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA 452/R-07-009, July).

U.S. EPA. 2007c. "Review of the National Ambient Air Quality Standards for O_3: Policy Assessment of Scientific and Technical Information. OAQPS Staff Paper." Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA report number EPA-452/R-07-007a. < http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_cr_sp.html >.

U.S. EPA. 2008a." National Air Quality: Status and Trends Through 2007." Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document EPA-454/R-08-006, November). <http://www.epa.gov/airtrends/2008/index.html>.

U.S. EPA. 2008b. "EPA’s Report on the Environment." Washington, DC. (EPA document number EPA/600/R-07/045F, May 2008). <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=190806>.

U.S. EPA. 2008c. "Risk and Exposure Assessment to Support the Review of the NO2 Primary National Ambient Air Quality Standard." (EPA document number EPA-452/R-08-008a, November 2008). <http://www.epa.gov/ttn/naaqs/standards/nox/data/20081121_NO2_REA_final.pdf>.

U.S. EPA. 2009a. *Integrated Review Plan for the* O_3 *National Ambient Air Quality Standards Review: External Review Draft*. Research Triangle Park, NC: Environmental Media Assessment Group, National Center for Environmental Assessment and Health and Environmental Impacts Division, Office of Air Quality Planning and Standards. (EPA document number EPA 452/D-09-001, September).

U.S. EPA. 2009b. "Risk and Exposure Assessment to Support the Review of the SO2 Primary National Ambient Air Quality Standard." (EPA document number EPA-452/R-09-007, August 2009). <http://www.epa.gov/ttn/naaqs/standards/so2/data/200908SO2REAFinalReport.pdf.>

U.S. EPA. 2010a. "Quantitative Risk and Exposure Assessment for Carbon Monoxide – Amended." EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-10-009, July). <http://www.epa.gov/ttn/naaqs/standards/co/data/CO-REA-Amended-July2010.pdf>.

U.S. EPA. 2010b. "Quantitative Health Risk Assessment for Particulate Matter." Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-10-005). <http://www.epa.gov/ttn/naaqs/standards/pm/data/PM_RA_FINAL_June_2010.pdf>.

U.S. EPA. 2011a. *Integrated Review Plan for the* O_3 *National Ambient Air Quality Standards*. Research Triangle Park, NC: National Center for Environmental Assessment, Office of Research and Development and Office of Air Quality Planning and Standards, Office of Air and Radiation. (EPA document number EPA 452/R-11-006, April).

U.S. EPA. 2011b. O_3 *National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/P-11-001, April).

U.S. EPA. 2011c. O_3 *National Ambient Air Quality Standards: Scope and Methods Plan for Welfare Risk and Exposure Assessment*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/P-11-002, April).

U.S. EPA. 2013. "Environmental Benefits Mapping Analysis Program Community Edition (BenMAP-CE v1.0)." Research Triangle Park, NC: Health and Environmental Impacts Division, Office of Air Quality Planning and Standards. <http://www.epa.gov/air/benmap/>.

Wegman, L. 2012. Updates to information presented in the Scope and Methods Plans for the O_3 NAAQS Health and Welfare Risk and Exposure Assessments. Memorandum from Lydia Wegman, Division Director, Health and Environmental Impacts Division, Office of Air Quality Planning and Standards, Office of Air and Radiation, US EPA to Holly Stallworth, Designated Federal Officer, Clean Air Scientific Advisory Committee, US EPA Science Advisory Board Staff Office. May 2, 2012.

World Health Organization. 2008. Harmonization Project Document No. 6. Part 1: Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment. <http://www.who.int/ipcs/methods/harmonization/areas/exposure/en/.>.

4 AIR QUALITY CONSIDERATIONS

4.1 INTRODUCTION

Air quality information is used in Chapters 5-8 to assess risk and exposure resulting from recent O_3 concentrations, as well as to estimate the relative change in risk and exposure that could result from just meeting the existing O_3 standard of 75 ppb and the potential alternative standard levels of 70 ppb, 65 ppb, and 60 ppb[1]. The same air quality data are used to examine fifteen[2] urban case study areas in the population exposure analyses discussed in Chapter 5 and the lung function risk assessment based on application of results from clinical studies discussed in Chapter 6: Atlanta, GA; Baltimore, MD; Boston, MA; Chicago, IL; Cleveland, OH; Dallas, TX; Denver, CO; Detroit, MI; Houston, TX; Los Angeles, CA; New York, NY; Philadelphia, PA; Sacramento, CA; St. Louis, MO; and Washington, DC. The epidemiology-based risk assessment discussed in Chapter 7 examines twelve[3] of the fifteen urban case study areas evaluated in the population exposure analyses. Finally, Chapter 8 includes an assessment of the national-scale O_3 mortality risk burden associated with recent O_3 concentrations, and characterizes the representativeness of the 15 urban case study areas compared to the rest of the U.S. This chapter describes the air quality information developed for these analyses, providing an overview of monitoring data and air quality (section 4.2) and an overview of air quality inputs to the risk and exposure assessments (section 4.3).

4.2 OVERVIEW OF O_3 MONITORING AND AIR QUALITY DATA

To determine whether or not the NAAQS have been met at an ambient O_3 monitoring site, a statistic commonly referred to as a "design value" must be calculated based on 3 consecutive years of data collected from that site. The form of the existing O_3 NAAQS design value statistic is the 3-year average of the annual 4th highest daily maximum 8-hour O_3 concentration in parts per billion (ppb), with decimal digits truncated. The existing primary and secondary O_3 NAAQS are met at an ambient monitoring site when the design value is less than

[1] For a subset of urban areas and analyses, the REA evaluates a standard level of 55 ppb, consistent with recommendations from CASAC to also give consideration to evaluating a level somewhat below 60 ppb.

[2] In the first draft REA, we proposed to include 16 urban areas in the second draft REA. However, further analysis of the air quality information available for Seattle, WA has prompted us to not include that city. This decision and supporting analysis are discussed in more detail in Appendix 4-E.

[3] These cities are Atlanta, GA; Baltimore, MD; Boston, MA; Cleveland, OH; Denver, CO; Detroit, MI; Houston, TX; Los Angeles, CA; New York, NY; Philadelphia, PA; Sacramento, CA; and St. Louis, MO.

or equal to 75 ppb.[4] In counties or other geographic areas with multiple monitors, the area-wide design value is defined as the design value at the highest individual monitoring site, and the area is said to have met the NAAQS if all monitors in the area are meeting the NAAQS.

Air quality monitoring data from 1,468 U.S. ambient O_3 monitoring sites were retrieved by EPA staff for use in the risk and exposure assessments. The initial dataset consisted of hourly O_3 concentrations in ppb collected between 1/1/2006 and 12/31/2010 from these monitors. Data for nearly 1,400 of these monitors were extracted from EPA's Air Quality System (AQS) database[5], while the remaining data came from EPA's Clean Air Status and Trends Network CASTNET) database which consists of primarily rural monitoring sites. While CASTNET monitors did not begin reporting regulatory data to AQS until 2011, it is generally agreed that data collected from these monitors prior to 2011 is of comparable quality to the data reported to AQS.

These data were split into two design value periods, 2006-2008 and 2008-2010, and all subsequent analyses based on these data were conducted independently for these two periods. Observations flagged in AQS as having been affected by exceptional events were included the initial dataset, but were not used in design value calculations in accordance with EPA's exceptional events policy. Missing data intervals of 1 or 2 hours in the initial dataset were filled in using linear interpolation. These short gaps often occur at regular intervals in the ambient data due to an EPA requirement for monitoring agencies to perform routine quality control checks on their O_3 monitors. Quality control checks are typically performed between midnight and 6:00 AM when O_3 concentrations are low. Missing data intervals of 3 hours or more were not replaced. Interpolated data values were not used in design value calculations.

Figures 4-1 and 4-2 show the design values for the existing O_3 NAAQS for all regulatory monitoring sites in the U.S. for the 2006-2008 and 2008-2010 periods, respectively. In general, O_3 design values were lower in 2008-2010 than in 2006-2008, especially in the Eastern U.S. There were 518 O_3 monitors in the U.S. with design values above the existing standard in 2006-2008, compared to only 179 in 2008-2010.

[4] For more details on the data handling procedures used to calculate design values for the current O_3 NAAQS, see 40 CFR Part 50, Appendix P.

[5] EPA's Air Quality System (AQS) database is a national repository for many types of air quality and related monitoring data. AQS contains monitoring data for the six criteria pollutants dating back to the 1970's, as well as more recent additions such as PM2.5 speciation, air toxics, and meteorology data. At present, AQS receives hourly O_3 monitoring data collected from nearly 1,400 monitors operated by over 100 state, local, and tribal air quality monitoring agencies.

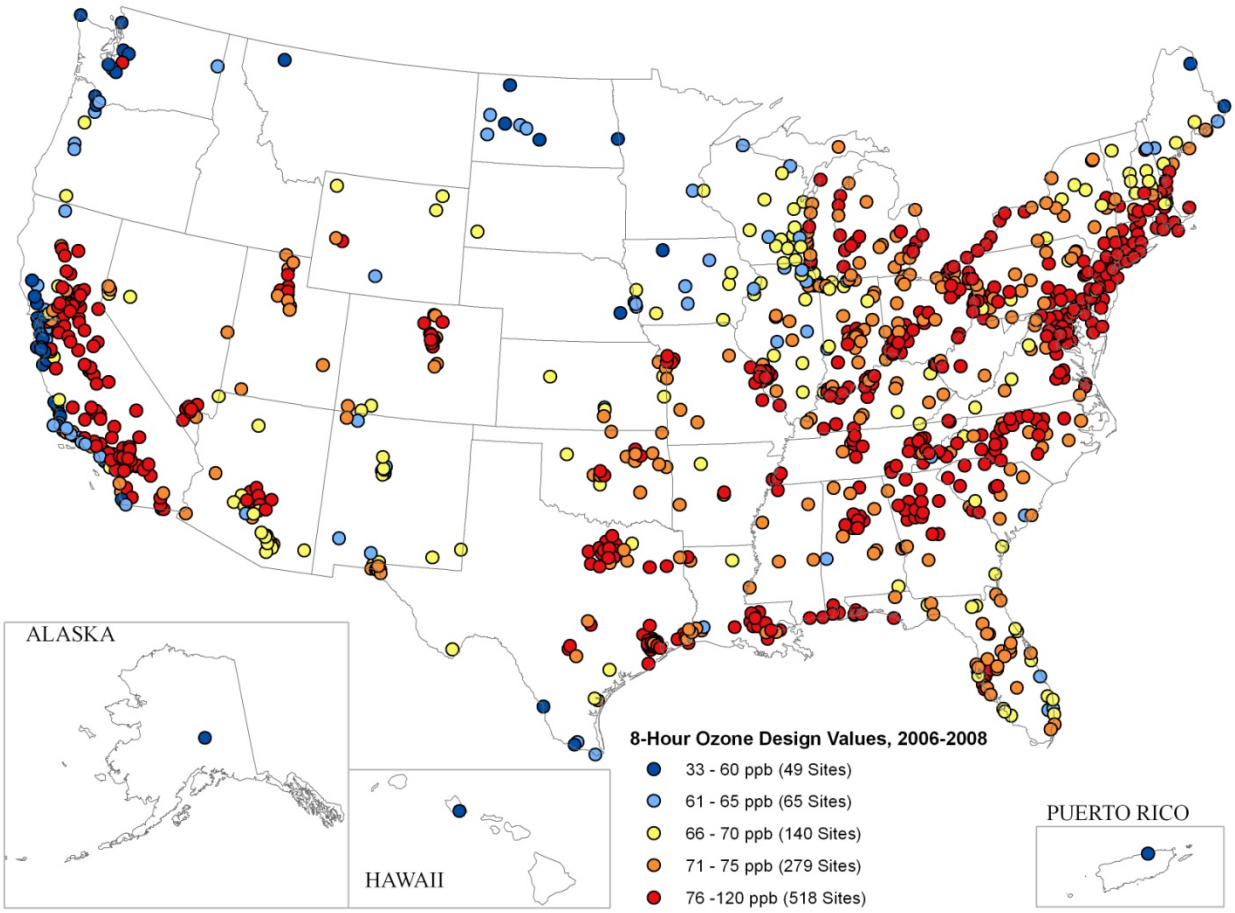


Figure 4-1 Map of Monitored 8-hour O_3 Design Values for the 2006-2008 Period

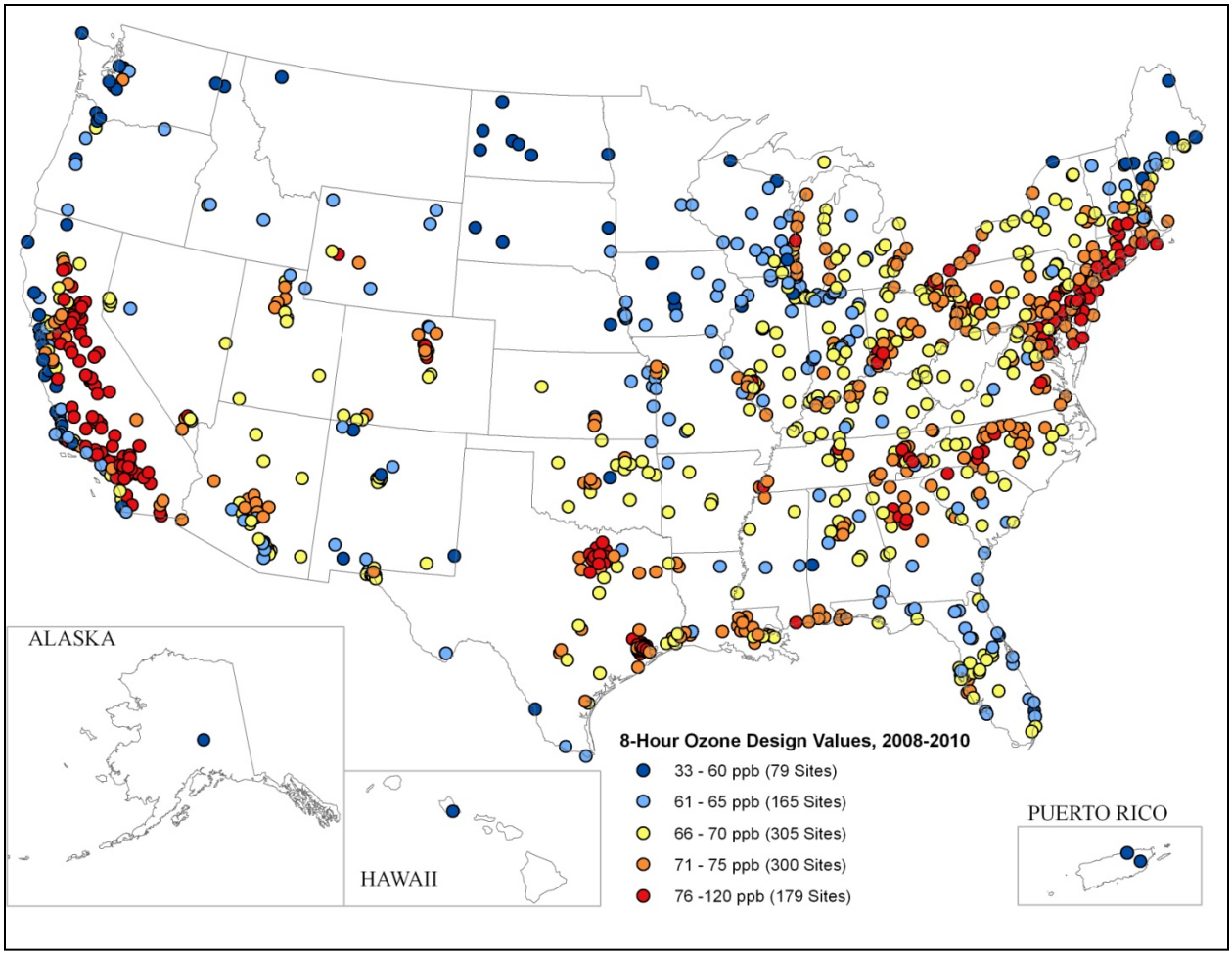


Figure 4-2 Map of Monitored 8-hour O_3 Design Values for the 2008-2010 Period

4.3 OVERVIEW OF URBAN-SCALE AIR QUALITY INPUTS TO RISK AND EXPOSURE ASSESSMENTS

The air quality information input into the urban-scale risk and exposure assessments includes both recent air quality data from the years 2006-2010, as well as air quality data adjusted to reflect just meeting the existing and potential alternative standard levels. In this section, we summarize these air quality inputs and discuss the methodology used to adjust air quality to meet the existing and potential alterative standards.

Figure 4-3 presents a flowchart of air quality data processing steps for the urban-scale analyses. The rest of section 4.3.1 will provide more details on each step depicted in the flow diagram. Additional information is provided in Appendices 4-A, 4-B and 4-D.

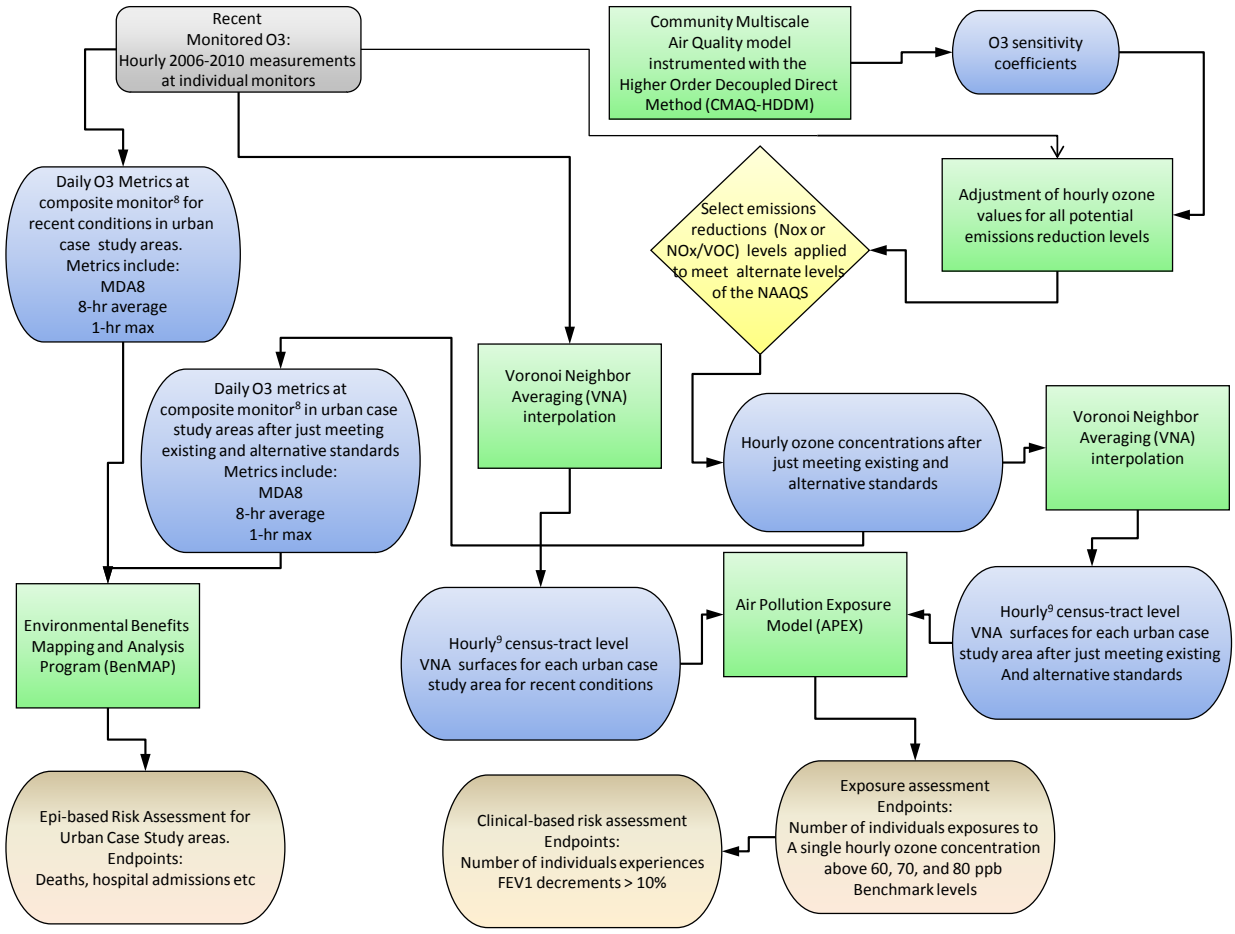


Figure 4-3 Flowchart of Air Quality Data Processing for Different Parts of the Urban-scale Risk and Exposure Assessments

4.3.1 Urban Case Study Areas

4.3.1.1 Exposure Modeling and Controlled Human Study Based Lung Function Risk Assessment

The 15 urban case study areas in the exposure modeling and lung function risk assessments covered a large spatial extent, with boundaries generally similar to those covered by the respective Combined Statistical Areas (CSA) defined by the U.S. Census Bureau. Table 4-1 gives some basic information about the 15 urban case study areas in the exposure assessment, including the number of ambient monitoring sites, the required O_3 monitoring season, and the 2006-2008 and 2008-2010 design values for each study area. All 15 of the urban case study areas had 8-hour O_3 design values above the existing standard in 2006-2008, while 13 urban areas had

[6] Composite monitors do not always include the highest design value monitor in every urban area.

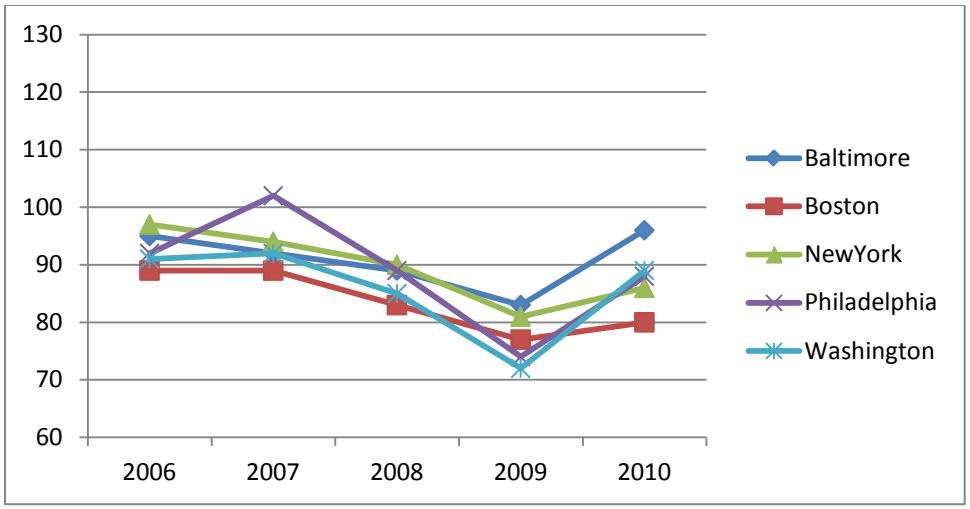
[7] 4800 VNA surfaces were created for each urban area/alternative standard level pair: 24 hrs × 365 days × 5 years.

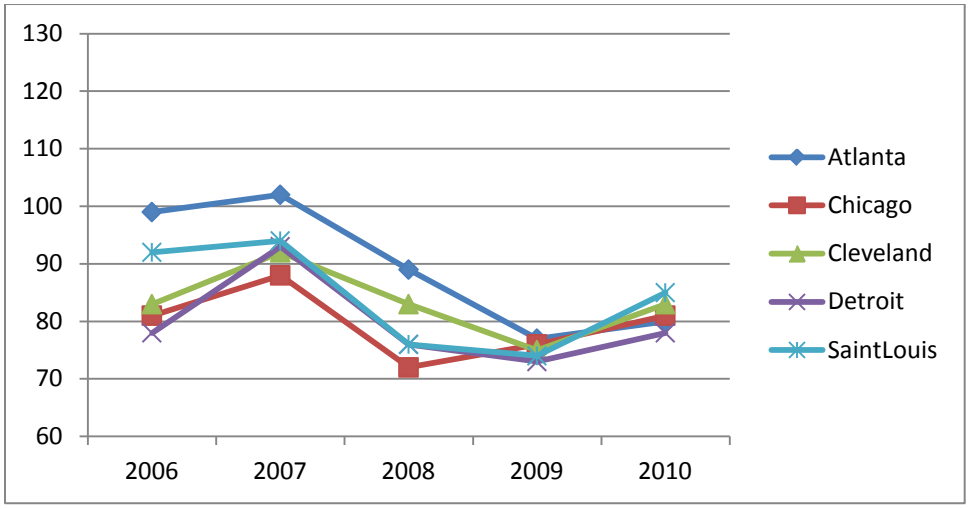
8-hour O_3 design values above the existing standard in 2008-2010. Chicago (74 ppb) and Detroit (75 ppb) had design values meeting the existing standard during the 2008-2010 period. The design values in the 15 urban areas decreased by an average of 6 ppb between 2006-2008 and 2008-2010, ranging from no change in Sacramento to a decrease of 15 ppb in Atlanta.

Table 4-1 Monitor and Area Information for the 15 Urban Case Study Areas in the Exposure Modeling and Clinical Study Based Risk Assessment

Area Name	# of Counties	# of O_3 Monitors	Population (2010)	Required O_3 Monitoring Season	2006-2008 DV (ppb)	2008-2010 DV (ppb)
Atlanta	33	13	5,618,431	March - October	95	80
Baltimore	7	7	2,710,489	April - October	91	89
Boston	10	14	5,723,468	April - September	83	77
Chicago	16	26	9,686,021	April - October	78	74
Cleveland	8	13	2,881,937	April - October	82	77
Dallas	11	20	6,366,542	January - December	89	86
Denver	13	26	3,390,504	March - September	86	77
Detroit	9	12	5,218,852	April - September	81	75
Houston	10	22	5,946,800	January - December	91	85
Los Angeles	5	54	17,877,006	January - December	119	112
New York	27	31	21,056,173	April - October	90	84
Philadelphia	15	19	7,070,622	April - October	92	83
Sacramento	7	26	2,755,972	January - December	102	102
St. Louis	17	17	2,837,592	April - October	85	77
Washington	26	22	5,838,518	April - October	87	81

Since O_3 design values are based on the annual 4th highest 8-hour daily maximum O_3 concentrations from 3-consecutive years, it is useful to look at inter-annual variability. In general, the annual 4th highest 8-hour O_3 concentrations decreased in 11 of the 15 urban areas from 2006 to 2010, while remaining relatively constant in the other 4 areas (Figure 4-4). The average decrease in the annual 4th highest daily maximum concentration from 2006 to 2010 was 8 ppb. However, there was significant year-to-year variability, and some areas showed increases in some years relative to 2006, even though the 2010 values were generally lower.





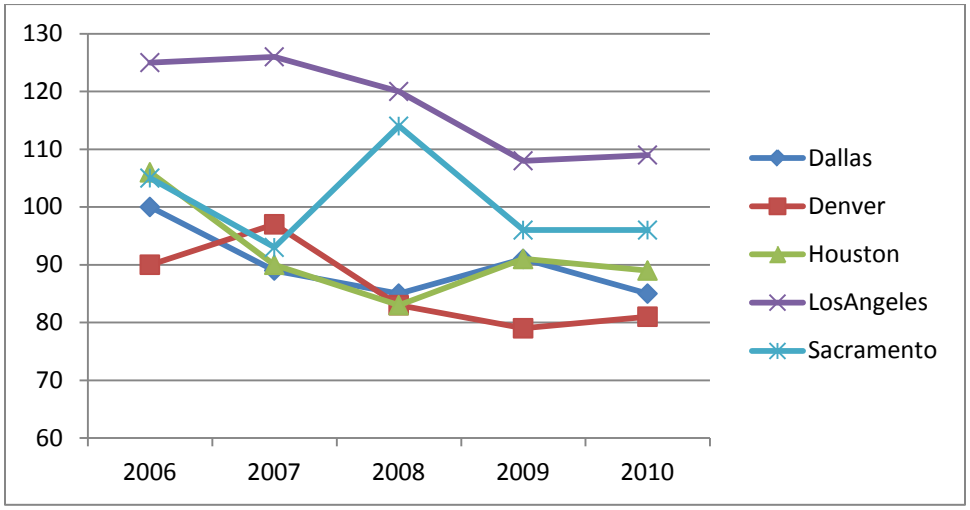


Figure 4-4 Trends in Annual 4th Highest 8-hour Daily Maximum O_3 Concentrations in ppb for the 15 Urban Case Study Areas for 2006-2010. Urban areas are grouped into 3 regions: Eastern (top), Central (middle), and Western (bottom).

4.3.1.2 Epidemiology Based Risk Assessment

Table 4-2 gives some basic information on the 12 urban case study areas in the epidemiology-based risk assessment for each set of area boundaries. The spatial extent of each urban case study area was based on the respective Core Based Statistical Area (CBSA)[8]. The CBSAs were generally smaller than the study areas used in the exposure modeling and clinical study based risk assessments, except for Baltimore and Houston, where the two study areas were identical. The rationales for the definitions of the spatial areas used in each type of analysis are provided in the corresponding chapters. The final two columns in Table 4-2 show the annual 4th highest daily maximum 8-hour O_3 concentration in ppb for the monitors within each urban case study area in 2007 and 2009.

It should be noted that the CBSA boundaries used for the urban case study areas in this assessment are different than those used in the 1st draft of the REA, where the study areas were derived from the Zanobetti and Schwartz (2008) study. The change to the CBSA boundaries was intended to capture a larger portion of the urban area populations by including some surrounding suburban counties, rather than focusing strictly on the urban population centers. Two sensitivity analyses were conducted to determine the effect of changing the spatial extent of the urban case study areas on the epidemiology-based risk estimates. These sensitivity analyses are presented in Chapter 7, and a summary of the two alternative sets of boundaries for the 12 urban case study areas are provided in Appendix 4-A.

Since O_3 is not directly emitted but is formed through photochemical reactions, precursor emissions may continue to react and form O_3 downwind of emissions sources, thus the highest O_3 concentrations are often found downwind of the highest concentrations of precursor emissions near the urban population center. There were some instances where the highest monitor occurred outside of the CBSA, but within the exposure area, which was designed to always include the monitor associated with the area-wide design value. For example, in Los Angeles, the CBSA includes Los Angeles and Orange counties, but the highest O_3 concentrations are typically measured further downwind in Riverside and San Bernardino counties. Thus, the values reported in Table 4-2 may not match the values shown in Figure 4-4.

[8] Core Based Statistical Areas (CBSAs) are used by the Office of Management and Budget (OMB) to group U.S. counties into urbanized areas. These groupings are updated by OMB every 5 years. The CBSAs used in the epidemiology based risk assessment are based on the OMB deliniations from 2008. For more information see: http://www.whitehouse.gov/sites/default/files/omb/assets/bulletins/b10-02.pdf

Table 4-2 Monitor and Area Information for the 12 Urban Case Study Areas in the Epidemiology Based Risk Assessment

Area Name	# of Counties	# of O_3 Monitors	Population (2010)	2007 4th high (ppb)	2009 4th high (ppb)
Atlanta	28	13	5,268,860	102	77
Baltimore	7	7	2,710,489	92	83
Boston	7	11	4,552,402	89	75
Cleveland	5	10	2,077,240	83	72
Denver	10	16	2,543,482	97	79
Detroit	6	8	4,296,250	93	73
Houston	10	22	5,946,800	90	91
Los Angeles	2	21	12,828,837	105	108
New York	23	22	18,897,109	94	81
Philadelphia	11	15	5,965,343	102	74
Sacramento	4	17	2,149,127	93	96
St. Louis	16	17	2,812,896	94	74

4.3.2 Recent Air Quality

The sections below summarize the recent air quality data input into the epidemiological study-based risk assessment, and the exposure and controlled human exposure study-based risk assessment. Additional details on these inputs are provided in Appendix 4-A.

4.3.2.1 Exposure Modeling and Controlled Human Exposure Study Based Risk Assessment

As discussed in more detail in Chapter 5, the REA uses the Air Pollutants Exposure (APEX) model (U.S. EPA, 2012a, b) to simulate exposure and to estimate lung function decrements based on application of results of controlled human exposure studies to populations in the 15 urban case study areas. The APEX model uses spatial fields of hourly O_3 concentrations at each census tract within an urban area to simulate exposure. In the first draft REA, these hourly spatial fields were generated for four urban areas using the concentrations from the nearest neighboring O_3 monitor. In this draft, we use Voronoi Neighbor Averaging (VNA) (Gold, 1997; Chen et al, 2004) to estimate hourly O_3 concentrations at each census tract in all 15 urban case study areas, for recent measured air quality, air quality meeting the existing standard of 75 ppb, and air quality meeting potential alternative standards. The VNA fields were

estimated using ambient hourly O_3 concentrations from monitors in each urban area, as well as monitors within a 50 km buffer region around the boundaries of each area. Additional details on the procedure used to generate the VNA fields, and a technical justification for the change from nearest neighbor fields to VNA fields are included in Appendix 4-A.

Figure 4-5 shows county-level maps of the 15 urban case study areas. Counties colored pink indicate the study area boundaries used in the Zanobetti & Schwartz (2008) and/or Smith et al (2009b) studies[9], where applicable. Counties colored gray indicate additional counties within the CBSA boundaries, and counties colored peach indicate any additional counties included in the exposure and lung function risk assessments. The X's indicate locations of the O_3 monitors used in the risk and exposure assessments, including those within the 50 km buffer region used to create the VNA fields.

[9] The Zanobetti and Schwartz (2008) and Smith et al (2009) study area boundaries were identical for 6 of the 12 urban case study areas, and had at least one county in common for all 12 urban case study areas. The 'Epidemiology Study Area' labels in figures 4-5 refer to counties included in either of these two studies.

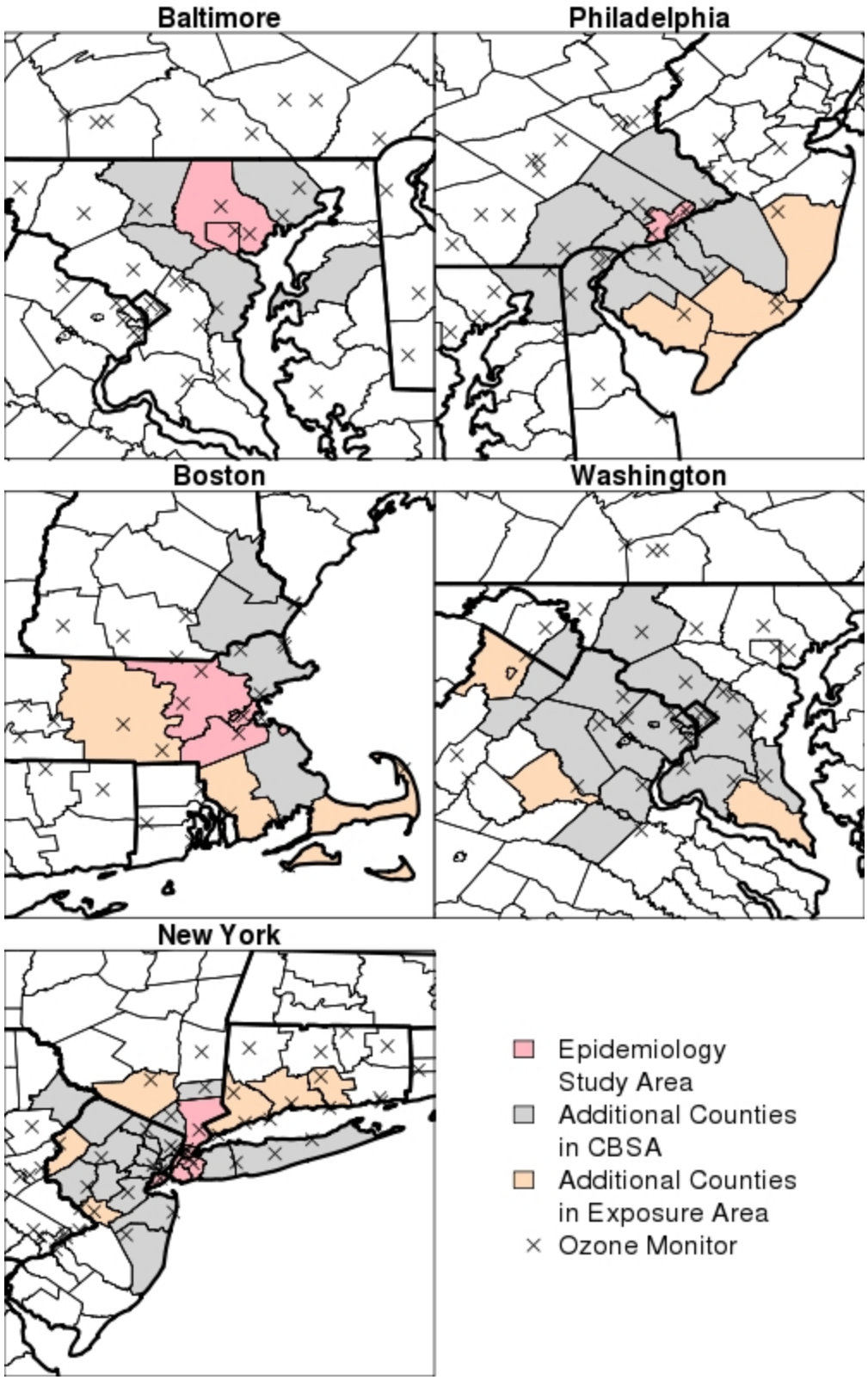


Figure 4-5a Maps of the 5 Eastern U.S. Urban Case Study Areas Including O_3 Monitor Locations

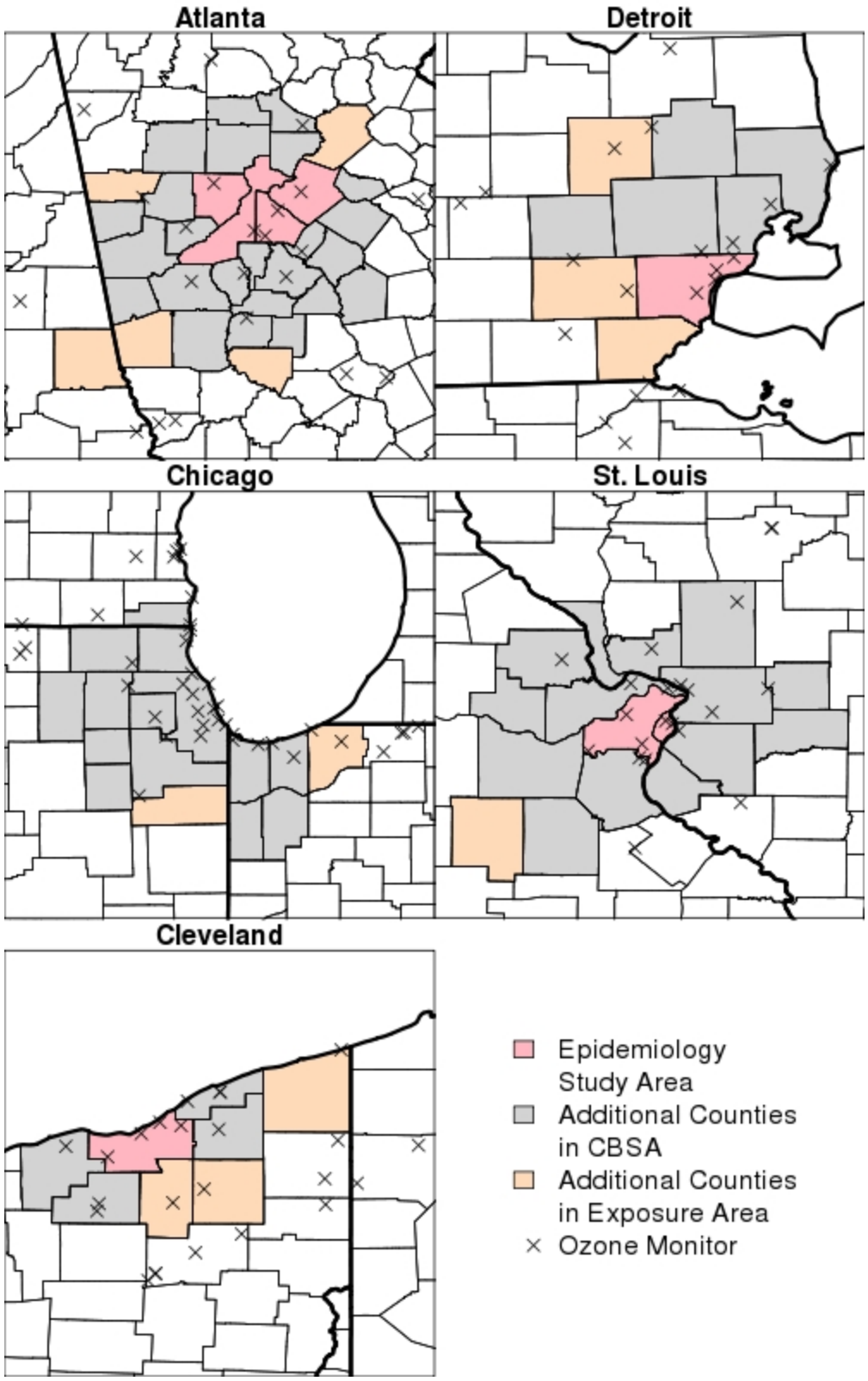


Figure 4-5b Maps of the 5 Central U.S. Urban Case Study Areas Including O_3 Monitor Locations

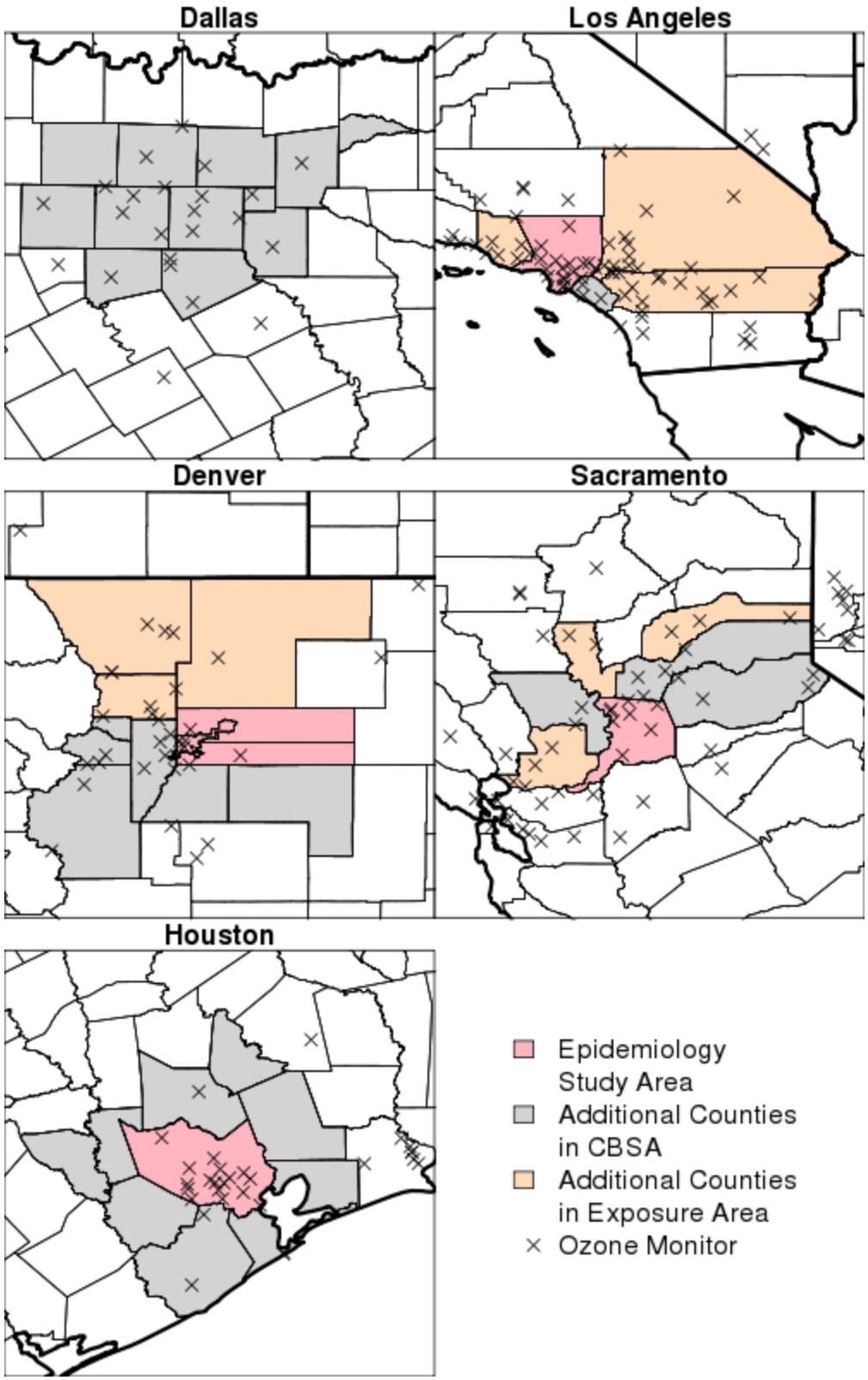


Figure 4-5c Maps of the 5 Western U.S. Urban Case Study Areas Including O_3 Monitor Locations

4.3.2.2 Epidemiology Based Risk Assessment

We input O_3 air quality concentration data for the epidemiology-based risk analyses into the environmental Benefits Mapping and Analysis Program Community Edition (BenMAP-CE) (U.S. EPA, 2013) for assessment. We used BenMAP to analyze four different daily O_3 metrics in 12 of the 15 urban case study areas, which were the basis for concentration-response relationships derived in various epidemiology studies:

1. Daily maximum 1-hour concentration
2. Daily maximum 8-hour concentration
3. Daytime 8-hour average concentration (10:00AM to 6:00PM)
4. Daily 24-hour average concentration

The air quality monitoring data used in BenMAP were daily time-series of "composite monitor" values for each of the 12 urban areas for years 2007 and 2009, which were chosen to represent years with high and low O_3 concentrations, respectively. The composite monitor values were calculated by first averaging the hourly O_3 concentrations for all monitors within the area-of-interest (resulting in a single hourly time-series for each urban area), then calculating the four daily metrics listed above. More details on the composite monitor value calculations and a presentation of the resulting concentrations can be found in Appendices 4-A and 4-D, respectively.

4.3.3 Air Quality Adjustments for "Just Meeting" Existing and Potential Alternative O_3 Standards

The focus of the risk and exposure assessments is the evaluation of risks and exposures after just meeting existing and alternative standards, and the change in risk between just meeting existing standards and just meeting alternative standards. These evaluations require estimation of the change in hourly O_3 concentrations that may occur in each urban area when "just meeting" the existing and potential alternative O_3 standards.

The first draft REA and the previous O_3 NAAQS review used the "quadratic rollback" method to adjust ambient O_3 concentrations to simulate just meeting existing and alternative standards (U.S. EPA, 2007; Wells et al., 2012).Although the quadratic rollback method replicates historical patterns of air quality changes better than some alternative methods (e.g. simply shaving peak concentrations off at the NAAQS level and the proportional rollback technique), its implementation relies on a statistical relationship instead of on a mechanistic characterization of physical and chemical processes in the atmosphere. Because of its construct as a statistical fit to measured O_3 values, the quadratic rollback technique cannot capture spatial and temporal heterogeneity in O_3 response and also cannot account for nonlinear atmospheric chemistry that

causes increases in O_3 during some hours and in some locations as a result of emissions reductions under some circumstances.

Photochemical grid models are better able to simulate these phenomena and therefore the first draft REA proposed to replace quadratic rollback with a model-based O_3 adjustment methodology and presented a test case for Atlanta and Detroit using modeling for July/August 2005 (Simon et al., 2012). The section below summarizes the methodology applied in this second draft REA to adjust air quality for attainment of existing and alternative standards. This new methodology applies Higher-Order Decoupled Direct Method (HDDM) capabilities in the Community Multi-scale Air Quality (CMAQ) model to simulate the response of O_3 concentrations to reductions in US anthropogenic NO_x and VOC emissions. The model incorporates anthropogenic U.S., Canadian, Mexican and other international emissions, as well as emissions from non-anthropogenic sources. Since sources of background O_3 are incorporated explicitly in the modeling, specifying U.S. background concentrations is unnecessary. Application of this approach also addresses the recommendation by the National Research Council of the National Academies (NRC, 2008) to explore how emissions reductions might effect temporal and spatial variations in O_3 concentrations, and to include information on how NO_x versus VOC control strategies might affect risk and exposure.

4.3.3.1 Methods

The EPA has developed an HDDM-adjustment methodology to estimate hourly O_3 concentrations that could occur at each monitor location if urban case study areas were to meet the existing and various alternative levels of the O_3 standard. An early version of this methodology was proposed in the first draft REA (Simon et al., 2012). The methodology was later improved and published in a peer-reviewed journal (Simon et al., 2013). The methodology and its application to hourly O_3 concentrations in the urban case study areas is summarized below and described in more detail in Appendix 4-D.

The HDDM-adjustment methodology uses the CMAQ photochemical model to determine monitoring site-specific response of hourly O_3 concentrations to reductions in US anthropogenic NOx and VOC emissions. These responses are then applied to ambient data to create a 5-year time-series of hourly O_3 concentrations at each monitor location which is consistent with meeting various potential levels of the O_3 NAAQS for the two three-year attainment periods 2006-2008 and 2008-2010. The steps are outlined in Figure 4-6 and summarized below:

- **Step 1:** Run CMAQ simulation with HDDM to determine hourly O_3 sensitivities to NO_x emissions and VOC emissions for the grid cells containing monitoring sites in an urban area.

- **Inputs:** Model-ready emissions and meteorology data
- **Outputs:** O_3 concentrations and sensitivities at locations of monitoring sites for each hour in January and April-October, 2007

- **Step 2:** For each monitoring site, season, and hour of the day use linear regression to relate first order sensitivities of NO_x and VOC (S_{NOx} and S_{VOC}) to **modeled** O_3 and second order sensitivities to NOx and VOC (S^2_{NOx} and S^2_{VOC}) to the first order sensitivities.
  - **Inputs:** Step 1 outputs
  - **Outputs:** Functions to calculate typical sensitivities based on monitor location, O_3 concentration, season, and hour of the day
- **Step 3:** For each **measured** hourly O_3 value between 2006 and 2010, calculate the first and second order sensitivities based on monitoring site-, season-, and hour-specific functions derived in Step 2.
  - **Inputs:** Step 2 outputs and hourly ambient data for 2006-2010.
  - **Outputs:** Hourly O_3 observations paired with modeled sensitivities for all hours in 2006-2010 at all monitor locations
- **Step 4:** Adjust measured hourly O_3 concentrations for incrementally increasing levels of emissions reductions using assigned sensitivities and then recalculate design values until an emissions reduction level is reached at which all monitors in an urban area are below the existing and potential alternative levels of the standard.
  - **Inputs:** Step 3 outputs
  - **Outputs:** Adjusted hourly O_3 values for 2006-2010 at monitor locations to show compliance with the existing and potential alternative standard levels based on the three year average of the 4th highest 8-hour daily max O_3 value. For each standard, two sets of data are created: 2006-2008 and 2008-2010. Because the emissions reductions used to attain standards in the two time periods might be different, adjusted 2008 O_3 values are different for the two sets of data.

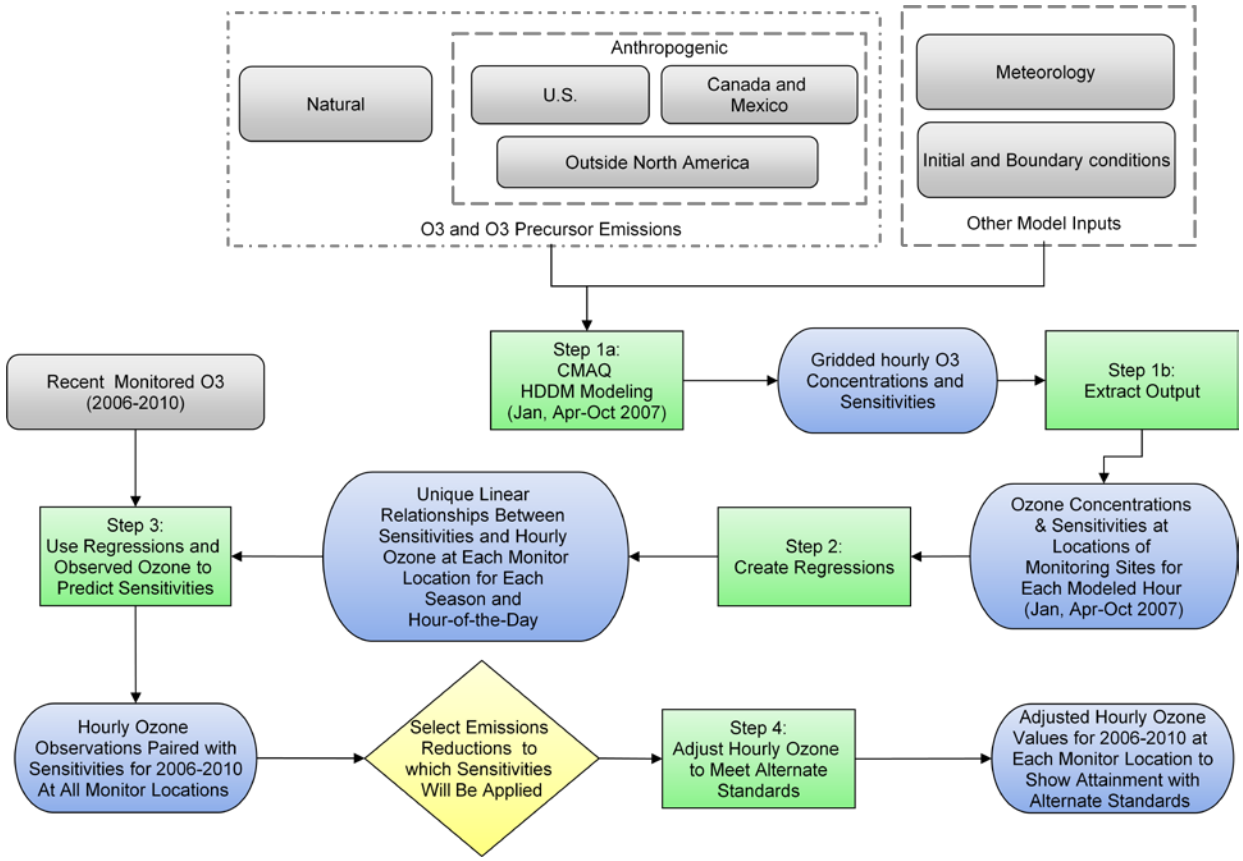


Figure 4-6 Flowchart of HDDM adjustment methodology to inform risk and exposure assessments.

We chose to adjust air quality for just meeting the existing and alternative standards by decreasing U.S. anthropogenic emissions of NOx and VOC throughout the U.S. For the purpose of this analysis we used the Community Multiscale Air Quality (CMAQ) model version 4.7.1 equipped with HDDM to simulate 8 months in 2007 (January and April-October). This time period was chosen to cover the full O_3 season and also includes at least one month from each season of the year. A full description of the model inputs, model set-up, and operational model evaluation against ambient data is available in Appendix 4-B. Sensitivities derived from the 2007 model simulation were applied to the two 3-year periods of ambient data (2006-2008 and 2008-2010) described in section 4.3.1.1. By applying equal proportional decreases in emissions throughout the U.S., we were able to estimate how O_3 would respond to changes in ambient NOx and VOC concentrations without simulating a specific control strategy. The model was set up to track response in hourly O_3 concentrations to these across-the-board changes in US anthropogenic NOx and VOC emissions. In choosing to apply across the board reductions throughout the modeling domain, we recognize that not all emissions across the domain contribute equally to nonattainment in each urban area. However, by decreasing emissions

across the domain, we allow for the possibility of contribution from both regional and local emissions sources to nonattainment and to the overall distribution of O_3 concentrations in urban areas. The modeling included sources which contribute to background O_3 such as biogenic emissions, wildfire emissions, and transport of O_3 and its precursors from international source regions. In addition, the HDDM tool was set-up to specifically calculate the changes in O_3 that would occur from changes in US anthropogenic emissions alone, yet to account for the effects of background sources on this response. Consequently, it is not necessary to set a "floor" background O_3 concentration as was done for quadratic rollback because background sources are explicitly accounted for in the model estimates of O_3 response to US anthropogenic emissions.

As described in more detail in Appendix 4-D, the HDDM adjustment methodology estimates hourly O_3 concentrations that would be associated with attaining a targeted level of the standard either though reductions in US anthropogenic NOx emissions alone or through reductions of both US anthropogenic NOx and VOC emissions in equal percentages. Because the combined NOx/VOC cuts are constrained to equal percentage cuts of both precursors, this is not an optimized NOx/VOC control scenario but rather a sensitivity analysis to characterize the range of results that could be obtained with alternate assumptions. In most of the urban areas, although the NOx/VOC scenario affected O_3 response on some days, it did not affect O_3 response at the highest design value (or controlling) monitor in such a way to reduce the total required emissions cuts. However, for the two cities of Chicago and Denver, the NOx/VOC scenarios allowed for lower percentage emissions cuts (applied to both NOx and VOC) to reach targeted standard levels than the NOx only scenario. Because of this, the core analyses presented in Chapters 5, 6, and 7 were based on the NOx only assumption for all cities except for Chicago and Denver which used the NOx/VOC equal percentage reduction assumption. Sensitivity analyses were performed to compare the NOx only and the NOx/VOC cases in 9 cities: Denver, Detroit, Houston, Los Angeles, New York, Philadelphia, and Sacramento. The effects of these sensitivity analyses on air quality and on the epidemiology-based risk assessment are discussed in more detail in Appendix 4-D and Chapter 7, respectively.

For New York and Los Angeles it should also be noted that a somewhat different approach was used for the HDDM-adjustment application. The HDDM adjustment methodology produces estimates of hourly O_3 concentrations with standard error bounds for every potential emission reduction scenario. Uncertainties in the application of the methodology to very large emissions perturbations along with the fact that the mean estimate does not capture the variability in modeled responses on similar days resulted in the inability of this methodology to estimate O_3 distributions in these two cities which would meet lower alternative standard levels (65 ppb for New York, 60 ppb for Los Angeles). This does not indicate that these two areas would not be able to meet these lower standard levels in reality, but simply reveals the

limitations of this adjustment methodology. Consequently for these two cities, we used the 95th percent confidence interval lower bound estimate of hourly O_3 concentrations to capture a scenario in which these cities could meet lower standard levels based on the range of responses in O_3 concentrations to emissions reduction predicted by the model for each city (See Appendix 4-D for more details). Estimates of risk for these two cities for these alternative standards will be significantly more uncertain, reflecting the use of the lower bound O_3 predictions.

4.3.3.2 Resulting Air Quality

The HDDM adjustment technique tended to have several effects on the distribution of air quality values. First, adjusted hourly O_3 concentrations at night and during the morning rush-hour tended to be higher than the recent observed concentrations (additional details are provided in Appendix 4-D). The CMAQ model predicts that, in general, these times have NOx titration conditions meaning that a reduction in NOx causes an increase in O_3 concentrations. The NOx titration effect was most pronounced in urban core areas which have higher volume of mobile source NOx emissions from vehicles than do the surrounding areas. Response of daytime concentrations was more varied. In general, O_3 tended to increase on low days and decrease on high days. However, specific monitors that were either always heavily VOC limited or always heavily NOx limited showed consistent increases and decreases respectively regardless of whether O_3 concentrations were high or low on a particular day. It should be noted that locations which were heavily VOC limited tended to have much lower observed O_3 concentrations than downwind areas. The tendency of the model to predict O_3 increases on lower concentration days and decreases on higher concentration days also leads to more compressed O_3 distributions in the HDDM adjustment cases. The variability in predicted daily O_3 concentrations decreased when meeting lower standard levels. The following paragraphs summarize a comparison of O_3 distributions from application of the quadratic rollback and HDDM adjustment approach for a case where the existing standard is estimated to be met, characterize the distribution of composite monitor O_3 values at different standard levels, and provide a discussion of the spatial distribution of O_3 changes in several cities. More details and figures for other case-study areas are provided in Appendix 4-D**.**

Figures 4-7 and 4-8 show a comparison of April-October composite monitor O_3 distributions for recent conditions (2006-2008) and for meeting the existing standard using the quadratic rollback technique versus the HDDM adjustment methodology. The composite monitor values in these plots are based on the monitors included in the composite monitor from the Zanobetti and Schwartz (2008) study which was used in the 1st draft REA and do not include all monitors in the CBSA as used in the main Chapter 7 analysis. In general, the O_3 distribution in

the HDDM adjustment case is shifted upward compared to the quadratic rollback case. The upward shift is more pronounced in the lower parts of the O_3 distribution. In all cities displayed in Figure 4-7, the 25th percentile, median, and mean of the 8-hour daily maximum O_3 concentrations are higher in the HDDM adjustment case than the quadratic rollback. In some cities (Sacramento and St. Louis) the 75th percentile values appear approximately equivalent in the two cases while in other cities the 75th percentile values are slightly higher in the HDDM adjustment case. In Houston, the very highest portion of the O_3 distribution is lower in the HDDM adjustment case than in the quadratic rollback case but in many cities the upper parts of the distributions for these two cases are roughly equivalent. Similar results are seen in the 2008-2010 time period; however there are more cases during this time period where HDDM adjustment and quadratic rollback have similar values in the upper half of the O_3 distribution. A comparison of Figure 4-7 and 4-8 shows that there is some seasonality to this effect. The two techniques appear to give very similar 8-hour daily maximum O_3 composite monitor distributions during the summer months (June-August) and most of the situations with higher O_3 levels with the HDDM adjustment come from cooler, lower O_3 time periods (April, May, September, and October). Although here we discuss composite monitor distributions based on April-October, the risk analyses in Chapter 7 are based on the required O_3 monitoring season, which is longer than April – October for some cities. We expect that the O_3 increases shown for spring and fall months here are also representative of the type of response in other "cool season" months. The exceptions to this occur in Denver, Houston, New York and Los Angeles which have higher composite monitor O_3 values from the HDDM adjustment compared to quadratic rollback even in the summer time period.

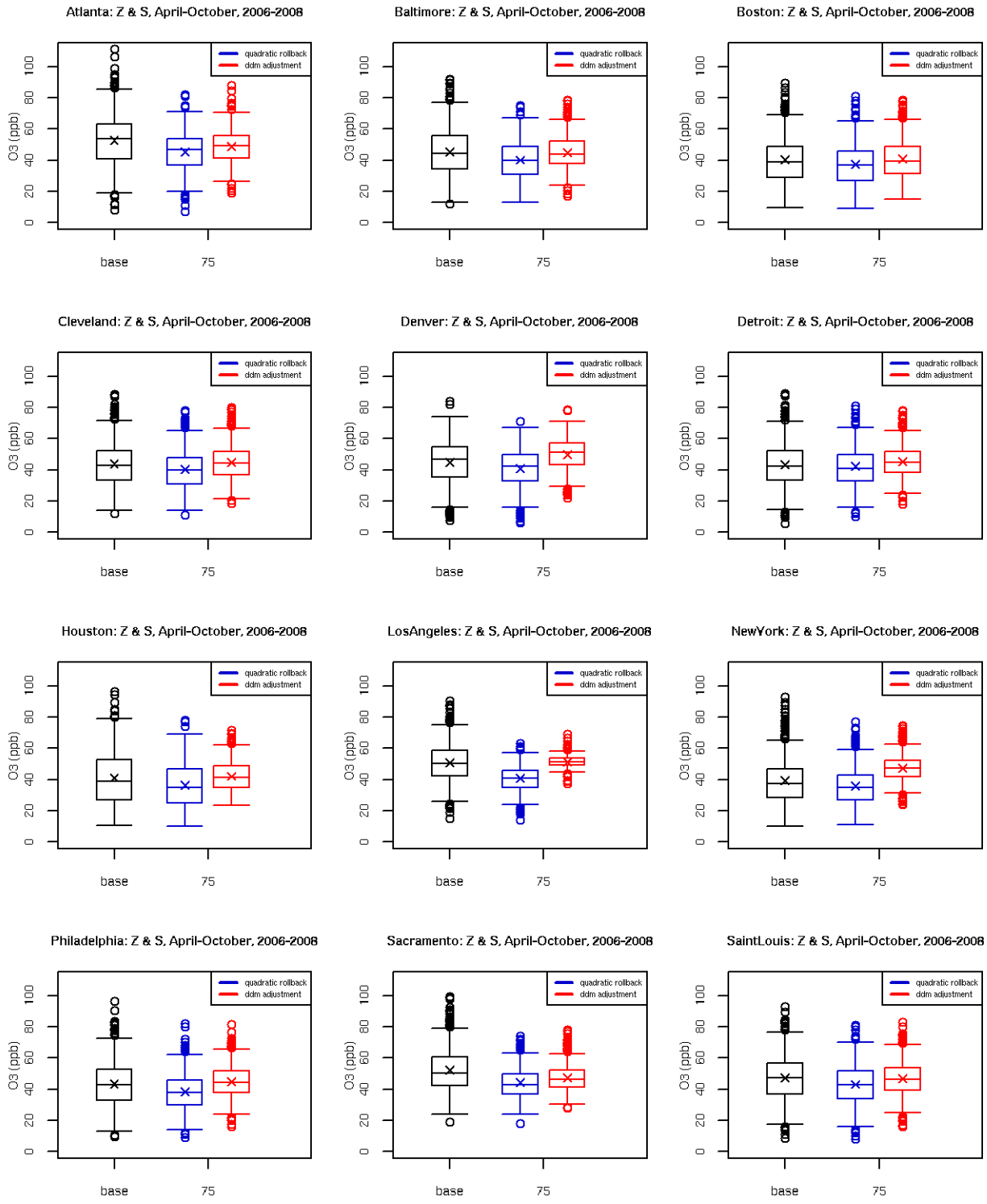


Figure 4-7 Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard. Values are based on the Zanobetti & Schwartz study areas for April-October of 2006-2008.

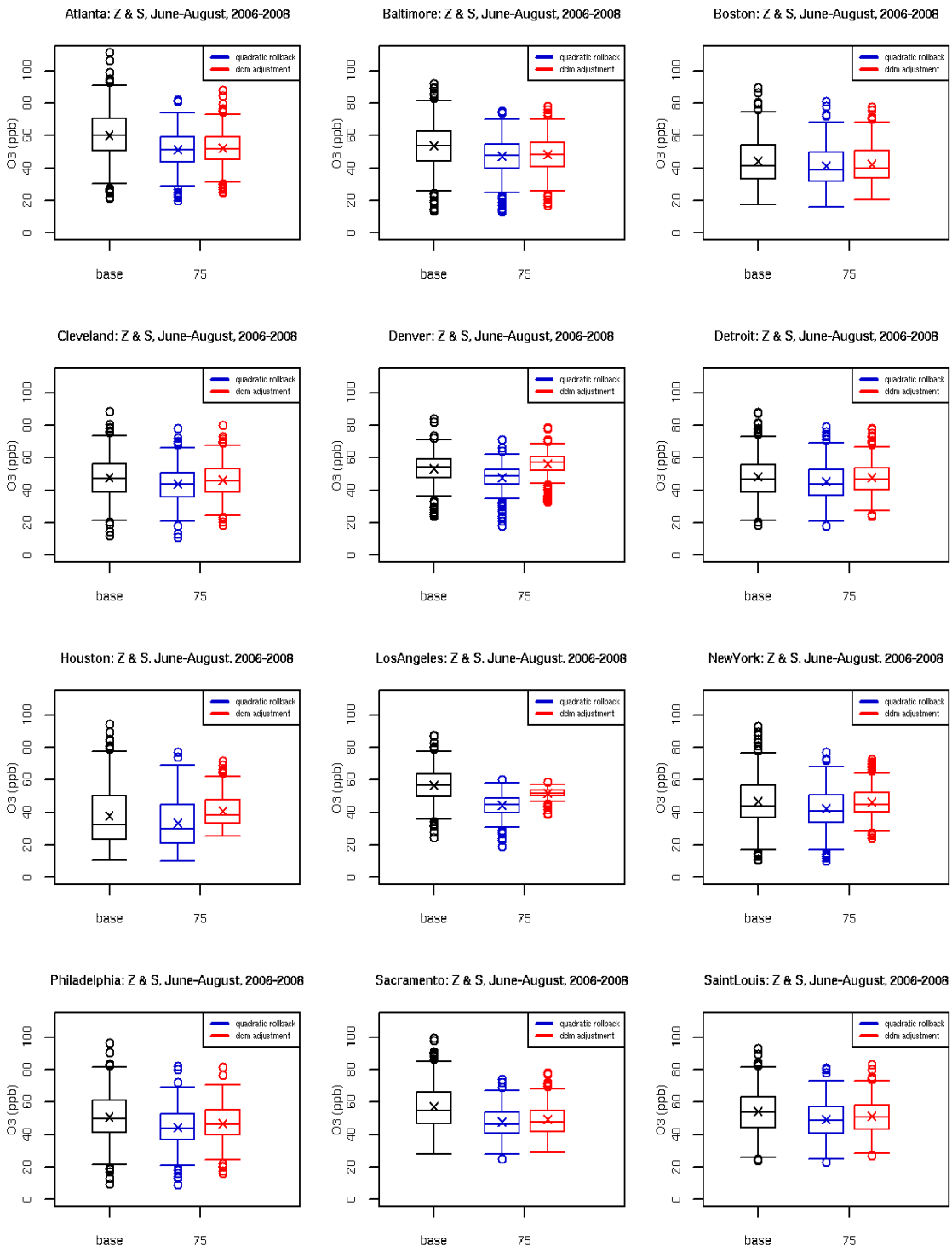


Figure 4-8 Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard. Values are based on the Zanobetti & Schwartz study areas for June-August of 2006-2008.

Figures 4-9 and 4-10 show "box-and-whisker" plots of the April-October composite monitor daily maximum 8-hour O_3 concentration distributions for the 12 urban case study areas evaluated in the epidemiology-based risk assessment; for recent air quality, and air quality adjusted to meet the existing and potential alternative standards. Figure 4-9 shows values from 2007, while figure 4-10 shows values from 2009. Appendix 4-D contains additional plots comparing the changes in the distribution of composite monitor values in each urban area due to the air quality adjustments across varying spatial extents, season lengths, and years. In general, the range of the composite monitor distributions decreased (i.e. the minimum value increased, while the maximum value decreased) in all 12 urban case study areas as the air quality data were adjusted to meet lower standard levels. However, the changes within the inter-quartile range of these distributions (represented by the "boxes") varied in response to the model-based air quality adjustments across the 12 urban areas. Three different types of responses are highlighted in the boxplots for Atlanta, New York, and Houston.

The Atlanta boxplots provide an example of an urban area in which all but the lowest composite monitor values decreased as the air quality data was adjusted to simulate compliance with progressively lower levels of the standard. The upper tail of the distribution (represented by the top whisker in each boxplot) decreased more quickly than the remainder of the distribution, resulting in less total variability in the composite monitor values with each progressively lower standard level. This type of response was also seen Sacramento and St. Louis, and to a lesser extent in Baltimore, Denver, and Philadelphia.

In New York, the boxplots showed an initial increase in the 25th percentile and median composite monitor values when the observed O_3 concentrations were adjusted to meet the existing standard. However, the median composite monitor value decreased relative to the existing standard as O_3 concentrations were adjusted to meet the 70 ppb standard, and both the median and 25th percentile values decreased when air quality were further adjusted to meet the 65 ppb standard. When the air quality were adjusted to meet 65 ppb, the median and mean composite monitor values were lower than under observed conditions. This type of response was also observed in Cleveland, Detroit, and Los Angeles.

In Houston, the median composite monitor value also increased between observed air quality and air quality adjusted to meet the existing standards. However, the pattern in Houston differed from New York and other cities as air quality was further adjusted to reflect meeting the potential alternative standards. The median value remained relatively constant relative to the existing standard, while the 25th percentile values continued to increase. Thus, in Houston, the air quality adjustments always resulted in a median composite monitor value higher than what

was seen in the observed data. The composite monitor distributions in Boston also exhibited this type of behavior.

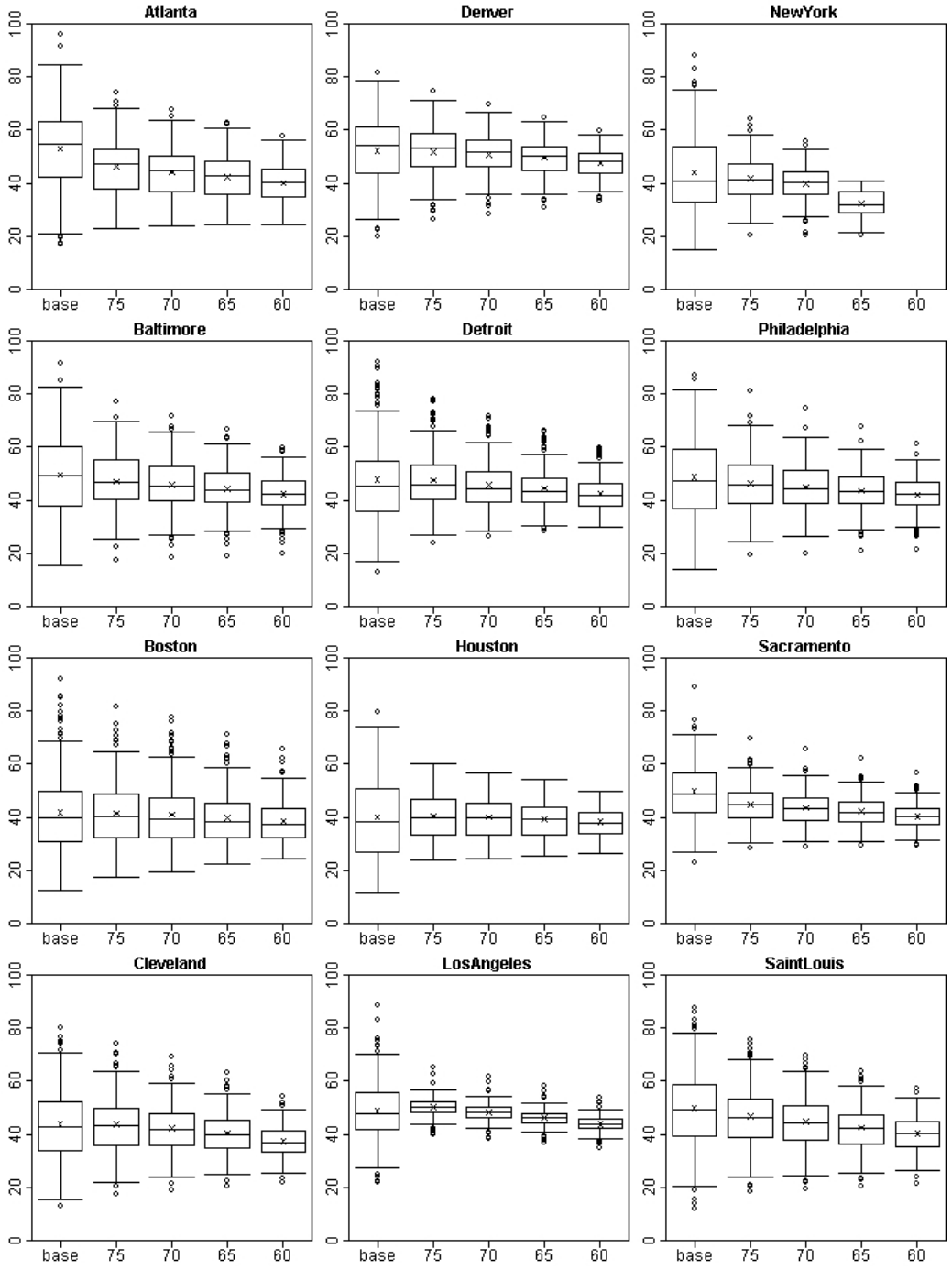


Figure 4-9 Distributions of composite monitor 8-hour daily maximum values for the 12 urban case study areas in the epidemiology-based risk assessment. Plots depict values based on ambient measurements (base), and values obtained with the HDDM adjustment methodology showing attainment of 75, 70, 65 and 60 ppb standards. Values shown are based on CBSAs for April-October of 2007. Note that the HDDM adjustment technique was not able to adjust air quality to show attainment of a 60 ppb standard in New York, so no boxplot is shown for that case.

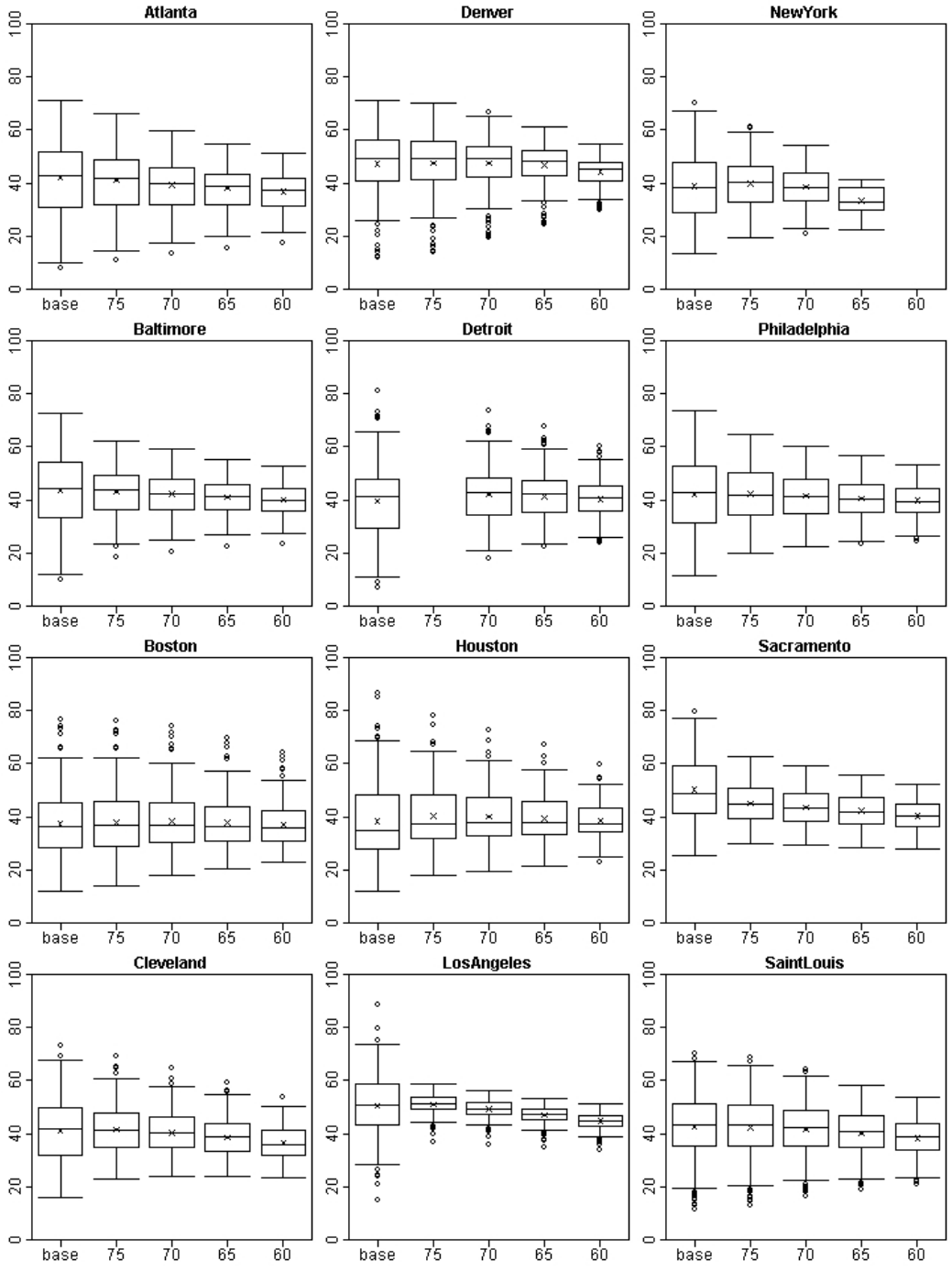


Figure 4-10 Distributions of composite monitor 8-hour daily maximum values for the 12 urban case study areas in the epidemiology-based risk assessment. Plots depict values based on ambient measurements (base), and values obtained with the HDDM adjustment methodology showing attainment of 75, 70, 65 and 60 ppb standards. Values shown are based on CBSAs for April-October of 2009. Note that Detroit air quality was meeting 75 ppb in 2008-2010, and the HDDM adjustment technique was not able to adjust air quality to show attainment of a 60 ppb standard in New York, so no boxplots are shown for those cases.

The exposure modeling and the clinical-based risk assessments used spatially varying surfaces of hourly O_3 concentrations estimated at the centroid of each census tract within the 15 urban case study areas. The maps in Figures 4-11, 4-12, and 4-13 depict the spatial distributions of the 2006-2008 average 4th highest (top) and May – September mean (bottom) daily maximum 8-hour (MDA8) O_3 concentrations for 3 of the 15 urban case study areas; for observed air quality (left), air quality adjusted to meet the existing standard (center), and air quality adjusted to meet the 65 ppb alternative standard (right). Appendix 4-A contains additional maps of the observed 4th highest MDA8 and May – September mean MDA8 concentrations in all 15 urban case study areas for 2006-2008 and 2008-2010. Appendix 4-D contains maps and related figures showing the changes in air quality that resulted from the HDDM adjustments for just meeting the existing standard, and just meeting the potential alternative standard of 65 ppb.

These maps portray the general pattern seen in all 15 urban case study areas for the 4th highest concentrations, which decreased when observed air quality were adjusted to meet the existing standard, and continued to decrease as the air quality were further adjusted to meet the various alternative standards. The May-September average values also generally decreased in suburban and rural areas surrounding the urban population center in all 15 areas. However, three different types of general behavior which were seen in the seasonal average values near the urban population centers, which are exemplified in Figures 4-11 (Atlanta), 4-12 (New York), and 4-13 (Houston).

In Atlanta, the observed May - September average were nearly constant across the entire study area. The observed values decreased nearly uniformly across the entire study area when observed air quality was adjusted to meet the existing standard, and continued to do so when air quality was further adjusted to meet the alternative standard of 65 ppb. The magnitudes of these decreases were slightly larger in suburban and rural areas than near the urban population center. This type of behavior was also seen in Sacramento and Washington, D.C.

In New York, the observed May – September average values were lower near the urban population center than in the surrounding suburban areas. When the observed air quality was adjusted to meet the existing standard, the seasonal average values increased near the urban population center and decreased in the suburban areas, so that the spatial pattern was reversed. When air quality was further adjusted to meet the 65 ppb alternative standard, large area-wide decreases in the seasonal average values were seen relative to the existing standard. While New York represents one of the most extreme examples, similar behavior was observed in 7 other urban areas: Baltimore, Cleveland, Dallas, Detroit, Los Angeles, Philadelphia, and St. Louis.

Houston started out in a similar fashion as New York. The observed May – September average concentrations were lower near the urban population center than in the surrounding

areas, and a similar pattern of increasing and decreasing seasonal average values occurred when observed air quality was adjusted to meet the existing standard. However, unlike New York, the seasonal average values near the Houston city center remained nearly constant relative to the existing standard when air quality were further adjusted to meet the 65 ppb standard. Boston, Chicago, and Denver exhibited this same type of behavior.

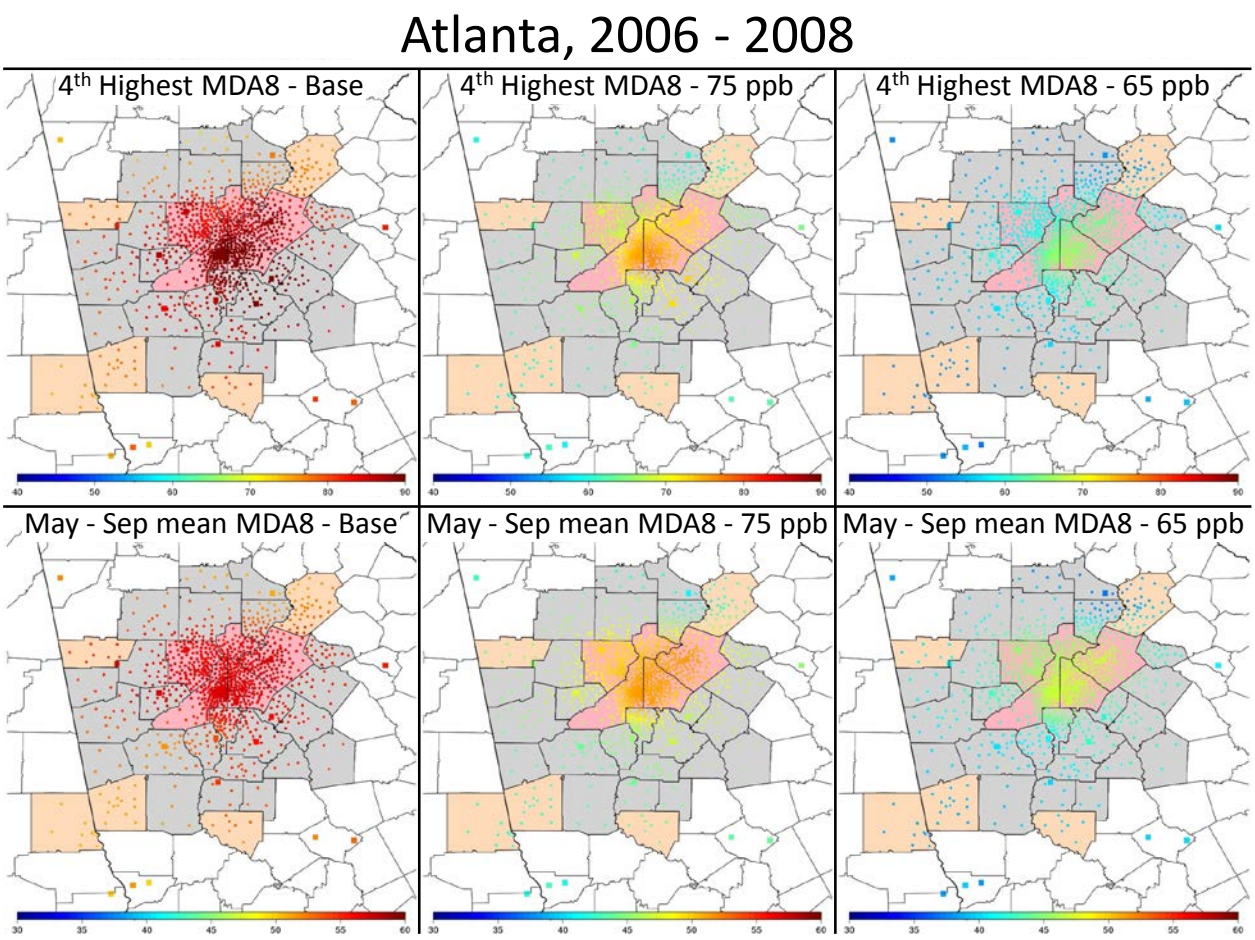


Figure 4-11 Maps showing the 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in Atlanta based on 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids.

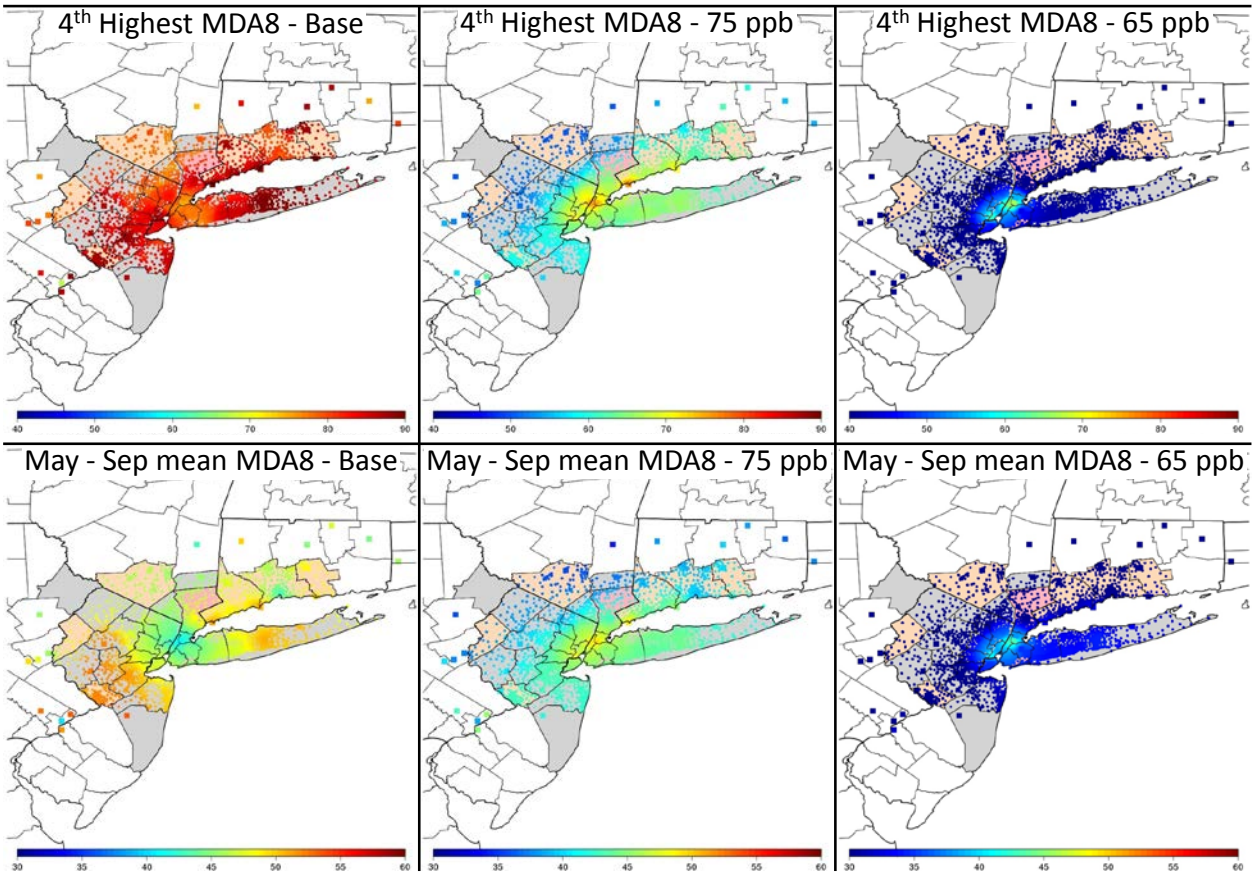


Figure 4-12 **Maps showing the 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in New York based on 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids.**

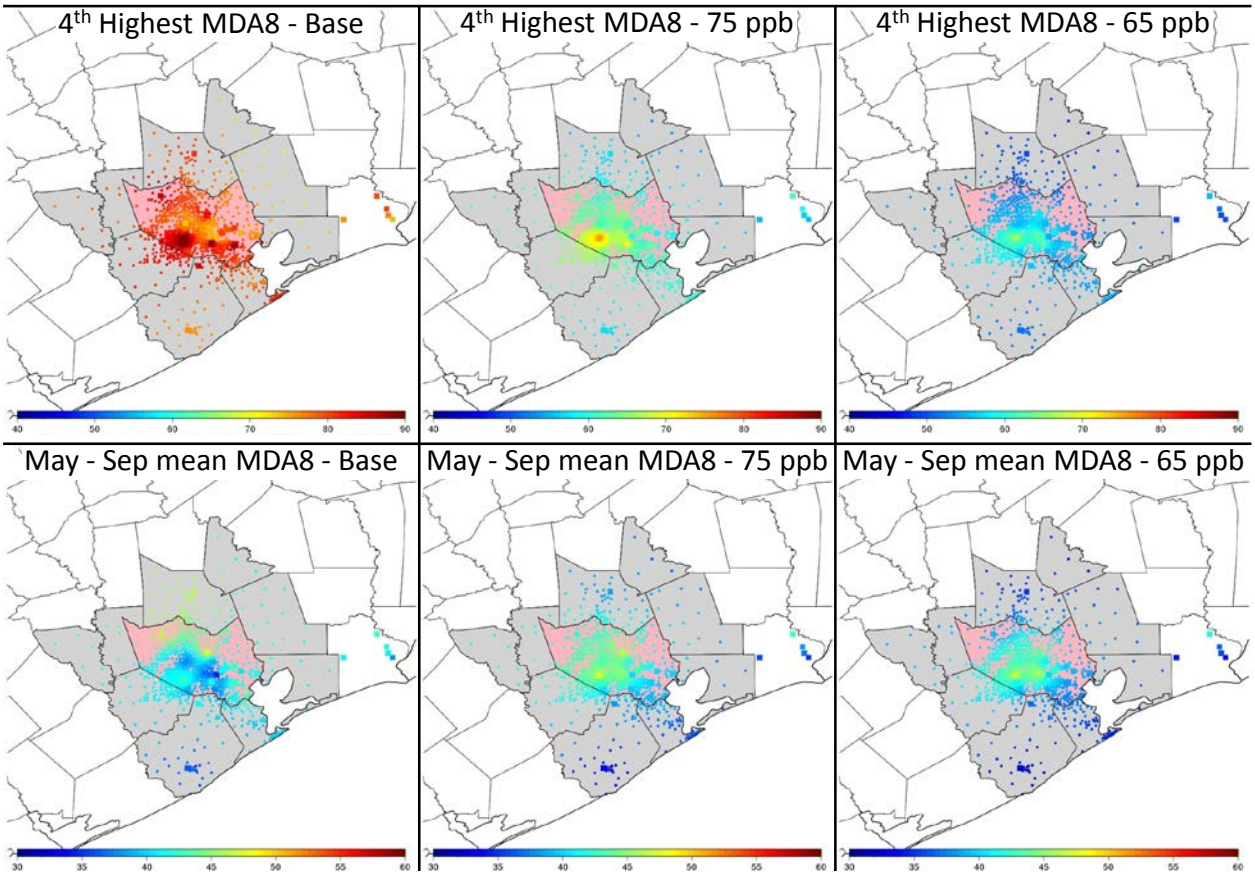


Figure 4-13 Maps of 4th highest (top) and May-September average (bottom) daily maximum 8-hour O_3 concentrations in Houston for 2006-2008 ambient measurements (left), HDDM adjustment to meet the existing standard (center), and HDDM adjustment to meet the alternative standard of 65 ppb (right). Squares represent measured values at monitor locations; circles represent VNA estimates at census tract centroids.

4.4 OVERVIEW OF NATIONAL-SCALE AIR QUALITY INPUTS

The national-scale analysis, presented in Chapter 8, is focused only on evaluating the total national burden of mortality risk associated with recent O_3 conditions. As such it uses a different approach to characterize air quality conditions throughout the U.S. The national-scale analysis employs a data fusion approach that takes advantage of the accuracy of monitor observations and the comprehensive spatial information of the CMAQ modeling system to create national-scale "fused" spatial surfaces of seasonal average O_3 concentrations. Measured O_3 concentrations from 2006-2008 were fused with modeled concentrations from a 2007 CMAQ

model simulation, run for a 12 km domain covering the contiguous U.S. In the first draft of the REA, the spatial surfaces were created using the enhanced Voronoi Neighbor Averaging (eVNA) technique (Timin et al, 2010), using the EPA's Model Attainment Test Software (MATS; Abt Associates, 2010b). In this draft, the spatial surfaces are created using EPA's Downscaler software (Berrocal et al, 2012). More details on the ambient measurements, the 2007 CMAQ model simulation, the Downscaler fusion technique, and a technical justification for changing from eVNA to Downscaler can be found in Appendix 4-C.

Three national "fused" spatial surfaces were created for:

1) the May-September average of the 8-hour daily maximum O_3 concentrations (consistent with the metric used by Smith et al. 2009);

2) the June-August average of the daily 10am-6pm mean O_3 concentrations (consistent with the metric used by Zanobetti and Schwartz 2008); and

3) the April-September average of the 1-hour daily maximum O_3 concentrations (consistent with the metric used by Jerrett et al 2009).

Figures 4-14 to 4-16 show the geographic distributions of these spatial surfaces. The spatial distributions of these three surfaces are very similar, with the highest levels occurring in Southern California for all three surfaces.

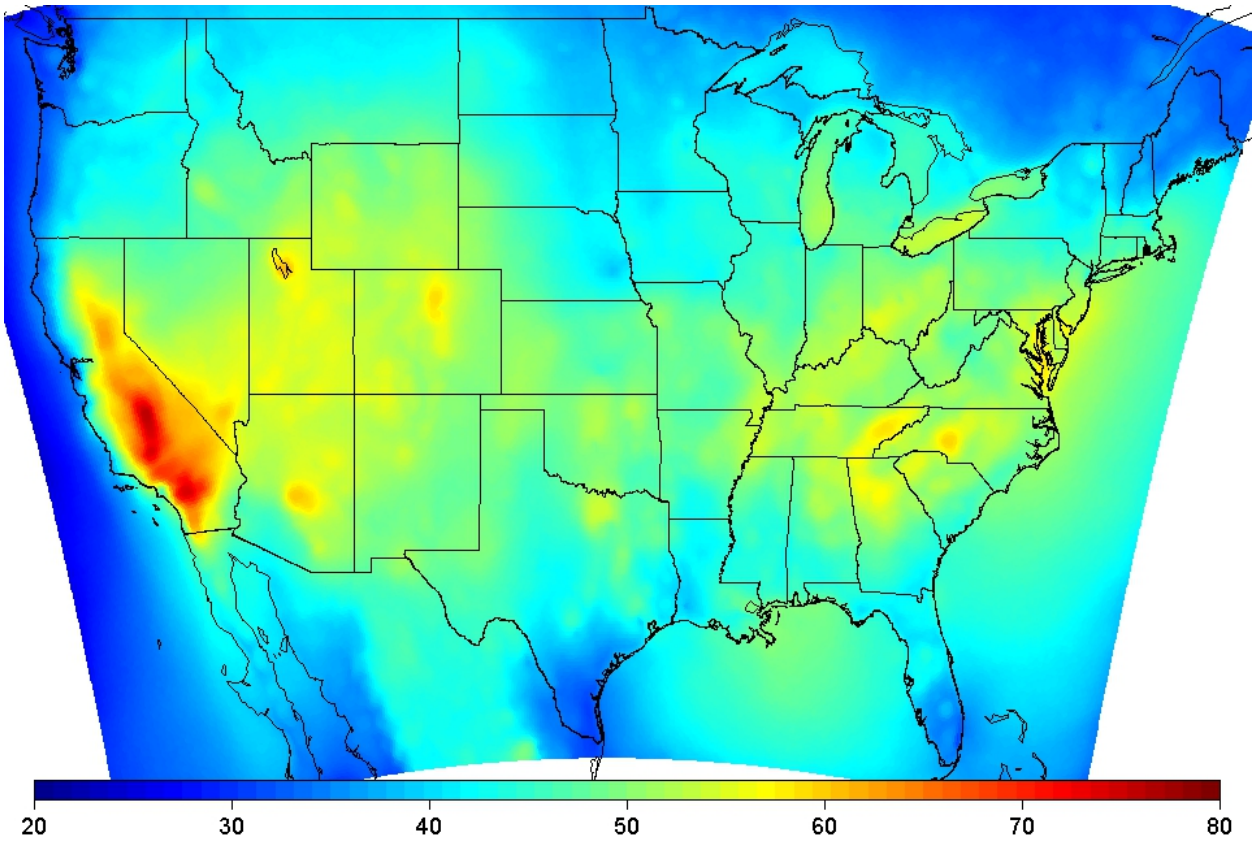


Figure 4-14 May-September average 8-hour daily maximum O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation.

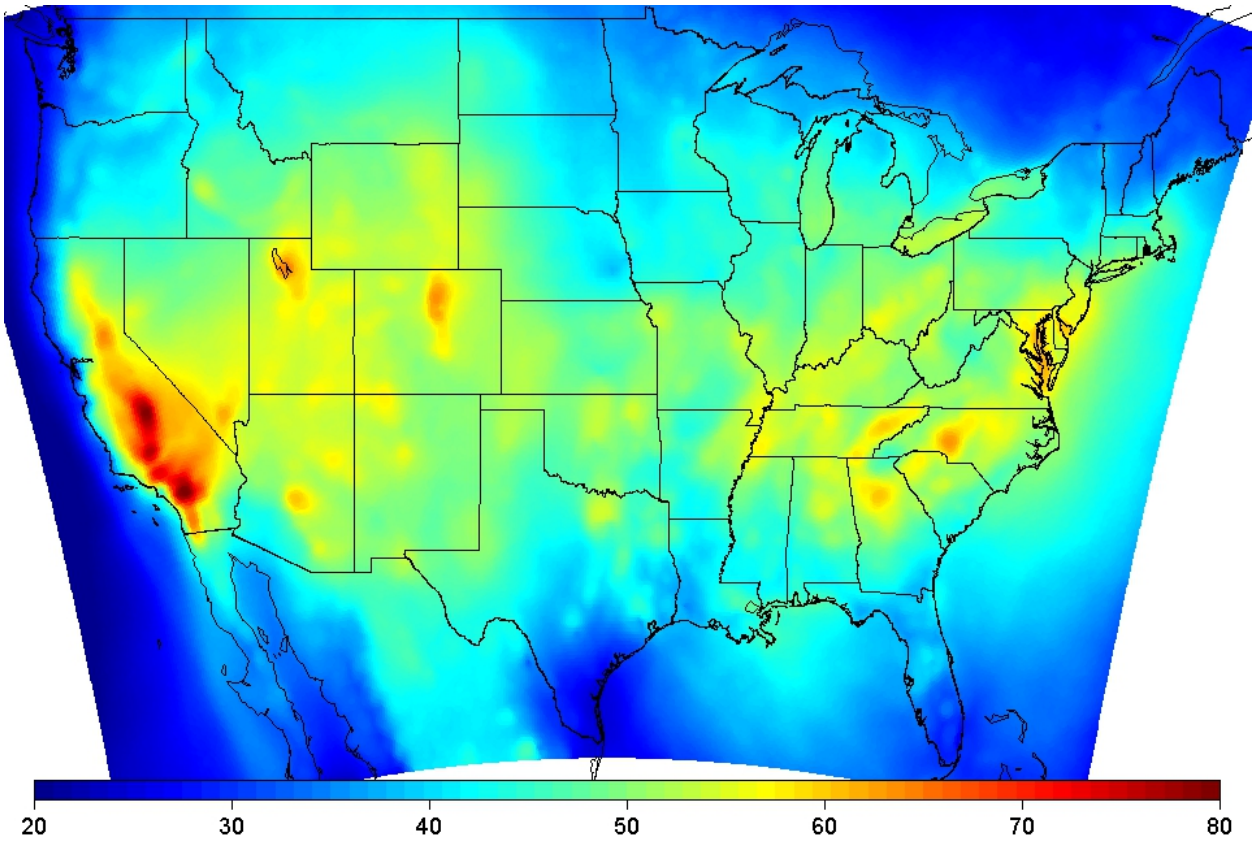


Figure 4-15 June-August average 8-hour daily 10am-6pm mean O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation.

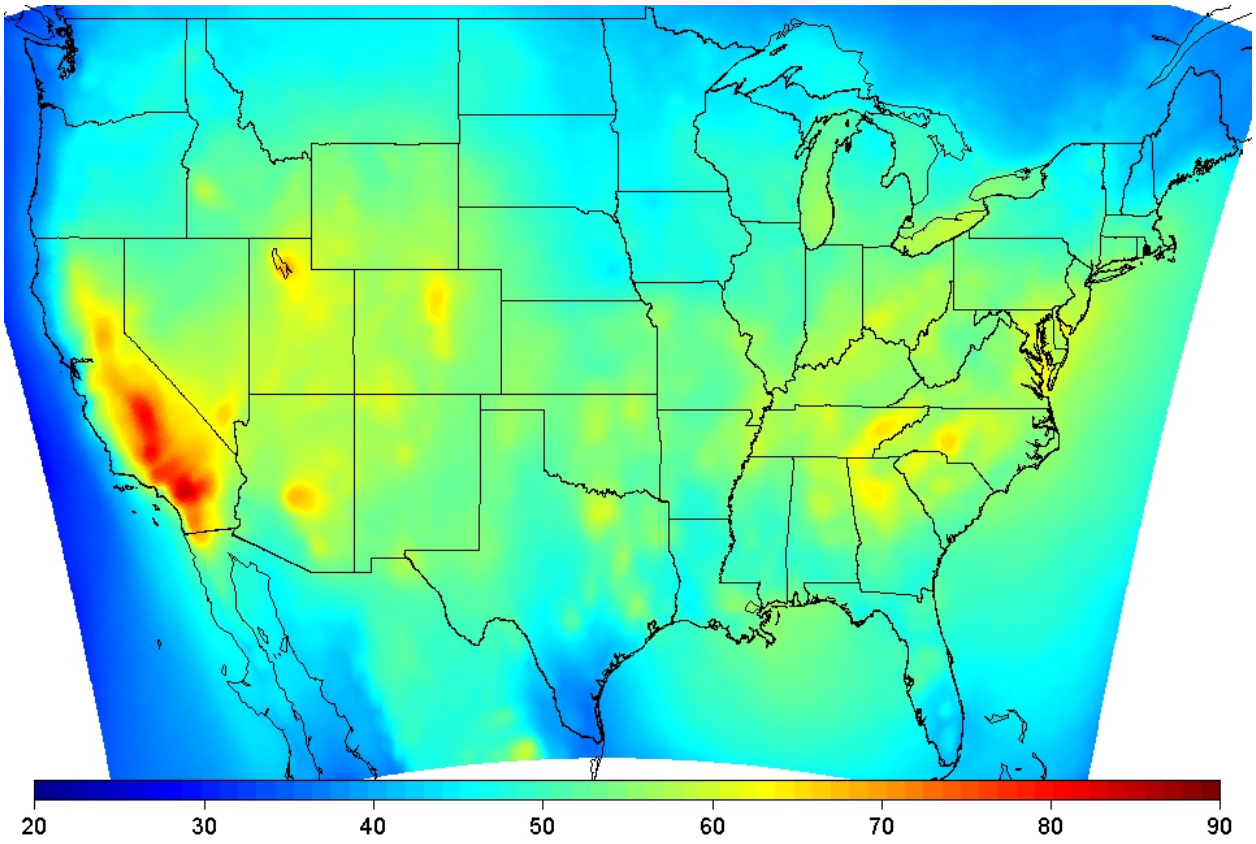


Figure 4-16 April-September average 1-hour daily maximum O_3 concentrations in ppb, based on a Downscaler fusion of 2006-2008 average monitored values with a 12km 2007 CMAQ model simulation.

Figure 4-17 shows the frequency and cumulative distributions of these three seasonal average O_3 surfaces based on all grid cells in the 12 km CMAQ modeling domain. The minimum, median, mean, 95th percentile, and maximum values for all three surfaces are shown in Table 4-3, and correlation coefficients between the three metrics are given in Table 4-4.

The May-September average 8-hour daily maximum concentrations were most frequently in the 30-60 ppb range, while the June-August average daily 10am–6pm mean concentrations were more evenly distributed across a range of 20-60 ppb. The April-September average 1-hour daily maximum concentrations were about 5 ppb higher on average than the May-September average 8-hour daily maximum concentrations, and about 8 ppb higher on average than the June-August average daily 10am-6pm mean concentrations. The correlation coefficients between these three metrics were all very high ($R > 0.97$).

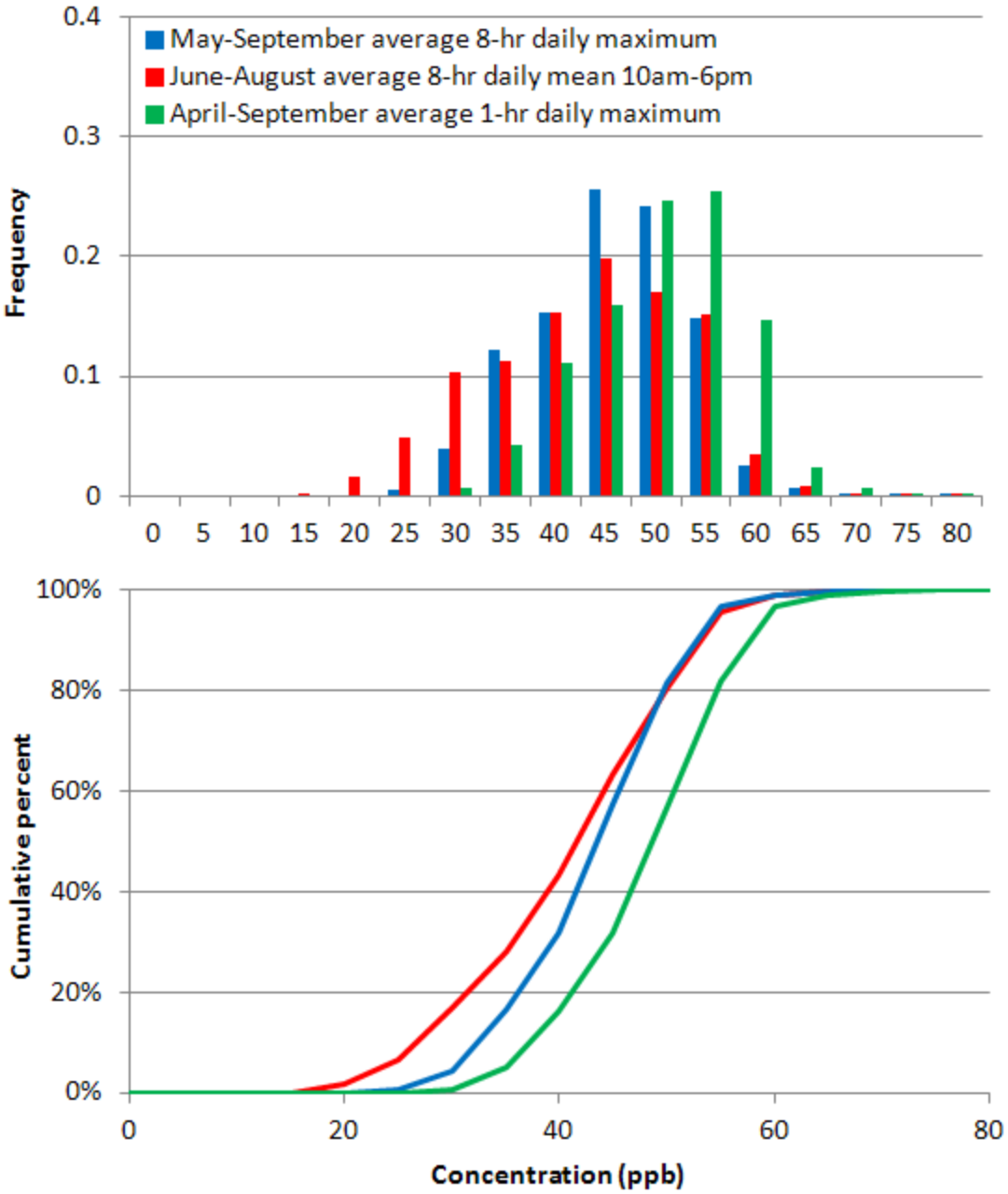


Figure 4-17 Frequency and Cumulative Distributions of the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells.

Table 4-3 Summary Statistics Based on the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells

Statistic	May-September average 8-hour daily maximum concentration (ppb)	June-August average daily 10am–6pm mean concentration (ppb)	April-September average 1-hour daily maximum concentration (ppb)
Minimum	21.8	14.9	26.2
Median	43.6	41.7	48.8
Mean	43.2	40.9	48.2
95th Percentile	54.3	54.8	59.0
Maximum	76.1	80.1	84.2

Table 4-4 Correlation Coefficients Between the Three Fused Seasonal Average O_3 Surfaces Based on all CMAQ 12 km Grid Cells

Seasonal metrics compared	Correlation coefficient
May-September average 8-hour daily maximum vs. June-August average daily 10am-6pm mean	0.974
May-September average 8-hour daily maximum vs. April-September average 1-hour daily maximum	0.995
June-August average daily 10am-6pm mean vs. April-September average 1-hour daily maximum	0.972

These seasonal average metrics are not equivalent to the form of the existing standard, which is based on the 4th highest value rather than on the seasonal mean. Thus, the values shown in the three fused surfaces should not be directly compared to the existing standard. Figure 4-18 shows comparisons between these three metrics and the 2006-2008 O_3 design values based on CMAQ 12 km grid cells containing O_3 monitors, and Table 4-5 presents correlation coefficients and summary statistics based on the ratios between the design values and these three metrics. The design values were, on average, approximately 50% higher than the seasonal average values, with substantial spatial heterogeneity, and some variation across the seasonal average metrics. The April-September average 1-hour daily maximum was the most strongly correlated with the design values (R = 0.75), followed by the May-September average 8-hour daily maximum (R = 0.71), and then the June-August average daily 10am-6pm mean (R = 0.69).

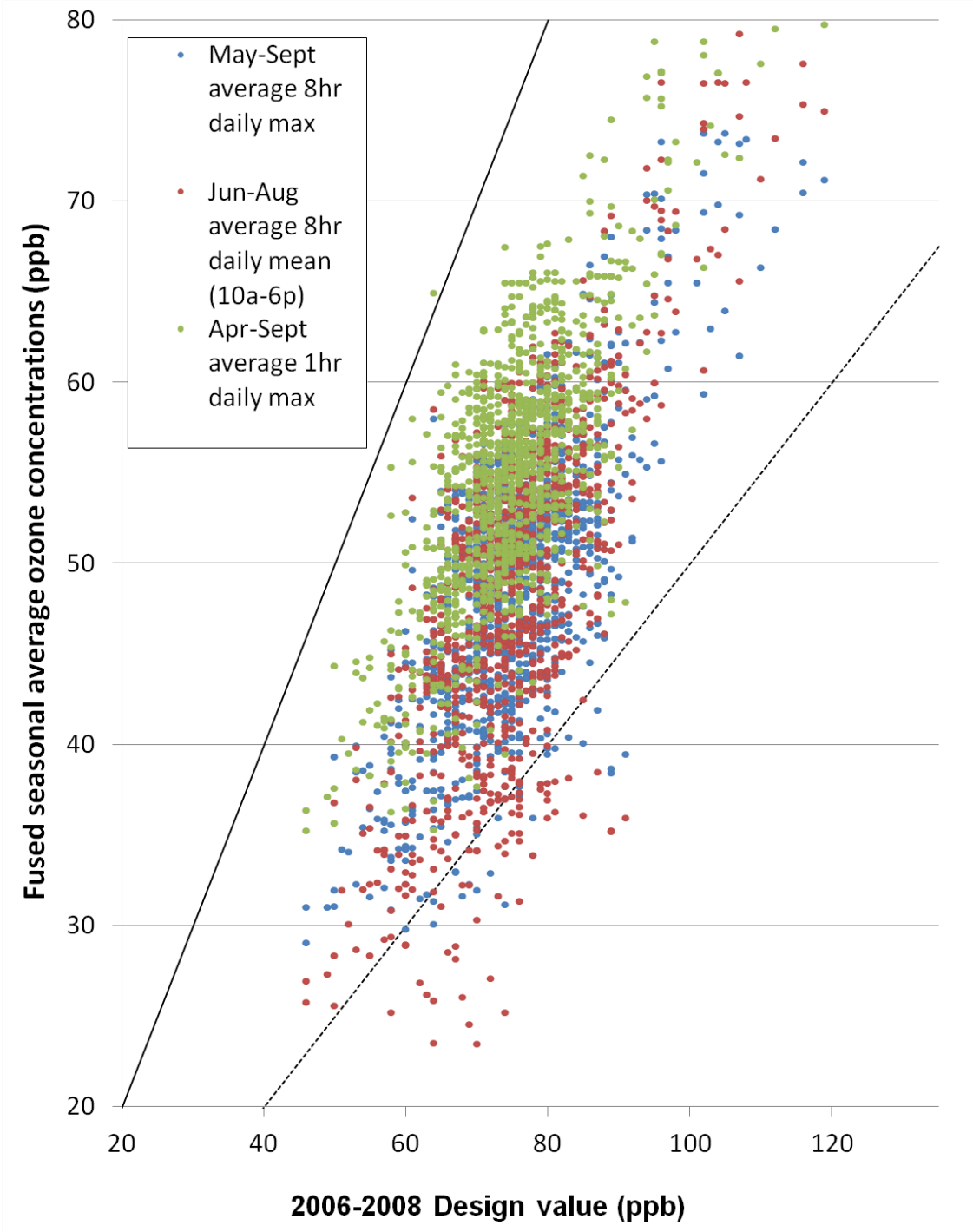


Figure 4-18 2006-2008 O_3 Design Values Versus 2006-2008 Fused Seasonal Average O_3 Levels for the CMAQ 12km Grid Cells Containing O_3 Monitors.

Table 4-5 Correlation Coefficients and Ratios of the 2006-2008 O_3 Design Values to the 2006-2008 Fused Seasonal Average O_3 Levels for the CMAQ 12km Grid Cells Containing O_3 Monitors

Statistic	May-September average 8-hour daily maximum	June-August average daily 10am-6pm mean	April-September average 1-hour daily maximum
Correlation	0.71	0.69	0.75
Ratios			
Minimum	1.1	1.1	1.0
2.5th Percentile	1.3	1.3	1.2
Median	1.5	1.5	1.4
Mean	1.6	1.6	1.4
97.5 Percentile	2.0	2.2	1.6
Maximum	2.4	3.0	1.9

4.5 UNCERTAINITIES IN MODELING OF RESPONSES TO EMISSION REDUCTIONS TO JUST MEET EXISTING AND POTENTIAL ALTERNATIVE STANDARDS

We recognize that there are sources of uncertainty in air quality measurements and the air quality estimates for each air quality scenario. These sources of uncertainty are described below and in Table 4-6 which discusses qualitatively the magnitude of uncertainty and potential for directional bias.

There is inherent uncertainty in all deterministic air quality models, such as CMAQ, the photochemical grid model which was used to develop the model-based O_3 adjustment methodology. Evaluations of air quality models against observed pollutant concentrations build confidence that the model performs with reasonable accuracy despite both structural and parametric uncertainties. A comprehensive model performance evaluation provided in Appendix 4-B shows generally acceptable model performance which is equivalent to or better than typical state-of-the science regional modeling simulations as summarized in Simon et al (2012). The use of the Higher Order Decoupled Direct Method (HDDM) within CMAQ to estimate O_3 response to emissions perturbations adds uncertainty to that inherent in the model itself. HDDM allows for the approximation of O_3 concentrations under alternate emission scenarios without re-running the model simulation with different inputs. This approximation becomes less accurate for larger emissions perturbations. To accommodate increasing uncertainty at larger emissions perturbations, the HDDM modeling was performed at three distinct emissions levels to allow for a better characterization of O_3 response over the entire range of emissions levels. The accuracy

of the HDDM estimates can be quantified at distinct emissions levels by re-running the model with modified emissions inputs and comparing the results. This method was applied to quantify the accuracy of 3-step HDDM O_3 estimates for 50% and 90% NOx cut conditions for each urban case study areas (as shown in Appendix 4-D). At 50% NOx cut conditions, HDDM using information from these multiple simulations predicted hourly O_3 concentrations with a mean bias and a mean error less than +/- 1 ppb in all case study areas compared to brute force model simulations. At 90% NOx cut conditions, HDDM using information from these multiple simulations predicted hourly O_3 concentrations with a mean bias less than +/- 3ppb and a mean error less than +/- 4 ppb in all case study areas. These small bias and error estimates show that uncertainty due to the HDDM approximation method is small up to 90% emissions cuts.

In order to apply modeled O_3 response to ambient measurements, regressions were developed which relate O_3 response to emissions perturbations with ambient O_3 concentrations for every season, hour-of-the-day, and monitor location. Applying O_3 responses based on this relationship adds uncertainty. Preliminary work showed that the relationships developed with these regressions were generally statistically significant for most season, hour-of-the-day, and monitor location combinations for 2005 modeling in Detroit and Charlotte (Simon et al, 2012). Statistical significance was not evaluated for each regression in this analysis since there were over 460,000 regressions created (322 monitors × 5 sensitivity coefficients × 3 emissions levels × 4 seasons × 24 hours = 463,680 regressions). Statistics can quantify the goodness of fit for the modeled relationships and can quantify the uncertainty in response at any given O_3 concentration based on variability in model results at that portion of the distribution for each regression. The regression model provided both a central tendency and a standard error value for O_3 response at each measured hourly O_3 concentration. The base analysis in all case study areas except New York and Los Angeles used the central tendency which will inherently dampen some of the variability in O_3 response. The standard error of each sensitivity coefficient was propagated through the calculation of predicted O_3 concentrations at various standard levels. These standard errors reflect the amount of variability that is lost due to the use of a central tendency. Since emissions reductions increased for lower standard levels the standard errors were larger for adjustments to lower standards. Mean (95th percentile) standard errors for the 75 ppb adjustment case ranged from 0.13 (0.26) to 1.18 (2.87) ppb in the 15 case study areas. Mean (95th percentile) standard errors for the 65 ppb adjustment case ranged from 0.54 (1.07) to 1.39 (2.98) ppb. The largest standard errors occurred in Los Angeles and New York due to the large emissions reductions applied in these cases. In cases where the use of the central tendency of response reduced the total estimated emissions reductions required to achieve a given standard level, in general we expect that the benefits of reducing high O_3 concentrations and the disbenefits of increasing low O_3 would both be underestimated. For the exposure assessment which estimates

health outcomes that occur at O_3 concentrations above 60, this would lead to an underestimation of risks. For the epidemiology-based risk assessment which is effected by the entire range of O_3 concentrations, the impact is undetermined since changes at both ends of the O_3 distribution in opposite directions would affect the results. The opposite would be true in cases where the use of the central tendency of response increased the total estimated emissions reductions required to achieve a given standard. However, given the small standard error values even in the case study areas with the greatest uncertainty (i.e. less than 1.5 ppb mean standard error), this source of uncertainty is not expected to substantially impact results.

Relationships between O_3 response and hourly O_3 concentration were developed based on 8 months of modeling: January and April-October 2007. These relationships were applied to ambient data from 2006-2010. Some locations monitor for months not included in this modeling (i.e., February, March, November, and December) while others do not. Seasonal relationships were developed between O_3 response to emissions reductions and O_3 concentration. Summer was the only season for which modeling data was created for all months (June, July, August). The winter relationships were developed based on January modeling, the spring relationships were developed based on April/May modeling, and the autumn relationships were developed based on September/October modeling. The reduction in data points (31 or 61 instead of ~90) increases uncertainty in the statistical fit for these seasons. In addition, the modeling generally showed more O_3 disbenefits to NOx decreases in cooler months. So applying April/May relationships to March and September/October relationships to November could potentially underestimate O_3 increases that would happen in those two months in the five case study areas which measure O_3 during March and/or November: Dallas, Denver, Houston, Los Angeles, and Sacramento. The eight months that were modeled capture a variety of meteorological conditions. In cases where other years have more frequent occurances of certain types of meteorological conditions, the regressions should be able to account for this. For instance, if a monitor only had 2-3 high O_3 days associated with sunny, high pressure conditions in the 2007 modeling but had 30-40 of those days in another year, the regression may be more uncertain at those high O_3 values but should still be able to capture the central tendency which can be applied to the more frequent occurances in other years. If, on the other hand, the meteorology/ O_3 conditions in another year were completely outside the range of conditions captured in the model, then the regression based on modeled conditions might not be able to capture those conditions. Finally, if emissions change drastically between the modeled period and the time of the ambient data measurements this could also change the relationship between O_3 response and O_3 concentrations. The regressions derived from the 2007 modeling period are only applied to measurements made within 3 years of the modeled time period. Although some emissions changes did occur over this

time period, we believe it is still reasonable to apply 2007 modeling to this relatively small window of measurements which occurs before and after the modeling.

O_3 response is modeled for across-the-board reductions in U.S. anthropogenic NOx (and VOC). These across-the-board cuts do not reflect actual emissions control strategies. The form, locations, and timing of emissions reductions that would be undertaken to meet various levels of the O_3 standard are unknown. The across-the-board emissions reductions bring levels down uniformly across time and space to show how O_3 would respond to changes in ambient levels of precursor species but do not reflect spatial and temporal heterogeneity that may occur in local and regional emissions reductions. In cases where VOC reductions were modeled, equal percentage NOx and VOC reductions were applied in the adjustment methodology. Regional NOx reductions are likely to be the primary means used to reduce high O_3 concentrations at DV monitors. In limited cases, VOC emissions reductions may also help lower high O_3 concentrations at these locations. In actual control strategies, NOx and VOC reductions may be applied in combination but are unlikely to be applied in equal percentages. The available modeling constrained the NOx/VOC case to this type of control scenario. The across-the-board cuts and the equal percentage NOx and VOC reductions scenario does not optimize the lowest cost or least total emissions combinations as state and local agencies will likely attempt to achieve.

Table 4-6 Summary of Qualitative Uncertainty Analysis of Key Air Quality Elements in the O_3 NAAQS Risk Assessment

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
A. O_3 measurements	O_3 concentrations measured by ambient monitoring instruments have inherent uncertainties associated with them. Additional uncertainties due to other factors may include: - monitoring network locations - O_3 monitoring seasons - monitor malfunctions - wildfire and smoke impacts - interpolation of missing data	Both	Low	Low	KB: O_3 measurements are assumed to be accurate to within ½ of the instrument's Method Detection Limit (MDL), which is 2.5 ppb for most instruments. EPA requires that routine quality assurance checks are performed on all instruments, and that all data reported to AQS are certified by both the monitoring agency and the corresponding EPA regional office. The CASTNET monitoring data were subject to their own set of QA requirements, and these data are generally believed to be of comparable quality to the data stored in AQS. KB: Monitor malfunctions sometimes occur causing periods of missing data or poor data quality. Monitoring data affected by malfunctions are usually flagged by the monitoring agency and removed from AQS. In addition, the AQS database managers run several routines to identify suspicious data for potential removal. KB: There is a known tendency for smoke produced from wildfires to cause interference in O_3 instruments. Measurements collected by O_3 analyzers were reported to be biased high by 5.1–6.6 ppb per 100 µg/m^3 of PM2.5 from wildfire smoke ,EPA, 2007). However, smoke concentrations high enough to cause significant interferences are infrequent and the overall impact is believed to be minimal.

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		KB: Missing intervals of 1 or 2 hours in the measurement data were interpolated, which may cause some additional uncertainty. However, due to the short length of the interpolation periods, and the tendency for these periods to occur at night when O_3 concentrations are low, the overall impact is believed to be minimal. INF: EPA's current O_3 monitoring network requirements have an urban focus. Rural areas where O_3 concentrations are lower tend to be under-represented by the current monitoring network. The network requirements also state that at least one monitor within each urban area must be sited to capture the highest O_3 concentrations in that area, which may cause some bias toward higher measured concentrations. INF: Each state has a required O_3 monitoring season which varies in length from May – September to year-round. Some states turn their O_3 monitors off during months outside of the required season, while others leave them on. This can cause discrepancies in the amount of data available, especially in months outside of the required monitoring season. The risk estimates attempt to minimize these impacts by focusing only on months where O_3 monitoring is required.

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
B. Veronoi Neighbor Averaging (VNA) spatial fields	VNA is a spatial interpolation technique used to estimate O_3 concentrations in unmonitored areas, which has inherent uncertainty	Both	Low-Medium	Low-Medium	KB: VNA interpolates monitored hourly O_3 concentrations to provide estimates of O_3 exposure at each census tract in the 15 urban areas. The VNA estimates are weighted based on distance from neighboring monitoring sites, thus the amount of uncertainty tends to increase with distance from the monitoring sites. KB: The 15 urban areas each had fairly dense monitoring networks which were generally sufficient to capture spatial gradients in O_3 concentrations. The use of hourly data to create the VNA fields instead of daily or other aggregates also served to reduce uncertainty by better capturing relationships in the diurnal patterns between O_3 monitors.
C.CMAQ modeling	Model predictions from CMAQ, like all deterministic photochemical models, have both parametric and structural uncertainty associated with them	Both	Low-Medium	Low-Medium	KB: Structural uncertainties are uncertainties in the representation of physical and chemical processes in the model. These include: choice of chemical mechanism used to characterize reactions in the atmosphere, choice of land surface model and choice of planetary boundary layer model. KB: Parametric uncertainties include uncertainties in model inputs (hourly meteorological fields, hourly 3-D gridded emissions, initial conditions, and boundary conditions) KB: Uncertainties due to initial conditions are minimized by using a 10 day ramp-up period

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
					from which model results are not used. KB: Evaluations of models against observed pollutant concentrations build confidence that the model performs with reasonable accuracy despite the uncertainties listed above. A comprehensive model evaluation provided in Appendix 4-B shows generally acceptable model performance which is equivalent or better than typical state-of-the science regional modeling simulations as summarized in Simon et al (2012). However, both under-estimations and over-estimations do occur at some times and locations. Generally the largest mean biases occur on low O_3 days during the summer season. In addition, the model did not fully capture rare wintertime high O_3 events occurring in the Western U.S.
D. Higher Order Decoupled Direct Method (HDDM)	HDDM allows for the approximation of O_3 concentrations under alternate emissions scenarios without re-running the model simulation multiple times using different emissions inputs. This approximation becomes less accurate for larger emissions perturbations	Both	Low-Medium	Low-Medium	KB: To accommodate increasing uncertainty at larger emissions perturbations, the HDDM modeling was performed at three distinct emissions levels to allow for a better characterization of O_3 response over the entire range of emissions levels. The replication of brute force hourly O_3 concentration model results by the HDDM approximation was quantified for 50% and 90% NOx cut conditions for each urban case study areas (as shown in Appendix 4-D). At 50% NOx cut conditions, HDDM using information from these multiple simulations predicted hourly

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
	especially under nonlinear chemistry conditions.				O_3 concentrations with a mean bias and a mean error less than +/- 1 ppb in all urban case study areas compared to brute force model simulations. At 90% NOx cut conditions, HDDM using information from these multiple simulations predicted hourly O_3 concentrations with a mean bias less than +/- 3ppb and a mean error less than +/- 4 ppb in all urban case study areas.
E. Application of HDDM sensitivities to ambient data	In order to apply modeled sensitivities to ambient measurements, regressions were developed which relate O_3 response to emissions perturbations with ambient O_3 concentrations for every season, hour-of-the-day and monitor location. Applying O_3 responses based on this relationship adds uncertainty.	Both	Medium	Medium	KB: Preliminary work showed that the relationships developed with these regressions were generally statistically significant for most season, hour-of-the-day, and monitor location combinations for 2005 modeling in Detroit and Charlotte. Statistical significance was not evaluated for each regression in this analysis since there were over 460,000 regressions created (322 monitors × 5 sensitivity coefficients × 3 emissions levels × 4 seasons × 24 hours = 463,680 regressions). Statistics can quantify the goodness of fit for the modeled relationships and can quantify the uncertainty in response at any given O_3 concentration based on variability in model results at that portion of the distribution for each regression. However it is not possible to quantify the applicability of this modeled relationship to the actual atmosphere. KB: The regression model provided both a central tendency and a standard error value for O_3 response at each measured hourly O_3

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		concentration. The base analysis used the central tendency which will inherently dampen some of the variability in O_3 response. The standard error of each sensitivity coefficient was propagated through the calculation of predicted O_3 concentrations at various standard levels. These standard errors reflect the amount of variability that is lost due to the use of a central tendency. Since emissions reductions increased for lower standard levels the standard errors were larger for adjustments to lower standards. Mean (95th percentile) standard errors for the 75 ppb adjustment case ranged from 0.13 (0.26) to 1.18 (2.87) ppb in the 15 case study areas. Mean (95th percentile) standard errors for the 65 ppb adjustment case ranged from 0.54 (1.07) to 1.39 (2.98) ppb. The largest standard errors occurred in Los Angeles and New York. INF: The NOx emissions reductions resulted in both increases and decreases in O_3 depending on the time and location. In cases where the use of the central tendency of response reduced the total estimated emissions reductions required to achieve a given standard level, in general we expect that the benefits of reducing high O_3 concentrations and the disbenefits of increasing low O_3 would be underestimated. For the exposure assessment which estimates health outcomes that occur at O_3 concentrations above

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
					60, this would lead to an underestimation of risks. For the epidemiology-based risk assessment which is effected by the entire range of O_3 concentrations, the impact is undetermined since changes at both ends of the O_3 distribution in opposite directions would affect the results. The opposite would be true in cases where the use of the central tendency of response increased the total estimated emissions reductions required to achieve a given standard.
F. Applying modeled sensitivities to un-modeled time periods	Relationships between O_3 response and hourly O_3 concentration were developed based on 8 months of modeling: January and April-October 2007. These relationships were applied to ambient data from 2006-2010. Some locations monitor for months not included in this modeling (February, March, November, and December) while others do not.	Both	Low-Medium	Low-Medium	KB: The eight months that were modeled capture a variety of meteorological conditions. In cases where other years have more frequent occurances of certain types of conditions, the regressions should be able to account for this. For instance, if a monitor only had 2-3 high O_3 days associated with sunny, high pressure conditions in the 2007 modeling but had 30-40 of those days in another year, the regression may be more uncertain at those high O_3 values but should still be able to capture the central tendency which can be applied to the more frequent occurances in other years. If, on the other hand, the meteorology/O_3 conditions in another year were completely outside the range of conditions captured in the model, then the regression based on modeled conditions might not be able to capture those conditions. KB: If emissions change drastically

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		between the modeled period and the time of the ambient data measurements this could also change the relationship between O_3 response and O_3 concentrations. The regressions derived from the 2007 modeling period are only applied to measurements made within 3 years of the modeled time period. Although some emissions changes did occur over this time period, we believe it is still reasonable to apply 2007 modeling to this relatively small window of measurements which occurs before and after the modeling. INF: Seasonal relationships were developed between O_3 response to emissions reductions and O_3 concentration. Summer was the only season for which modeling data was created for all months (June, July, August). The winter relationships were developed based on January modeling, the spring relationships were developed based on April/May modeling, and the autumn relationships were developed based on September/October modeling. The reduction in data points (31 or 61 instead of ~90) increases uncertainty in the statistical fit for these months. In addition, the modeling generally showed more O_3 disbenefits to NOx decreases in cooler months. So applying April/May relationships to March and September/October relationships to November could potentially underestimate O_3 increases that would happen in those two months in the five

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
					urban case study areas which measure O_3 during March and/or November: Dallas, Denver, Houston, Los Angeles, and Sacramento.
G. Assumptions of across-the-board emissions reductions	O_3 response is modeled for across-the-board reductions in U.S. anthropogenic NOx (and VOC). These across-the-board cuts do not reflect actual emissions control strategies.	Both	Low-Medium	Low-Medium	KB: The form, locations, and timing of emissions reductions that would be undertaken to meet various levels of the O_3 standard are unknown. The across-the-board emissions reductions bring levels down uniformly across time and space to show how O_3 would respond to changes in ambient levels of precursor species but do not reflect spatial and temporal heterogeneity that may occur in local and regional emissions reductions.
H. Assumption of equal percentage NOx and VOC reductions	In cases where VOC reductions were modeled, equal percentage NOx and VOC reductions were applied in the adjustment methodology.	Both	Low-Medium	Medium	KB: NOx reductions are likely to be the primary means used to reduce high O_3 concentrations at DV monitors. In limited cases, VOC emissions reductions may also help lower high O_3 concentrations at these locations. NOx and VOC reductions may be applied in combination but are unlikely to be applied in equal percentages. The available modeling constrained the NOx/VOC case to this unrealistic scenario. The equal percentage NOx and VOC reductions scenario does not optimize the lowest cost or least total emissions combinations as state and local agencies will likely attempt to achieve.
I. Downscaler	Downscaler combines monitored and modeled concentrations to produce a “fused” air quality	Both	Low-Medium	Low-Medium	KB: Downscaler combines modeled and monitored concentrations to provide estimates of O_3 concentrations in unmonitored areas while correcting model biases near monitors. The cross-

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
	surface. Uncertainties may occur in sparsely monitored regions, or in urban areas with dense monitoring networks and large spatial gradients.				validation analysis in Appendix 4-A shows that Downscaler generally gives more accurate estimates of air quality in monitored locations than either the monitored or modeled values alone. However, it is not possible to quantify the uncertainty associated with the estimates in unmonitored locations. KB: The air quality surfaces modeled by Downscaler for the national-scale risk assessment were seasonal average concentrations, which tend to have smaller spatial gradients than other metrics such as peak concentrations, and thus less uncertainty. INF: The cross-validation analysis in Appendix 4-A also shows that Downscaler tends to over-estimate low concentrations and under-estimate high concentrations. The mean bias in the estimates in monitored locations is nearly zero, but monitor locations are often chosen to capture the highest concentrations, thus there might be some bias towards higher concentrations in umonitored areas.

* Refers to the degree of uncertainty associated with our understanding of the phenomenon, in the context of assessing and characterizing its uncertainty. Sources classified as having a "low" impact would not be expected to impact the interpretation of risk estimates in the context of the O_3 NAAQS review; sources classified as having a "medium" impact have the potential to change the interpretation; and sources classified as "high" are likely to influence the interpretation of risk in the context of the O_3 NAAQS review.

4.6 REFERENCES

Abt Associates, Inc. 2010a. "Environmental Benefits and Mapping Program (Version 4.0)." Bethesda, MD. Prepared for U.S. Environmental Protection Agency Office of Air Quality Planning and Standards. Research Triangle Park, NC. Available on the Internet at <http://www.epa.gov/air/benmap>.

Abt Associates, Inc. 2010b. "Model Attainment Test Software (Version 2)." Bethesda, MD. Prepared for the U.S. Environmental Protection Agency Office of Air Quality Planning and Standards. Research Triangle Park, NC. Available on the Internet at: http://www.epa.gov/scram001/modelingapps.mats.htm.

Bell, M.L.; A. McDermott; S.L. Zeger; J.M. Samet and F. Dominici. 2004. "Ozone and Short-term Mortality in 95 U.S. Urban Communities," 1987-2000. *JAMA*, 292:2372-2378.

Berrocal, V.J.; A. E. Gelfand and D.M. Holland. 2012. "Space-Time Data Fusion Under Error in Computer Model Output: An Application to Modeling Air Quality." *Biometrics*, 68(3), 837-848.

Chen, J. R.; Zhao; Z. Li. 2004. "Voronoi-based k-order Neighbor Relations for Spatial Analysis." *ISPRS J Photogrammetry Remote Sensing*, 59(1-2), 60-72.

Duff, M.; R. L. Horst; T.R. Johnson, 1998. "Quadratic Rollback: A Technique to Model Ambient Concentrations Due to Undefined Emission Controls." San Diego, CA: Presented at the Air and Waste Management Annual Meeting, June 14-18, 1998.

Fann, N.; A.D. Lamson; S.C. Anenberg; K. Wesson; D. Risley; B.J. Hubbell. 2012. "Estimating the National Public Health Burden Associated with Exposure to Ambient $PM_{2.5}$ and Ozone." *Risk Analysis*, 32:81-95.

Gold, C. 1997. "Voronoi Methods in GIS," Vol. 1340. In *Algorithmic Foundation of Geographic Information Systems* (Kereveld M., J. Nievergelt, T. Roos, P. Widmayer eds). Lecture notes in *Computer Science*, Berlin: Springer-Verlag, 21-35.

Hall, E.; A. Eyth; S. Phillips. 2012. "Hierarchical Bayesian Model (HBM)-Derived Estimates of Air Quality for 2007: Annual Report." (EPA document number EPA/600/R-12/538). < http://www.epa.gov/heasd/sources/projects/CDC/AnnualReports/2007_HBM.pdf>.

Jerrett, M.; R.T. Burnett; C.A. Pope, III; K. Ito; G. Thurston; D. Krewski; Y. Shi, E. Calle and M. Thun. 2009. "Long-term O_3 Exposure and Mortality." *N. Eng. J. Med.*, 360:1085-1095.

Johnson, T. 2002. "A Guide to Selected Algorithms, Distributions, and Databases Used in Exposure Models Developed by the Office of Air Quality Planning and Standards," prepared by TRJ Environmental, Inc. for the U.S. Environmental Protection Agency. Research Triangle Park, NC: Office of Research and Development.

National Research Council of the National Academies. 2008. "Estimating Mortality Risk Reduction and Economic Benefits from Controlling Ozone Air Pollution." Washington, DC: *The National Academies Press*.

Rizzo, M. 2005. "A Comparison of Different Rollback Methodologies Applied to Ozone Concentrations," posted on November 7, 2005, http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_cr_td.html.

Rizzo, M. 2006. "A Distributional Comparison between Different Rollback Methodologies Applied to Ambient Ozone Concentrations," posted on May 31, 2006, http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_cr_td.html.

Simon, H.; K. Baker; N. Possiel; F. Akhtar; S. Napelenok; B. Timin; B. Wells. 2012. "Model-based Rollback Using the Higher Order Direct Decoupled Method (HDDM)," posted at < http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_td.html>.

Simon, H.; K. R. Baker; F. Akhtar; S.L. Napelenok; N. Possiel; B. Wells and B. Timin. 2013. "A Direct Sensitivity Approach to Predict Hourly Ozone Resulting from Compliance with the National Ambient Air Quality Standard" *Environmental Science and Technology*, Vol. 47, 2304-2313.

Smith, R.L., B. Xu, P. Switzer. 2009b. "Reassessing the Relationship Between Ozone and Short-term Mortality in U.S. Urban Communities," *Inhale Toxicol*, Vol. 21: 37-61.

Timin, B.; K. Wesson and J. Thurman. 2010. "Application of Model and Ambient Data Fusion Techniques to Predict Current and Future Year $PM_{2.5}$ Concentrations in Unmonitored Areas, " in D.G. Steyn and St Rao (eds), *Air Pollution Modeling and Its Application XX*, Netherlands: Springer, pp. 175-179.

U.S. Environmental Protection Agency. 2007. *Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information OAQPS Staff Paper*. Washington, DC: EPA Office of Air and Radiation. (EPA document number

U. S. EPA. 2012a. *Integrated Science Assessment for Ozone and Related Photochemical Oxidants: Third External Review Draft.* Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-07-007; EPA/600/R-10/076C).

U.S. EPA. 2012b. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume I: User's Guide*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001a). <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2012c. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume II: Technical Support Document*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001b). <http://www.epa.gov/ttn/fera/human_apex.html>.

Wells, B.; K. Wesson and S. Jenkins. 2012. "Analysis of Recent U.S. Ozone Air Quality Data to Support the O_3 NAAQS Review and Quadratic Rollback Simulations to Support the First Draft of the Risk and Exposure Assessment." <http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_td.html>.

Zhang, L.; D.J. Jacob; N.V. Smith-Downey; D.A. Wood; D. Blewitt; C.C. Carouge; A. van Donkelaar; D.B. A. Jones; L.T. Murray and Y. Wang. 2011. "Improved Estimate of the Policy-relevant Background Ozone in the United States Using the GEOS-Chem Global Model with 1/2°x2/3° Horizontal Resolution Over North America." *Atmos Environ*, Vol. 45, pp. 6769-6776.

Zanobetti, A. and J. Schwartz. 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*, 177:184-189.

5 CHARACTERIZATION OF HUMAN EXPOSURE TO OZONE

5.0 OVERVIEW

As part of the previous 2007 O_3 NAAQS review, EPA staff conducted exposure analyses for the general population, all school-age children (ages 5-18), all active school-age children,[1] and asthmatic school-age children (U.S. EPA, 2007a,b). Exposure estimates were generated for 12 urban study areas[2] for recent years of air quality and for just meeting the existing 8-hr standard and several alternative 8-hr standards. EPA also conducted a health risk assessment that produced risk estimates for the number and percent of all school-age children experiencing impaired lung function and other respiratory symptoms associated with the exposures estimated for these same 12 study areas.

The exposure analysis conducted for this current NAAQS review builds upon the methodology and lessons learned from the exposure analyses conducted in previous O_3 reviews (U.S. EPA, 1996a, 2007a,b) and information provided in the final ISA (U.S. EPA, 2013). Here, we estimate exposures for people residing in 15 urban study areas in the U.S.[3] The population exposures to ambient O_3 concentrations were modeled using EPA's Air Pollutants Exposure (APEX) (US EPA, 2012a,b). Exposures were calculated considering O_3 concentrations in recent years, using 2006 to 2010 spatially interpolated ambient monitoring data. Exposures were also estimated considering alternative air quality scenarios, that is, where O_3 concentrations just meet the existing 8-hr O_3 NAAQS and at several other standard levels considering the same indicator, form, and averaging time, based on adjusting data as described in Chapter 4. Exposures were modeled for 1) all school-age children (ages 5-18), 2) asthmatic school-age children (ages 5-18), 3) asthmatic adults (ages 19-95), and 4) all older adults (ages 65-95), each while at moderate or greater exertion level at the time of exposure.[4] The strong emphasis on children, asthmatics, and older adults reflects the finding of the last O_3 NAAQS review (U.S. EPA, 2007a) and the ISA (U.S. EPA, 2013, Chapter 8) that these are important at-risk groups. Exposure model output of interest for this chapter are the percent (and number) of persons exposed at or above 8-hr average

[1] In the previous 2007 exposure assessment, a study group of active school-age children was identified as children having their median daily physical activity index (PAI) over the exposure period ≥ 1.75, an activity level characterized by exercise physiologists as being "moderately active" or "active" (McCurdy, 2000).

[2] The twelve study areas evaluated in the 2007 exposure assessment were Atlanta, Boston, Chicago, Cleveland, Detroit, Houston, Los Angeles, New York, Philadelphia, Sacramento, St. Louis, Washington DC (an area which at that time was modeled to include Baltimore as part of the Baltimore-Northern Virginia MSA).

[3] In addition to the twelve study areas identified in the 2007 exposure assessment, staff has added Dallas and Denver, while also separately modeling Baltimore (from Washington DC) in this current assessment. Inclusion of Seattle, WA was considered but not included due to a lack of appropriate monitoring data.

[4] The "all school-age children" study group includes both asthmatic and non-asthmatic children ages 5 to 18. The "all older adults" includes both asthmatic and non-asthmatic older adults ages 65 to 95. Note also that the 8-hr average exposure of interest in both this and the previous assessment was concomitant with moderate or greater exertion for all study groups.

O_3 concentrations of concern, all while at moderate or greater exertion levels, based on adverse effects observed in human clinical exposure studies. Further, the complete time series of individual exposures estimated by APEX serves as input to a module that estimates human health risk (Chapter 6).

This chapter first provides a brief overview of human exposure and exposure modeling using APEX (section 5.1), the scope of this O_3 exposure assessment and key inputs used to model exposure in the 15 U.S. study areas selected (section 5.2), and followed by the main body exposure results (section 5.3). Then, section 5.4 presents an assemblage of targeted analyses designed to provide additional insight to the main body of exposure results by focusing on important data inputs, additional at-risk populations, lifestages, or scenarios, influential attributes in estimating exposures, and performance evaluations. The results of these and other exposure model targeted analyses are integrated in an uncertainty characterization section (section 5.5) along with a final section summarizing the key observations for this chapter (section 5.6).

5.1 SYNOPSIS OF O_3 EXPOSURE AND EXPOSURE MODELING

5.1.1 Human Exposure

Human exposure to a contaminant is defined as "contact at a boundary between a human and the environment at a specific contaminant concentration for a specific interval of time," and has units of concentration times duration (National Research Council, 1991). For air pollutants the contact boundary is nasal and oral openings in the body, and *personal exposure* of any individual to a chemical in the air for a discrete time period is fundamentally quantified as (Lioy, 1990; National Research Council, 1991):

$$E_{[t_1, t_2]} = \int_{t_1}^{t_2} C(t)dt \qquad (5\text{-}1)$$

where $E_{[t_1,t_2]}$ is the personal exposure or *exposure concentration* during the time period from t_1 to t_2, and $C(t)$ is the concentration at time t in the breathing zone. The breathing rate at the time of exposure will influence the dose received by the individual. While we do not directly estimate dose in this assessment, *intake* is the total O_3 inhaled (i.e., exposure concentration, duration, and ventilation combined).[5]

[5] In chapter 6, the estimation of risk combines the time series of both the personal exposure concentrations and ventilation rate, among other variables in essentially calculating a dose, though not explicitly output from the model.

5.1.2 Estimating O_3 Exposure

Exposure to O_3 can be directly estimated by monitoring the concentration of O_3 in a person's breathing zone (close to the nose/mouth) using a personal exposure monitor. Studies employing this measurement approach have been reviewed in the ISA and EPA O_3 Air Quality Criteria Documents (U.S. EPA, 1986, 1996b, 2006, 2013). Personal exposure measurements from these studies are useful in describing a general range of exposure concentrations (among other reported measurement data) and in identifying factors that may influence varying exposure levels. However, these measurement studies are largely limited by the disparity between sample measurement duration and exposure concentration averaging-times of interest and in appropriately capturing variability in population exposure occurring over large geographic areas , particularly when considering both concentration (e.g., spatial variability) and population (e.g., age, sex) attributes that influence exposure.

O_3 exposure for individuals, small groups of individuals or large populations can be calculated indirectly (or *modeled*) using Equation 5-1. When employing such an approach in a population exposure assessment, two basic types of input data are needed; a time-series of O_3 concentrations that appropriately represents spatial heterogeneity in O_3 concentrations and a corresponding time-series of locations visited by the persons exposed. When considering air pollutant concentrations, population exposure models are commonly driven by ambient concentrations. These ambient concentrations may be provided by monitoring data, by air quality model estimates, or perhaps by a combination of these two data sources. Then, an understanding of the relationships between ambient pollutants and the locations people occupy is needed. This is because human exposure, regardless of the pollutant or whether one is interested in individual or population exposure, depends on where an individual is located, how long they occupy that location, and what the pollutant concentration at the point of contact is. Furthermore, if interested in air pollutant intake rate or dose, one needs to know what activity the person is performing while exposed.

Thus, the types of measurement and modeling studies that provide information for more realistically estimating exposure to O_3 can be augmented from the above list to include studies of: 1) O_3 formation, deposition, and decay, 2) people's locations visited and activities performed, 3) human physiology, and 4) local scale meteorological measurements and/or modeling. Useful data derived from these varied studies are O_3 concentrations (i.e., fixed site, personal exposure, indoor and outdoor locations), built environment physical factors (i.e., air exchange rates (AERs), infiltration rates, decay and deposition rates), human time-location-activity patterns (minute-by-minute, hourly, daily, and longer-term), time-averaged or activity-specific breathing rates among varying sexes and/or lifestages, and hourly ambient temperatures.

When integrating these varied data (among others such as population demographics and disease prevalence) and understanding factors affecting exposure, exposure models can extend beyond the limited information given by measurement data alone. For example, an exposure model can reasonably estimate exposures for any perceivable at-risk population (e.g., asthmatics living in a large urban area) and considering any number of hypothetical air quality conditions (e.g., just meeting a daily maximum 8-hr average concentration of 70 ppb). Exposure models that account for variability in human physiology can also realistically estimate pollutant intake by using activity-specific ventilation rates. These types of measurements cannot realistically be performed for a study group or population of interest, particularly when considering time, cost, and other constraints. The following section provides an overview of how such exposure modeling can be done using APEX, the model developed by EPA to perform such calculations and used to estimate O_3 exposures in this REA.

5.1.3 Modeling O_3 Exposure Using APEX

EPA has developed the APEX model for estimating human population exposure to criteria and air toxic pollutants, used most recently in estimating exposures for the O_3 (U.S. EPA, 2007b), nitrogen dioxide (U.S. EPA, 2008), sulfur dioxide (U.S. EPA, 2009a), and carbon monoxide (U.S. EPA, 2010) NAAQS reviews. APEX is a probabilistic model designed to account for the numerous sources of variability that affect people's exposures. An overview of the approaches used by APEX to estimate exposure concentrations is found in Appendix 5A with details provided in U.S. EPA (2012a,b).

Briefly, APEX simulates the movement of individuals through time and space and estimates their exposure to a given pollutant while occupying indoor, outdoor, and in-vehicle locations. The model stochastically generates simulated individuals in selected study areas using census-derived probability distributions for demographic characteristics. Population demographics are drawn from the 2000 Census data[6] at a tract level, and a national commuting database based on 2000 Census data provides home-to-work commuting flows between tracts.[7] Any number of individuals can be simulated, and collectively they approximate a random sampling of people residing in a particular study area.

Daily activity patterns for individuals in a study area, an input to APEX, are obtained from detailed daily time-location-activity pattern survey data that are compiled in the Consolidated Human Activity Database (CHAD) (McCurdy et al., 2000; U.S. EPA, 2002). These daily diaries are used to construct a sequence of locations visited and activities performed for APEX simulated individuals consistent with their demographic characteristics, day-type (e.g.,

[6] Due to resource limitations and data availability, the 2010 Census data have not yet been processed to include in this 2nd draft REA.

[7] There are approximately 65,400 census tracts in the ~3,200 counties in the U.S.

weekend or weekday), and season of the year, as defined by ambient temperature regimes (Graham and McCurdy, 2004). The time-location-activity data input to APEX are linked with personal attributes of the surveyed individuals' such as age, sex, employment status, day-of-week surveyed, and daily maximum and daily mean temperature. These specific personal attribute data are then used by APEX to best match the daily diary with the simulated persons of interest, using the same variables as first-order diary selection characteristics. The approach is designed to capture the important attributes contributing to an individuals' time-location-activity pattern, and of particular relevance here, time spent outdoors (Graham and McCurdy, 2004). In using a diverse collection of time-location-activity diaries that capture the duration and frequency of occurrence of visitations/activities performed, APEX can simulate expected variability in human behavior, both within and between individuals. This, combined with exposure concentrations, allows for the reasonable estimation of the magnitude, frequency, pattern, and duration of exposures an individual experiences.

A key concept in modeling exposure using APEX is the *microenvironment*, a term that refers to the immediate surroundings of an individual at a particular time. APEX has a flexible approach for modeling micro-environmental concentrations whereas the model user defines the type, number and characteristics of the microenvironments to be modeled. Typical microenvironments include indoors at home, indoors at school, near roadways, inside cars, and outside home. In this exposure assessment, all microenvironmental O_3 concentrations are derived from ambient O_3 concentrations input to APEX and are estimated using either a *mass-balance* or *transfer factors* approach, selected by the user. The mass balance approach assumes that the air in an enclosed microenvironment is well-mixed and that the air concentration is spatially uniform at a given time within the microenvironment. The approach employs indoor-to-outdoor AERs (i.e., number of complete air exchanges per hour) and considers removal mechanisms such as deposition to building surfaces and chemical decay rates. The transfer factors model is simpler than the mass balance model, and employs two variables, a *proximity factor*, used to account for proximity of the microenvironment to sources or sinks of pollution, or other systematic differences between concentrations just outside the microenvironment and the ambient concentrations, and a *penetration factor*, which quantifies the degree to which the outdoor air penetrates into the microenvironment.

Activity-specific simulated breathing rates of individuals are used in APEX to characterize intake received from an exposure. This is done because controlled human exposure studies have shown adverse health outcomes are associated with both elevated concentrations and study participant exertion levels. The breathing rates calculated by APEX are derived from the energy expenditure associated with each simulated persons' activity performed, adjusted for age- and sex-specific physiological parameters (Graham and McCurdy, 2005). The energy

expenditure estimates themselves are derived from distributions of METS[8] (or metabolic equivalents of work) associated with every activity performed (McCurdy et al., 2000, using Ainsworth et al., 1993).

An important feature of APEX is the ability to account for variability in exposure by representing input variables as statistical distributions along with dependent conditional variables, where appropriate. For example, the distribution of AERs in a home, office, or motor vehicle can depend on the type of heating and air conditioning present, which are also stochastic inputs to the model, as well as the ambient temperature on a given day. The user can choose to keep the value of a stochastic parameter constant for the entire simulation (appropriate for the volume of a house), or can specify that a new value shall be drawn hourly, daily, or seasonally from specified distributions.

Finally, APEX calculates a unique time-series of exposure concentrations on the order of minutes or smallest diary event duration that each simulated person may experience during the modeled time period, based in that individual's estimated microenvironmental concentrations and the time spent in each of sequence of microenvironments visited according to the time-location-activity diary of each individual. Then, hourly average exposures of each simulated individual are estimated using time-weighted averages of the within-hour exposures. From hourly exposures, APEX calculates any other time averaged exposure of interest (e.g., 8-hr or daily average) that a simulated individual experiences during the modeled period.

5.2 SCOPE OF THE EXPOSURE ASSESSMENT

This section broadly presents the scope of the exposure assessment including descriptions of the modeling domains, ambient concentrations used, time periods and populations modeled, as well as identifying key approaches, inputs and outputs used by APEX in estimating population O_3 exposures. Detailed descriptions regarding APEX modeling, model inputs and other supporting information are provided in Appendix 5A-5E and the APEX user's guide and technical support documents (U.S. EPA 2012a,b). Figure 5-1 illustrates the general conceptual framework for generating our population exposure concentrations, including the time series of exposure and ventilation rate output generated as input to population risk calculations in Chapter 6.

5.2.1 Urban Areas Selected

The selection of urban areas to include in the exposure assessment considered the location of O_3 epidemiological studies, the availability of ambient O_3 monitoring data, and the

[8] METS are a dimensionless ratio of the activity-specific energy expenditure rate to the basal or resting energy expenditure rate. The metric is used by exercise physiologists and clinical nutritionists to estimate work undertaken by individuals as they go through their daily activities (Montoye et al., 1996).

desire to represent a range of geographic areas, encompassing variability in climate and population demographics. Specifically, the criteria included the following:

- The overall set of urban locations should represent a range of geographic areas, urban population demographics, and climate, beginning with study areas selected in the 2007 O_3 NAAQS review.
- The locations should be focused on areas that do not meet or are close to not meeting the existing 8-hr O_3 NAAQS and should include areas having O_3 non-attainment designations.
- There must be sufficient O_3 ambient air quality data for the recent 2006-2010 period.
- The study areas should include the 12 cities modeled in the epidemiologic-based risk assessment (Chapter 7).

Based on these criteria, we chose the 15 study areas listed in Table 5-1 to develop our population exposure estimates. We then defined an *air quality domain* for each study area, broadly bounding the ambient concentration field where exposures were to be estimated. To do this, we evaluated 1) counties modeled in the previous 2007 O_3 NAAQS review common to current study areas, 2) political/statistical county aggregations (e.g., whether in a metropolitan statistical areas or MSAs), and 3) if the study area was designated as a non-attainment area (NAA), the counties that were part of the NAA list. We identified a final list of 215 counties[9] to comprise the air quality domain for the 15 study areas, the names of which are provided in Appendix 5B.

5.2.2 Time Periods Simulated

The exposure periods modeled are the O_3 seasons for which routine hourly O_3 monitoring data were available for years 2006 to 2010 (Table 5-1), and defined by 40 CFR part 58, Appendix D, Table D-3. These periods are designed to reasonably capture year-to-year variability in ambient concentrations and meteorology and include most of the high concentration events occurring in each area. Having this wide range of air quality data across multiple years allows us to more realistically estimate a range of exposures, rather than using a single year of air quality data. While the number of available O_3 monitors may vary slightly from year to year, we assumed constant representation by the available monitors and associated statistically interpolated data for each year over the simulation period (see section 5.2.3).

[9] Of the 215 counties defining the air quality domain, 207 remained in the exposure model domain.

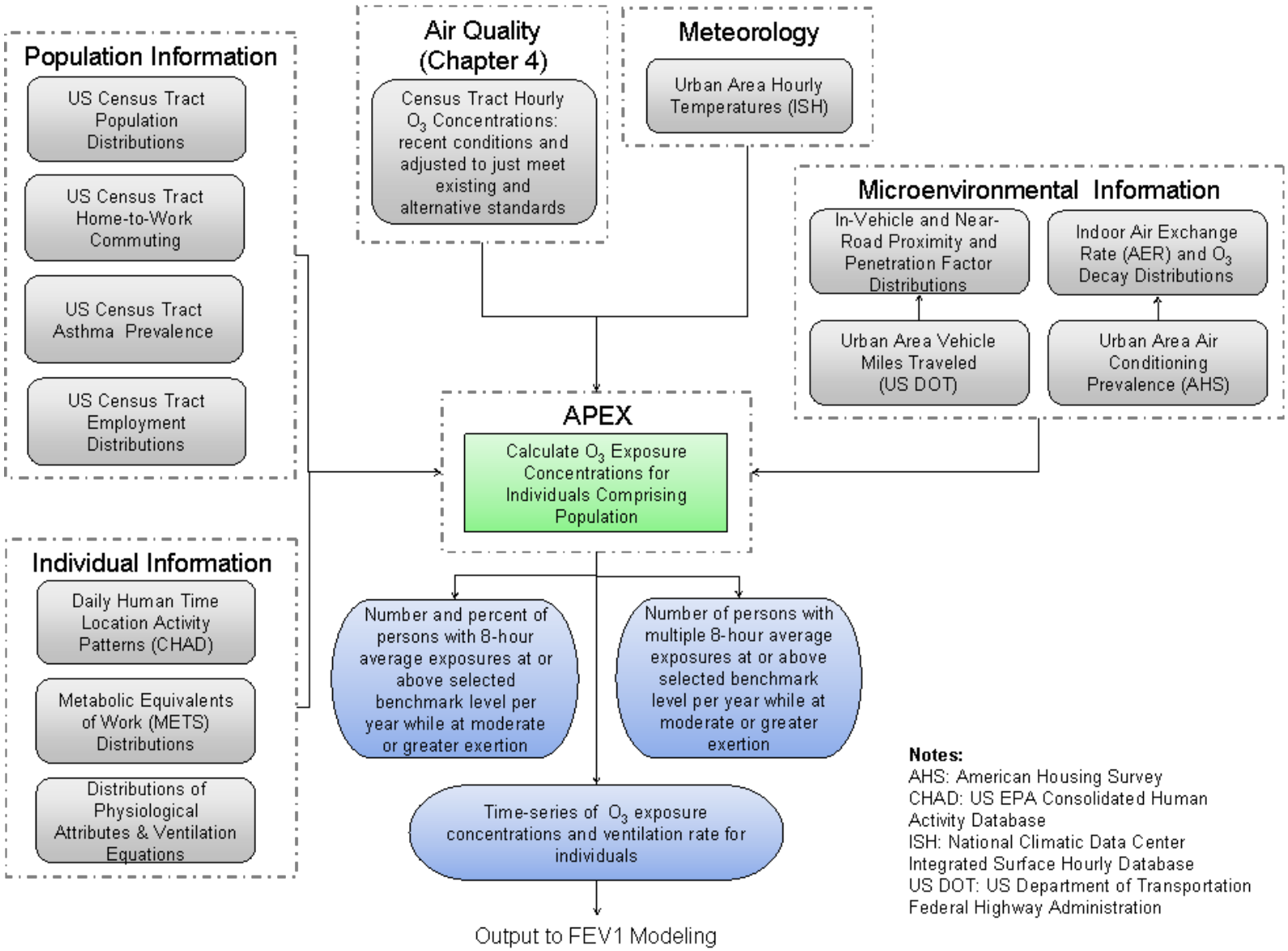


Figure 5-1 Conceptual Framework Used for Estimating Study Area Population O_3 Exposure Concentrations

Table 5-1 General Characteristics of the Population Exposure Modeling Domain Comprising Each Study Area

Study Area (Abbreviation)	O_3 Season[1]	Study Area Number of:					
						Persons	
		Counties	Ambient Monitors	APEX Air Districts	US Census Tracts	All School Age Children (age 5-18)	All Ages (age 5-95)
Atlanta (ATL)	Mar 1-Oct 31	32	14	664	678	860,649	3,850,951
Baltimore (BAL)	Apr 1-Oct 31	7	12	603	618	505,140	2,209,226
Boston (BOS)	Apr 1-Sep 30	7	13	1,005	1,028	905,208	4,449,291
Chicago (CHI)	Apr 1-Oct 31	16	28	1,882	2,055	1,899,073	8,345,373
Cleveland (CLE)	Apr 1-Oct 31	8	16	802	879	578,733	2,692,846
Dallas (DAL)	Mar 1-Oct 31	11	21	1,012	1,036	1,097,004	4,698,392
Denver (DEN)	Mar 1-Sep 30	12	25	655	675	560,137	2,626,239
Detroit (DET)	Apr 1-Sep 30	9	13	1,419	1,454	1,016,896	4,572,479
Houston (HOU)	Jan 1-Dec 31	10	19	779	802	970,528	3,925,054
Los Angeles (LA)	Jan 1-Dec 31	5	50	2,000	3,352	3,620,972	14,950,340
New York (NY)	Apr 1-Oct 31	27	32	1,900	4,889	3,843,450	18,520,868
Philadelphia (PHI)	Apr 1-Oct 31	15	19	1,452	1,555	1,231,052	5,506,954
Sacramento (SAC)	Jan 1-Dec 31	7	18	447	461	466,169	1,926,598
St. Louis (STL)	Apr 1-Oct 31	15	16	494	518	527,755	2,340,325
Wash., DC (WAS)	Apr 1-Oct 31	26	28	1,013	1,037	966,791	4,498,374
All Study Areas	-	**207**	**324**	**16,127**	**21,037**	**19,049,557**	**85,113,310**

[1] Each study area's O_3 monitoring season is defined by 40 CFR part 58, Appendix D, Table D-3.

5.2.3 Ambient Concentrations Used

We used the available hourly ambient monitor concentration data within and around each study area along with a statistical interpolation technique (Chapter 4) to estimate hourly census tract concentrations within the counties comprising each study area. These concentrations served as the 'base' air quality input for each study area year. Ambient concentrations were also adjusted to just meet the existing standard (75 ppb, 4th highest 8-hr average, averaged over a 3-year period) and alternative standard levels (70, 65, 60, and 55 ppb) using an air quality model and the statistical interpolation technique (Chapter 4).

These estimated hourly census tract O_3 concentrations served as the APEX *air districts*, the basic ambient concentrations from which each simulated persons microenvironmental concentrations are estimated. Having these temporally and spatially resolved air districts in each

study area allows for better utilization of APEX spatial and temporal capabilities in estimating exposure. Because APEX simulates where individuals are located at specific times of the day, more realistic exposure estimates are obtained in simulating the contact of individuals with these spatially and temporally diverse concentrations.

Even though we estimated O_3 ambient concentrations at all census tracts in each county-level study area, the study area *exposure modeling domain* was defined as a subset of these census tracts by using the ambient monitoring sites within the urban core of each study area's air quality domain and a 30 km radius of influence. This zone of influence is consistent with what was done in the 1st draft O_3 REA, though in that exposure assessment, only the ambient monitoring data sites themselves were used to represent the APEX air districts, hence concentrations measured at a particular monitoring site would be directly extrapolated outwards to all census tracts within 30 km of that site. In contrast, by incorporating the VNA estimated concentrations and retaining the same 30 km radius of influence, we are stressing the significance of the monitor information in defining the urban core air quality while also reasonably estimating concentration gradients (where such gradients exist) with increasing distance from monitoring locations.

Thus, all air districts[10] and census tracts that fall within the 30 km radius of each ambient monitor were used to estimate the exposures, defining the final exposure modeling domain in each study area (Table 5-1). The monitor IDs used to select the census tracts to be modeled are provided in Appendix 5B, while the complete list of census tract IDs where exposures are modeled are within the APEX control files for each study area (and are the same for each simulation year).

5.2.4 Meteorological Data Used

APEX uses study area temperature data to select representative diaries for a particular day and in selecting an appropriate air exchange rate used to calculate indoor residential microenvironmental concentrations. APEX uses the data from the closest weather station to each Census tract. To ensure reasonable coverage for each study area, a few to several meteorological stations recording hourly surface temperature measurements were identified using data obtained from the National Weather Service ISH data files.[11] Details regarding the meteorological stations selected and data processing are given in Appendix 5B. Briefly, APEX requires the temperature input data to be 100% complete. In general, any missing values were filled using a linear

[10] The original number of air quality districts for New York and Los Angeles needed to be reduced by about half due to exceeding personal computer memory capacity when APEX used > 2,000 air districts. See Appendix 5B for details.

[11] http://www.ncdc.noaa.gov/oa/climate/surfaceinventories.html

interpolation or regression approach that employs information from proximal meteorological stations.

5.2.5 Populations Simulated

Exposure was estimated for four at-risk study groups residing in each study area: all school-age children (ages 5-18), asthmatic school-age children, asthmatic adults (ages 19-95), and all older adults (ages 65-95). Due to the increased amount of time spent outdoors engaged in relatively high levels of physical activity (which increases intake), school-age children as a group are particularly at risk for experiencing O_3-related health effects (U.S. EPA, 2013, Chapter 8). We report results for all school-age children down to age five, recognizing an increasing trend for younger children to attend school. Some U.S. states allow 4-year-olds to attend kindergarten, and most states have preschool programs for children younger than five. In 2000, six percent of U.S. children ages 3 to 19 who attend school were younger than five years old (2000 Census Summary File 3, Table QT-P19: School Enrollment). Currently we do not estimate exposure for these younger children due to a lack of information that would let us confidently characterize these younger aged children. While EPA guidance recommends, for certain instances, an upper age group of children ages 16 through 21 (U.S. EPA, 2005), we restricted our upper age classification of children through age 18. In considering the expected variability in activity patterns over the span of ages 16 through 21 (e.g., time spent outdoors, time in school, each in contrast to time spent working) and the relatively small difference in respiratory physiology over that same age span compared with that of adults (e.g., Figure 5-17), factors critical for high O_3 exposure and dose, we assumed simulated persons age 19 to 21 would be best included in our adult study group. The number of persons represented in each of the 15 study areas is given in Table 5-1 and, considering all study areas together, captures approximately 32.8 % of all children ages 5 to18 and 32.0 % of the total U.S. population ages 5 to 95.

The number of asthmatic school-age children and asthmatic adults in each study area was estimated using asthma prevalence from the Center for Disease Control (CDC) and Prevention's National Health Interview Survey (NHIS).[12] Briefly, years 2006-2010 NHIS survey data were combined to calculate asthma prevalence, defined as the probability of a "*Yes*" response to the question: "*do you still have asthma?*" among those that responded "*Yes*" to the question "*has a doctor ever diagnosed you with asthma?*". The asthma prevalence was first stratified by NHIS defined regions (Midwest, Northeast, South, and West), sex, age (single years for ages 0-17) or age groups (ages ≥ 18), and by a family income/poverty ratio.[13] These new asthma prevalence estimates were then linked to U.S. census tract level poverty ratio probabilities (U.S. Census

[12] See http://www.cdc.gov/nchs/nhis.htm (accessed October 4, 2011).

[13] The income/poverty ratio threshold used was 1.5, that is the surveyed person's family income was considered either ≤ or > than a factor of 1.5 of the U.S. Census estimate of poverty level for the given year.

Bureau, 2007), also stratified by age and age groups, to generate a final database consisting of census tract level asthma prevalence for the entire U.S. A detailed description of how the data base was developed is presented in Appendix 5C, while the estimated asthma prevalence used for each census tract is provided in the APEX asthma prevalence input file. A summary of the asthma prevalence calculated for each study area simulation is provided here in Table 5-2.

Table 5-2 Asthma Prevalence for Children and Adults Estimated by APEX in Each Simulated Study Area

Study Area	Asthma Prevalence (%)		
	Children (5-18)	Adults (18-95)	All Persons (5-95)
Atlanta	9.6	6.5	7.2
Baltimore	9.7	6.6	7.3
Boston	11.4	7.9	8.6
Chicago	10.7	7.8	8.4
Cleveland	10.9	7.7	8.4
Dallas	9.9	6.5	7.3
Denver	8.9	7.7	7.9
Detroit	11.1	7.7	8.5
Houston	10.1	6.5	7.4
Los Angeles	9.0	7.7	8.0
New York	12.2	8.1	9.0
Philadelphia	11.3	7.9	8.7
Sacramento	9.0	7.8	8.1
St. Louis	11	7.6	8.4
Washington DC	9.5	6.4	7.1
All Areas	**10.5**	**7.6**	**8.2**

All simulated persons (either asthmatic or non-asthmatic) used time-location-activity data from CHAD, the most complete, high quality source of human activity data for use in exposure modeling. The current CHAD database contains over 53,000 individual daily diaries including time-location-activity patterns for individuals of both sexes across a wide range of ages (<1 to 94). The database is geographically diverse, containing diaries from individuals residing in several major cities, suburban, and rural areas across the U.S. Time spent performing activities within particular locations can be on a minute-by minute basis, thus avoiding the smoothing of potential peak exposures longer event durations would yield.

Table 5-3 summarizes the studies and number of diaries in CHAD used in this assessment, noting that the total CHAD diaries used by APEX is restricted to just over 41,000 given our simulation age range (5-95) and additionally selected usability requirements.[14] Additional context regarding the representativeness of the CHAD data in estimating exposure is provided in section 5.3.1 and Appendix 5G.

APEX creates a sequence of daily diaries across the entire O_3 season for each simulated individual using a method designed to capture the tendency of individuals to repeat activities, based on reproducing realistic variation in a key diary variable (Glen et al., 2008). For this O_3 analysis, the key variable selected is the amount of time an individual spends outdoors each day, one of the most important determinants of exposure to high levels of O_3 (see section 5.3.2). The longitudinal method targets two statistics, a population diversity statistic (*D*) and a within-person autocorrelation statistic (*A*). Values of 0.2 for *D* and 0.2 for *A* were initially developed based on analyses by Geyh et al. (2000) and Xue et al. (2004), with both studies evaluating groups of children ages 7 to 12 in a single study area. We adjusted values for *D* upwards to 0.5 to reflect a broader range of ages and to better estimate repeated activities.[15] Further details regarding the development of the longitudinal methodology can be found in U.S. EPA (2012a, b).

[14] In this assessment, the CHAD diaries must be from persons having a known age, sex, day-of-week, and daily temperature. In addition, diaries must have no more than 3 hours total of missing location and/or activity data.

[15] A small *D* means that the overall variability between people in the key diary statistic is smaller than the variability observed over days within the same person. A *D* closer to 1 means that each person shows little variation over time relative to the variability between persons.

Table 5-3 Consolidated Human Activity Database (CHAD) Study Information and Diary-days Used by APEX

Activity Pattern Study (Abbrev.)	General Study Area	Study Years	Subject Ages	Diary-days (ages 4-94)	Diary-days (ages 4-18)
Baltimore Retirement Home (BAL)	Baltimore, MD	1997-98	72 - 93	304	0
California Youth (CAY)	California	1987-88	12 - 17	182	182
California Adults (CAA)	California	1987-88	18 - 94	1,555	36
California Children (CAC)	California	1989-90	<1 - 11	771	771
Cincinnati (CIN)	Cincinnati, OH	1985	<1 - 86	2,259	727
Detroit Exposure and Aerosol Research (DEA)[1,2]	Detroit, MI	2005-06	18 - 74	331	5
Denver CO Personal Exposure (DEN)	Denver, CO	1982-83	18 - 70	714	7
EPA Longitudinal (EPA)[1,2]	RTP, NC	1999-2000, 2002, 06-08	<1 - 60	1,386	0
LA O_3 Exposure: Elementary School (LAE)	Los Angeles, CA	1989	10 - 12	50	50
LA O_3 Exposure: High School (LAH)	Los Angeles, CA	1990	13 - 17	42	42
National Human Activity Pattern Study: Air (NHA)	National	1992-94	<1 - 93	4,129	693
National Human Activity Pattern Study: Water (NHW)	National	1992-94	<1 - 93	4,099	745
National-Scale Activity Survey (NSA)	7 US metro. areas	2009	35 - 92	6,825	0
Population Study of Income Dynamics I (ISR)[1]	National	1997	<1 - 13	3,507	3,507
Population Study of Income Dynamics II (ISR)[1]	National	2002-03	5 - 19	4,800	4,793
Population Study of Income Dynamics III (ISR)[1,2]	National	2007-08	10 - 19	2,619	2614
RTI O_3 Averting Behavior (OAB)[1]	35 US metro. areas	2002-03	2 - 12	2,187	2,187
RTP Panel (RTP)[1]	RTP, NC	2000-01	55 - 85	871	0
Seattle (SEA)[1]	Seattle, WA	1999-2002	6 - 91	1,624	317
Study of Use of Products and Exposure Related Behavior (SUP)[1,2]	Sac/San Fran, CA Counties	2006-10	1 - 88	2,533	994
Washington, D. C. (WAS)	Wash., DC	1982-83	18 - 71	686	10
Totals		**1982 - 2010**	**<1 - 94**	**41,474**	**17,680**

[1] Study data added after 2007 O_3 NAAQS review.

[2] Study data added after 2012 1st Draft O_3 REA.

5.2.6 Key Physiological Processes And Personal Attributes Modeled

The modeling of physiological processes relevant to the O_3 exposure and intake is complex, particularly when representing inter- and intra-personal variability in energy expenditure (EE) and ventilation rates (VE). APEX has a module capable of estimating several variables associated with every activity performed by simulated individuals. Briefly, the module links the diary indicated activities to specific energy expended, the rate of oxygen consumed (VO_2) and the associated ventilation rate, all considering the unique sequence of events individuals go through each simulated day. The activity-specific time-series of VE estimates ultimately serve as an important variable used in estimating O_3 intake as well as in identifying when simulated individuals performing activities at moderate or greater exertion. In addition, age, sex, and body mass related physiological differences are specifically taken into account by the ventilation algorithm, derived using ventilation data obtained from several human studies (see Graham and McCurdy, 2005):

$$\ln(\dot{V}_E / BM) = b_0 + b_1 \ln(\dot{V}_{O2} / BM) + b_2 \ln(1 + age) + b_3\, sex + e_b + e_w \qquad (5\text{-}2)$$

where,

ln = natural logarithm of variable

$\dot{V}_E / BM$ = activity specific ventilation rate, body mass normalized (liter air/kg)

b_i = see Table 5-4

$\dot{V}_{O2} / BM$ = activity specific oxygen consumption rate, body mass normalized (liter/O_2/kg)

age = age of the individual (years)

sex = sex (-1 for males, +1 for females)

e_b = randomly sampled error term for between persons N{0, se}, (liter air/kg)

e_w = randomly sampled error term for within persons N{0, se}, (liter air/kg)

As indicated by Equation 5-2, the random error (ε) is allocated to two variance components used to estimate the between-person (inter-individual variability) residuals distribution (e_b) and within-person (intra-individual variability) residuals distribution (e_w). The regression parameters b_0, b_1, b_2, and b_3 are assumed constant over time for all simulated persons, e_b is sampled once per person by APEX, while whereas e_w varies from event to event. Point estimates of the regression coefficients and standard errors of the residuals distributions are given in Table 5-4. See Appendix 5A, Isaacs et al. (2008), and Chapter 7 of the APEX TSD (US

EPA, 2012b) for further discussion of this module. See also section 5.4.4 for a limited performance evaluation of this module in estimating ventilation rates.

Table 5-4 Ventilation equation coefficient estimates (b_i) and residuals distributions (e_i)

Age group	Ventilation Equation Coefficients[1]				Random Error[1]	
	b_0	b_1	b_2	b_3	e_b	e_w
<20	4.3675	1.0751	-0.2714	0.0479	0.0955	0.1117
20-<34	3.7603	1.2491	0.1416	0.0533	0.1217	0.1296
34-<61	3.2440	1.1464	0.1856	0.0380	0.1260	0.1152
61+	2.5826	1.0840	0.2766	-0.0208	0.1064	0.0676

[1] These are values of the coefficients and residuals distributions described by Equation (5-2) and described in Graham and McCurdy (2005).

Two key personal attributes determined for each simulated individual in this assessment are body mass (BM) and body surface area (BSA). Each simulated individual's body mass is randomly sampled from age- and sex-specific body mass distributions generated from National Health and Nutrition Examination Survey (NHANES) data for the years 1999-2004.[16] Details in their development and the parameter values are provided by Isaacs and Smith (2005). Then age- and sex-specific body surface area can be estimated for each simulated individual based on logarithmic relationships developed by Burmaster (1998) using body mass as an independent variable as follows:

$$BSA = e^{-2.2781}\ BM^{0.6821} \qquad (5\text{-}3)$$

5.2.7 Microenvironments Modeled

APEX is designed to estimate human exposure by using algorithms that attempt to capture the full range of O_3 concentrations expected within several microenvironments. Broadly aggregated, these can be either indoor, inside a motor vehicle, near road, or outdoor locations. The two methods available in APEX for calculating pollutant concentrations within microenvironments are a *mass balance model* and a *transfer factor* approach. Table 5-5 lists the 28 microenvironments selected for this analysis and the exposure calculation method used for each.

The importance of modeling indoor microenvironments (e.g., homes, offices, schools) is underscored by research indicating that personal exposure measurements of O_3 may not be well-

[16] Demographic (Demo) and Body Measurement (BMX) datasets for each of the NHANES studies were obtained from http://www.cdc.gov/nchs/nhanes/nhanes_questionnaires.htm.

correlated with ambient measurements and indoor concentrations are usually much lower than ambient concentrations (U.S. EPA, 2013, Section 4.3.3). We used mass balance modeling to estimate O_3 concentrations in all indoor microenvironments, considering probabilistic distributions of temperature-dependent (where data were available) building air exchange and chemical decay rates. Parameter settings for each of these variables are provided in Appendix 5B, while additional discussion regarding updates made to air exchange rates using more recent study data is given in Appendix 5E.

The remaining microenvironments were modeled using a transfer factors approach. Outdoor microenvironmental concentrations were assumed equivalent to ambient concentrations, near-road concentrations were adjusted considering whether or not O_3 concentrations were reduced by atmospheric reactions (e.g., scavenging by NO_X) or other processes, and vehicular microenvironments considered both the outdoor concentration attenuation and infiltration/removal in the concentration estimation. Specific parameter settings for each of these variables are provided in Appendix 5B.

Table 5-5 Microenvironments Modeled, Calculation Method Used, and Variables Included

Microenvironment	Calculation Method	Variables
Indoor: *Residence, Community Center or Auditorium, Restaurant, Hotel/Motel, Office building/Bank/Post Office, Bar/Night Club/Café, School, Shopping Mall/Non-Grocery Store, Grocery Store/Convenience Store, Metro-Subway-Train Station, Hospital/Medical/Care Facility, Industrial Factory/Warehouse, Other Indoor*	Mass balance	Building air exchange & chemical decay rates
Outdoor: *Residential, Park/Golf Course, Restaurant/Café, School Grounds, Boat, Other Outdoor Non-Residential*	Factors	None
Near-road: *Metro-Subway-Train Stop, Within 10 Yards of street, Parking Garage (covered or below ground), Parking lot (open)/Street parking, Service Station*	Factors	Proximity factors
Vehicle: *Cars/Light Duty Trucks, Heavy Duty Trucks, Bus, Train/Subway*	Factors	Proximity & penetration factors

5.2.8 Model Output

APEX estimates the complete time series of exposure concentrations for every simulated individual and can summarize data using standardized time metrics (e.g., hourly or daily average, daily maximum 8-hr average) or can output the minute-by-minute exposure concentrations (as is needed for the risk estimation in Chapter 6). As an indicator of exposure to O_3 air pollution, we selected the daily maximum 8-hr average O_3 exposure[17] for every simulated individual and

[17] It is important to stress here that only the maximum 8-hr exposure concentration is retained for each day simulated, per person. While every day could contain twenty-four unique 8-hr averages and that it is entirely

stratified these exposures by exertion level at the time of exposure. This indicator was selected based on controlled human exposure studies where reported adverse health responses were associated with exposure to O_3 and while the study subject was exercising.[18] Factors important in calculating this indicator includes the magnitude, duration, frequency of exposures, and the breathing rate of individuals at the time of exposure. As a reminder, the calculated daily maximum 8-hr average exposure concentrations are distinct from that of daily maximum 8-hr average ambient concentrations by accounting for simulated individual's time-location-activity patterns and O_3 concentration decay/variation occurring within the occupied microenvironments.

Benchmark levels used in this assessment include 8-hr average O_3 exposure concentrations of 60, 70 and 80 ppb; the same benchmark levels used for the 2007 O_3 exposure assessment (U.S. EPA, 2007b). Estimating exposures to ambient O_3 concentrations at and above these benchmark levels is intended to provide perspective on the public health impacts of O_3-related health effects observed in human clinical and toxicological studies, but that cannot currently be evaluated in quantitative risk assessments (e.g., lung inflammation, increased airway responsiveness, and decreased resistance to infection). The 80 ppb-8hr benchmark level represents an exposure level where there is substantial clinical evidence demonstrating a range of O_3-related effects including lung inflammation and airway responsiveness in healthy individuals. The 70 ppb-8hr benchmark level reflects evidence that asthmatics have larger and more serious effects than healthy people as well as a substantial epidemiological evidence of adverse effects associated with O_3 levels that extend below 80 ppb-8hr. The 60 ppb-8hr benchmark level represents the lowest exposure level at which O_3-related effects have been observed in clinical studies of healthy individuals. See ISA section 6.2.1 for further discussions regarding the body of evidence supporting the selection of these benchmark levels.

The level of exertion of individuals engaged in particular activities is approximated by an equivalent ventilation rate (EVR), that is, ventilation normalized by body surface area (BSA, in m^2) and is calculated as VE/BSA, where VE is the ventilation rate in liters/minute. For identifying moderate or greater exertion occurring during any 8-hr average exposure period in this assessment, we used the lower bound EVR value of 13 (liters/min-m^2) based on a range of EVRs used by Whitfield et al. (1996) to categorize persons engaged in moderate exertion activities for an 8-hr period. Whitfield et al. (1996) developed this range from EVR data reported in a 6.6-hr controlled human exposure study conducted by McDonnell et al. (1991).

---

possible multiple benchmark exceedances could occur for an individual on certain high O_3 concentration days, staff judge this is not a practical output for the purposes of this assessment.

[18] It is worth noting that the adverse health responses in the human clinical studies are generally based on 6.6 hour exposure to O_3. Therefore, it is possible that the number of benchmark exceedances is underestimated because of the lesser likelihood of an 8-hr exposure above the same threshold due to the longer averaging time.

APEX then calculates two general types of exposure estimates for the population of interest: the estimated number of people exposed to a specified O_3 concentration level and, the number of days per O_3 season that they are so exposed; the latter metric is expressed in terms of *person-days*. The former highlights the number of individuals exposed *one or more* times per O_3 season at or above a selected benchmark level. The person-days measure estimates the number of times per season the simulated individuals are exposed at or above a selected benchmark level and summed across individuals comprising the population. We note that a person-days metric conflates people and occurrences: one occurrence for each of 10 people would be counted the same as 10 occurrences for one person (i.e., 10 person-days at or above benchmark level). In this assessment we are more interested in reporting *multiday* exposures rather than total person-days, that is, the number of times an individual experiences multiple exposures at or above a benchmark level during an O_3 season. Given the complexities of the exposure modeling, the four study groups considered, the 15 study areas, the 5 years of ambient air quality, the multiple air quality scenarios simulated, and ultimately the output data generated, including both single and multiday exposures for simulated individuals, the consolidation of the results and the related graphic depictions used in this assessment requires additional discussion.

To begin, a simple example of summary results is the estimated percent of asthmatic school-age children experiencing exposures at or above a single 8-hr benchmark level when considering base air quality stratified by year (e.g., Figure 5-2, left panel). This presentation largely depicts the variability in O_3 exposure across the 15 study areas within years, along with an illustration of broad year-to-year temporal variability. A general finding regarding temporal variability extracted from this graph would be that fewer asthmatic school-age children exceed daily maximum 8-hr average exposures of 60 ppb considering 2009 base air quality when compared with other simulation years. An observation regarding the spatial variability could include the range of exposures within years (i.e., the study area variability) spans between 15 to 35 percentage points, dependent on the particular simulation year (Figure 5-2, left panel). One could also stratify the same exposure results by study area (e.g., Figure 5-2, right panel), thus depicting variability in estimated exposures across years within each study area, along with having broad study area comparisons. A general finding regarding temporal variability in this type of presentation would be that the range of exposures within study areas spans about 20 percentage points, though some study areas have a generally small range (<5 percentage points) for most simulated years. An observation regarding spatial variability could be that Chicago largely has the fewest asthmatic school-age children at or above benchmark levels, having a mean about 15%, while Los Angeles consistently has the most asthmatic school-age children, having a mean about 35%, at or above benchmark levels while at moderate or greater exertion.

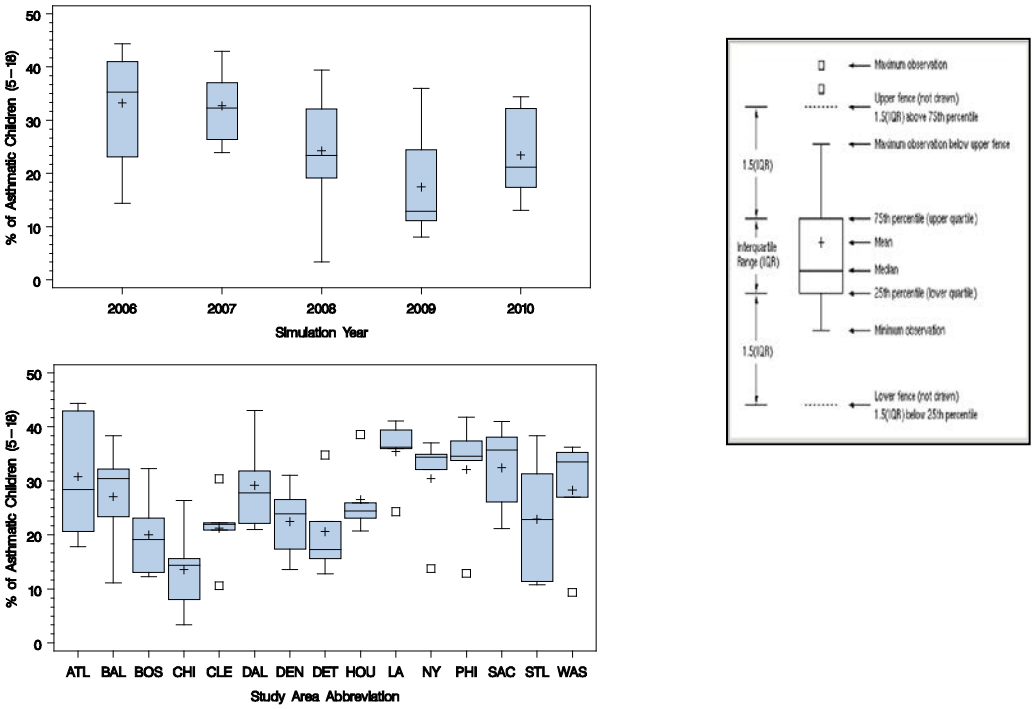


Figure 5-2 Percent of asthmatic school-age children in all study areas with at least one O_3 exposure at or above 60 ppb-8hr while at moderate or greater exertion using base air quality (2006-2010), stratified by year (top left panel) or by study area (bottom left panel).

While these boxplots are an efficient tool that summarize potentially complex data sets by illustrating important statistical aspects of data analysis results (e.g., means, ranges, occasional upper percentile data values), at times important features of the data may be masked (e.g., trends or patterns within consolidated variables) and the presentation of other aspects of the exposure results would require the generation of additional graphs (e.g., results for additional benchmark levels). A tabular format could be one way to present all possible data, though given the number of APEX simulations performed (i.e., > 1,000) and aforementioned dimensions of the assessment, linking the trends and patterns across all study areas, years, and benchmark levels from the numerous output tables would be visually challenging.

This discussion regarding properly representing temporal and spatial variability in the exposure results can be further extended to include the added dimension of the five air quality scenarios (base, existing standard, and three alternative standard levels). Mindful of these complexities, we elected to use a multi-panel graphing approach to succinctly summarize the exposure output data, while also retaining as much information as possible in a single page format to allow for visual analysis of trends and patterns. As an example, Figure 5-3 (top panels)

illustrates boxplots for Atlanta similar to those presented above, though the exposure results are for the three exposure benchmark levels of interest, with each stratified by the particular adjusted air quality scenario. As expected with increasing stringency of the 8-hr standard level, fewer asthmatic school-age children are exposed at or above a given benchmark level. Also expected is the fewer percent of asthmatic school-age children exposed to higher benchmark levels when compared with lower benchmark levels. While these three graphs can provide a clear depiction of the exposure results for a single study area, the six years encompassing the two averaging periods 2006-2008 and 2008-2010 are combined in the graphic and difficulty would remain in simultaneously exhibiting all 15 areas.

To overcome these limitations, Figure 5-3 (lower panel) exhibits all of the dimensions of the exposure results mentioned above (i.e., year, benchmark level, study area) along with distinguishing between the two standard averaging periods for each the existing (75 ppb-8hr) and alternative standard levels (60, 65, 70 ppb-8hr). The nomenclature above each subgraph indicating the particular air quality scenario requires defining. For example, a panel heading of "75" contains the exposures estimated when air quality was adjusted to just meet the existing standard level of 75 ppb-8hr (4th highest daily maximum 8-hr average O_3 concentration averaged over a three year period) either using years 2006, 2007, and 2008 ambient air quality data or for a second averaging period that extended from 2008 through 2010 (with results for each given by two separate lines on the same plot). Exposure results are readily observed for any air quality scenario, year, or benchmark level of interest. For example, when considering the 75ppb standard 2006-2008 averaging time scenario, 20% of asthmatic school-age children in Atlanta experience at least one daily maximum 8-hr average exposure of 60 ppb occurs when considering year 2006 air quality, while only about 5% experience exposures at or above the same benchmark level considering 2008 air quality (though when considering the 2008-2010 averaging period, approximately 20% of asthmatic children are estimated experience at least one exposure at or 60 ppb-8hr). Fewer than 5% of asthmatic school-age children in Atlanta experience at least one benchmark exposure of 70 ppb-8hr considering any year and any air quality scenario, including just meeting the existing O_3 standard.

Because APEX simulates the complete time series of exposure for every simulated individual, also output is the number of times an individual experiences a benchmark exceedance over the duration of the simulation (i.e., the entire O_3 season simulated in each study area). These data can also be summarized in a similar multi-panel format, though differ slightly in composition from that of Figure 5-3. Instead of displaying the percent of persons with at least one exceedance of each of the three benchmarks, presented are the percent of persons with multiple exposures at or above a single benchmark within an O_3 season. For example, Figure 5-4 illustrates the percent of asthmatic school-age children in Atlanta having multiple days where

exposures (≥2, ≥4, and ≥6 per O_3 season) were at or above 60 ppb-8hr considering the 2006-2010 air quality adjusted to just meet the existing and alternative standards levels. When considering 2006 air quality adjusted to just meet the existing standard, approximately 10% of asthmatic school-age children experienced at least two days where their daily maximum 8-hr average exposure was at or above 60 ppb, though fewer than 5% experienced such exposures in 2009. When collectively considering all simulated air quality scenarios and years, fewer than 3% of asthmatic school-age children experienced at least four exposures at or above 60 ppb and virtually no asthmatic school-age children experienced six or more such exposures over the O_3 season.

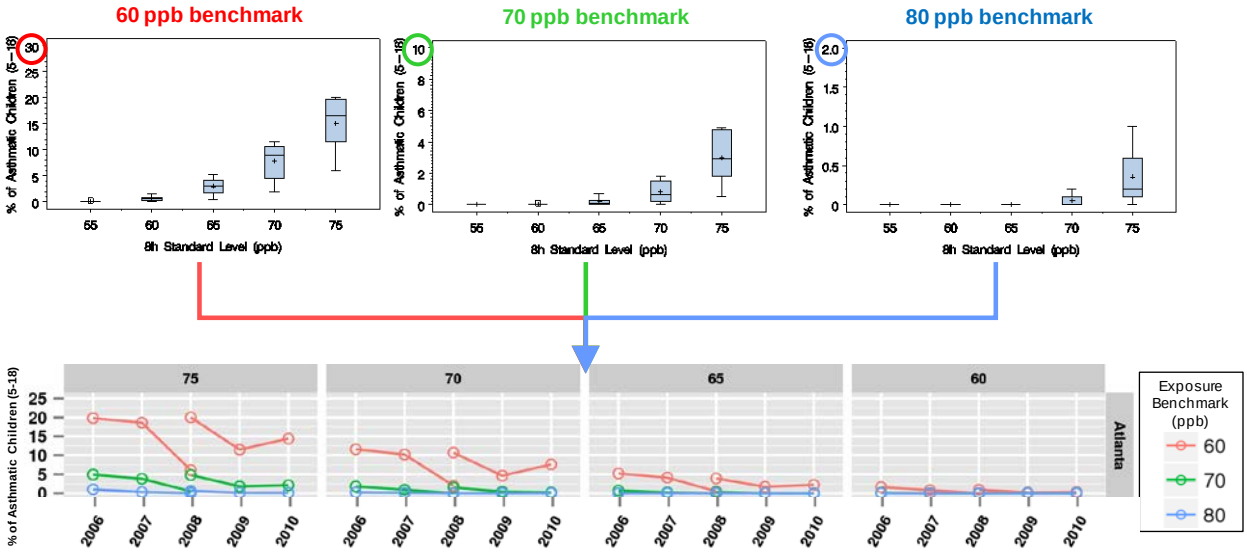


Figure 5-3 Percent of asthmatic school-age children in Atlanta with at least one O_3 exposure at or above 60 ppb-8hr (left top panel), 70 ppb-8hr (middle top panel), and 80 ppb-8hr (right top panel while at moderate or greater exertion, years 2006-2010 air quality adjusted to just meet the existing and alternative O_3 standard levels. The multi-panel display (bottom) illustrates the same exposure results expanded to reflect individual data points by year, standard averaging period, and benchmark level.

Also worth discussing is the appearance of a similar pattern between the benchmark level results (Figure 5-3) and the number of exceedances of a single benchmark (Figure 5-4). Because the ambient concentration is an important determinant in exposure concentrations, it is not surprising to see that the trend over years for persons having at least one exposure at or above a particular benchmark level (e.g., 60 ppb-8hr) is similar to those experiencing at least two exposures above 60 ppb-8hr (though a smaller percentage of persons). This is because years having the highest peak concentrations will yield the greatest percent of persons above benchmark levels, and when one year has a day with the highest concentration, it is likely that

year also has a second day with a similarly and relatively high concentration, and so on. Using the same logic, one might also conclude that there could be a pattern between the percent of persons experiencing a single exceedance of 70 ppb-8hr and multiple exceedances (e.g., four) of 60 ppb-8hr also driven by the overall ambient concentration distribution. However, given that very few persons experience these types of benchmark exceedances, determining the relationship between the two (if present) may not be of practical significance. For brevity, the complete multiday exposure results for all APEX simulations are presented in Appendix 5F, with results presented for one study group (e.g., all school-age children) in the main body of the REA.

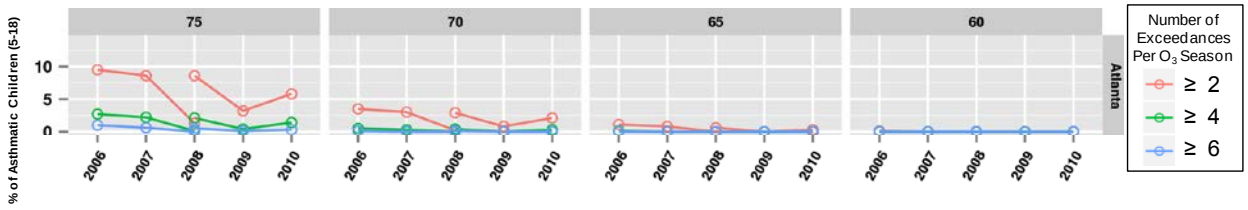


Figure 5-4 Percent of asthmatic school-age children in Atlanta with multiple O_3 exposures at or above 60 ppb-8hr while at moderate or greater exertion, years 2006-2010 air quality adjusted to just meet the existing and alternative O_3 standard levels.

5.3 EXPOSURE ASSESSMENT RESULTS

5.3.1 Overview

The results of the exposure analysis are presented as a series of figures focusing on the defined range of benchmark levels (i.e., persons experiencing daily maximum 8-hr average O_3 exposure concentrations at or above 60 ppb, 70 ppb, and 80 ppb), noted as being of particular health concern (Section 5.2.8). A range of concentrations in the air quality data over the five year period (2006-2010) were used in the exposure model, providing a range of estimated exposures output by APEX. The adjusted air quality was developed using two distinct 3-year period design values (2006-2008 and 2008-2010), as described in Chapter 4.[19] Exposures were estimated for four study groups of interest (i.e., all school-age children (5-18), asthmatic school-age children, asthmatic adults (19-95), and older adults (65-95)) in each of the 15 study areas.

In this exposure assessment, we are primarily interested in O_3 exposures associated with the ambient air quality adjusted to just meet the existing and potential alternative O_3 standards.

[19] Thus, the year 2008 will have two sets of estimated exposures, one from each of the two sets of design values. In Figure 5-2, the greater temporal variability observed for 2008 is driven in part by differences in some study areas resulting from the air quality adjustment period. Exposure results for both 2008 averaging periods are provided when presenting data by year. Where mean results are presented in subsequent results sections, the two values given for year 2008 were first averaged to give a single exposure value for 2008 before averaging across all years.

Thus, most of the exposure results presented and discussed are for where ambient air quality was adjusted to just meet these particular scenarios. While understanding exposures and health risks associated with historical and existing air quality is important, the primary goal of this and any REA is to evaluate to what extent the existing NAAQS, and its associated air quality, protect health and to what extent alternative NAAQS protect health. Exposure results associated with recent (base) air quality are briefly discussed here first, though largely reported in Appendix 5F.

5.3.2 Exposure Modeling Results for Base Air Quality

The exposure results for the base air quality are distinguished from the other air quality scenario results primarily due to the wide ranging variability in estimated exposures across the study areas and years. The variability in exposures are the result of the wide ranging variability in ambient concentration levels, with perhaps some years in some study areas exhibiting air quality at or near that just meeting the current 8-hr standard, while other study areas and years exhibiting air quality levels much higher than the existing 8-hr standard. These exposures are informative in describing the existing or recent health risks associated with a unique air quality scenario, but because they variably diverge from a set concentration level of interest (such as the existing 8-hr standard), they are of limited relevance in evaluating the adequacy of either the existing NAAQS as well as potential alternative air quality standards. That said, detailed tabular and graphic presentations of exposure results associated with the base air quality (years 2006-2010) are provided in Appendix 5F, with only key findings summarized in the following discussion.

Consistent with the previously discussed observations regarding year-to-year variability in ambient concentrations (Chapter 4), most study areas have the greatest percent of all school-age children experiencing concentrations at or above the three benchmark levels during 2006 or 2007 along with having the lowest percent of all school-age children exposed during 2009. In general, between 20-40% of all school-age children experience at least one O_3 exposure at or above 60 ppb-8hr, 10-20% experience at least one O_3 exposure at or above 70 ppb-8hr, and 0-10% experience at least one O_3 exposure at or above 80 ppb-8hr, all while at moderate or greater exertion (i.e., an 8-hr EVR $\geq$ 13 L/min-m^2) and considering the base air quality (2006-2010). Year-to-year variability observed for asthmatic school-age children and the percent of asthmatic school-age children were similar to exposure results for all school-age children, largely a function of having both simulated study groups using an identical time-location-activity diary pool to construct each simulated individual's time series of activities performed and locations visited.

The overall year-to-year pattern of exposure for asthmatic adults is similar to that observed for all school-age children, though the percent of the asthmatic adult study group

exposed is lower by a factor of about three or more. Having a lower percent of asthmatic adults exposed is expected given that outdoor time expenditure is an important determinant of O_3 exposure (section 5.4.2) and that adults spend less time outdoors than children (section 5.4.1), as well as adults having a lower outdoor participation rate. The percent of all older adults experiencing exposures at or above the selected benchmark levels is lower by a fewer percentage points when compared with the results for asthmatic adults. Again, older adults, on average, would tend to spend less time outdoors and do so with less frequency when compared with both adults and children (section 5.4.1), in addition to fewer older adults performing activities at moderate or greater exertion for extended periods of time, thus leading to fewer persons exposed to O_3 concentrations of concern.

The year-to-year patterns of the single and multiple exposure occurrences considering base air quality (2006-2010) were similar among the four exposure study groups, therefore only results for all school-age children will be summarized here. Depending on the year and study area, about 10-25% of all school-age children could experience at least two exposures above the 60 ppb-8hr benchmark during the O_3 season, while about 5-10% school-age children could experience at least four. Most study areas and years are estimated to have fewer than 5% of all school-age children experience six or more exposures above 60 ppb-8hr considering the base air quality. When considering the multi-day exposures for all school-age children at or above the 70 ppb-8hr benchmark, about 2-10% of all school-age children could experience at least two exposures during the O_3 season, while four or more exposures were generally limited to fewer than 4% of all school-age children. Almost half of the study area-year combinations had no school-age children experiencing two or more exposures at or above the 80 ppb-8hr benchmark, with the other half estimated to have about 1% of all school-age children experiencing two or more exposures at or above the 80 ppb-8hr benchmark. School-age children having four or more 80 ppb-8hr benchmark exceedances were limited to only a few study area years and, where a non-zero value was estimated, were limited to $\leq 0.5\%$ of the study group.

5.3.3 Exposure Modeling Results for Simulations of Just Meeting Existing and Alternative O_3 Standards

In this section, we present the exposures estimated when considering the air quality adjusted to just meeting the existing O_3 NAAQS standard, as well as when considering potential alternative standard levels (55, 60, 65, 70 ppb 8-hr) of the existing standard. Comprehensive multi-panel displays of exposure results are presented for each of the study groups of interest, i.e., all school-age children (5-18), asthmatic school-age children, asthmatic adults (19-95), and all older adults (ages 65-95; Figure 5-5 to Figure 5-8, respectively). Included in each display are the three benchmark levels (60, 70, and 80 ppb-8hr), the five years of air quality (2006-2010), for the 15 study areas. A single multi-panel display is used to present the results for each of the four

study groups, beginning with the estimated percent of persons exposed at least one time at or above the selected benchmark levels. Modeled exposures in the 15 study areas and considering each benchmark level are presented on the same scale to allow for direct comparisons across the multi-panel display. The most notable patterns in the exposure results are described here using one study group (i.e., all school-age children), as there is a general consistency in the year-to-year variability within each study area across all four study groups. Any deviation from the observed pattern will be discussed for the subsequent study group.

We note that after adjusting to just meet a potential 8-hr ambient standard level of 55 ppb, there were nearly no persons exposed at or above any of the selected benchmark levels, thus these data, while modeled, are not presented in detail here. In addition, in one study area (Chicago), O_3 ambient monitor design values were below that of the existing standard during the 2008-2010, therefore APEX simulations could not be performed for meeting the existing standard for that 3-year period. And finally, we were not able to simulate just meeting a standard level of 60 ppb-8hr or below in the New York study area (see Chapter 4 for details), thus APEX simulations for these air quality scenarios could not be performed in New York.

Figure 5-5 illustrates the exposures estimated for all school-age children in each study area with general observations as follows. After adjusting air quality to just meet the existing and alternative standards, there are virtually no school-age children exposed at or above 80 ppb-8hr, with very few school-age children exposed at or above the 70 ppb-8hr benchmark. For example, out of 87 possible study area and year combinations considering air quality adjusted to just meet the existing standard (the least stringent standard level considered here), only 29 resulted in > 0.1% estimated percent of all school-age children exposed at least once at or above the 80 ppb-8hr benchmark with the maximum percent of all school-age children exposed estimated for St. Louis (1.1%). Ninety-four percent of study area and year combinations had fewer than 5% of all school-age children experiencing at least one daily maximum 8-hr average exposure $\geq$ 70 ppb considering ambient air quality adjusted to just meeting the existing standard, again with a maximum of 8.1% occurring in St. Louis. When considering air quality adjusted to just meet an 8-hr ambient standard level of 70 ppb, $\leq$ 0.2% of all school-age children experience at least one 80 ppb-8hr exposure benchmark exceedance for all study area and year combinations, while for 76 or 90 study area and year combinations, $\leq$ 1% of all school-age children experience a 70 ppb-8-hr exposure benchmark exceedance. This pattern of having very few school-age children experiencing exposures at or above 70 and 80 ppb-8hr is as expected given the nature of the air quality adjustment procedure that limits 8-hr ambient concentrations at or above the selected potential alternative standard level.

In contrast, approximately 10-20% percent of all school-age children are estimated to be exposed to at least one 60 ppb-8hr concentration when considering air quality just meeting the

existing standard (Figure 5-5). And similar to that mentioned above regarding exposures associated with the base air quality, a general year-to-year exposure pattern emerges with respect to study area and year. For the Northeastern (Boston, New York), Mid-Atlantic (Philadelphia, Washington DC, Cleveland) and Mid-Western (Chicago, Detroit, and St. Louis) study areas, the maximum percent of all school-age children exposed generally occurs during year 2007. For the Southern (Atlanta, Dallas, Houston) and Western (Denver, Los Angeles, Sacramento) study areas, the maximum exposure occurs during year 2006. Deviations from this temporal exposure pattern appear mostly as a result of the standard averaging period, with the 2008-2010 period producing equal or greater maximum exposures during either 2008, 2010, or both years and most prevalent in the Northeastern and Mid-Atlantic study areas (Baltimore, Boston, New York, Philadelphia, Washington DC; note also a trend in Atlanta, Denver, St. Louis).

These 60 ppb-8hr exposure patterns remain consistent when considering air quality adjusted to just meet a 70 ppb-8hr ambient standard, though the percent of all school-age children exposed is less than that observed when considering the air quality adjusted to just meet existing standard. Further, 75 of 90 study area and year combinations are estimated to have ≤ 10% of all school-age children experience a 60 ppb-8hr or greater exposure, though between 10-20% of all school-age children were estimated to be exposed for a few study area and year combinations (e.g., Atlanta-2006, Chicago-2007 and -2010, and Houston-2009). When considering air quality adjusted to just meet a 65 ppb standard level, the percent of all school-age children experiencing an exposure at or above 60 ppb-8hr diminishes to 5% or less for most study areas and years (i.e., 81 of 90 study area year combinations).

All of what has been described regarding the estimated exposures to school-age children (i.e., the year-to-year and benchmark level patterns, and the percent of the study group exposed) also applies to the exposures estimated for asthmatic school-age children (Figure 5-6). Different however would be the relative number of asthmatic school-age children exposed in each study area if compared with all school-age children, as the asthma prevalence rates vary by study area (Table 5-2), though on average are about 10% of the population of children.

The percent of asthmatic adults (Figure 5-7) experiencing daily maximum 8-hr average exposures above the selected benchmark levels is sharply lower than that estimated for all school-age children. For example, only three of a possible 84 study area and year combinations (Chicago-2007, Houston-2009, and St. Louis-2007) were estimated have > 0.1% of asthmatic adults experience a daily maximum 8-hr average exposure ≥ 80 ppb, and only six of a possible 84 study area and year combinations were estimated have >1% of asthmatic adults experience an daily maximum 8-hr average exposure ≥ 70 ppb, all occurring when considering air quality just meeting the existing standard. No study area or year combination has more than 10% of asthmatic adults estimated to experience an exposure at or above 60 ppb-8hr when considering

air quality just meeting the existing standard, with 67 of 84 study area and year combinations estimated to have 5% or less asthmatic adults experiencing such exposures.

When considering air quality adjusted to just meeting a standard level of 70 ppb-8hr, no asthmatic adults experience an exposure at or above 80 ppb-8hr and ≤ 0.6% experience a daily maximum 8-hr average exposure ≥ 70 ppb for any study area or year combination. Less than 5% of asthmatic adults could experience an exposure at or above 60 ppb-8hr when considering air quality adjusted to just meet a standard level of 70 ppb-8hr for 88 or 90 possible study area year combinations, with the maximum percent of adult asthmatics exposed outside this range occurring in Denver (6.8%-2008) and St. Louis (5.5%-2007).

Older adults are estimated to have the fewest exposures above the two highest benchmark levels when considering the adjusted air quality. For example, only two of a possible 84 study area and year combinations (St. Louis-2007 and Washington DC-2008) were estimated have > 0.1% of asthmatic adults experience a daily maximum 8-hr average exposure ≥ 80 ppb, and only six of a possible 84 study area and year combinations were estimated have > 1% of asthmatic adults experience a daily maximum 8-hr average exposure ≥ 70 ppb, all occurring when considering air quality just meeting the existing standard (Figure 5-8). Also, exceeding the 60 ppb-8hr exposure benchmark appears to be limited to fewer than 5% of all older adults when considering air quality adjusted to just meet the existing standard and a standard level of 70 ppb-8hr, and occurs in < 2% of all older adults when considering a standard level of 65 ppb-8hr.

An example of multi-day exposure results associated with adjusted air quality is provided in Figure 5-9. The percent of all school-age children estimated to experience multi-day exposures above benchmark levels during each study area's O_3 season is largely limited to two air quality scenarios: the existing standard and air quality adjusted to just meeting a standard level of 70 ppb-8hr. This is because of the small percent of school-age children experiencing even a single exposure above the lowest benchmark level when considering standard levels at or below 65 ppb-8hr. In addition, when experiencing multiple exposures, most school-age children appear to have at most two days above benchmark levels per O_3 season, even when considering the lowest benchmark level of 60 ppb-8hr. For example, 81 of 87 possible study area and year combinations have < 10% of all school-age children experiencing two or more exposures ≥ 60 ppb-8hr when considering an ambient standard level of 75 ppb-8hr, while 83 of 90 possible study area and year combinations have < 5% of all school-age children experiencing two or more exposures ≥ 60 ppb-8hr when considering an ambient standard level of 70 ppb-8hr. With increasing stringency in the standard level to 65 ppb-8hr, 81 of 90 possible study area and year combinations have < 1% of all school-age children experiencing two or more exposures ≥ 60 ppb-8hr.

Multi-day exposure to the higher exposure benchmarks (either the 70 or 80 ppb-8hr) is a rare occurrence, even when considering the air quality adjusted to the existing O_3 standard. For

example, there were no school-age children experiencing two or more exposures above 80 ppb-8hr in all but one study area year combination and, and when considering that one study year having a non-zero value (St. Louis-2007), the estimated percent of all school-age children at or above the exposure benchmark was only 0.1%. Further, 83 of 87 possible study area and year combinations have < 1% of all school-age children experiencing two or more exposures ≥ 70 ppb-8hr, also when considering an ambient standard level of 75 ppb-8hr.

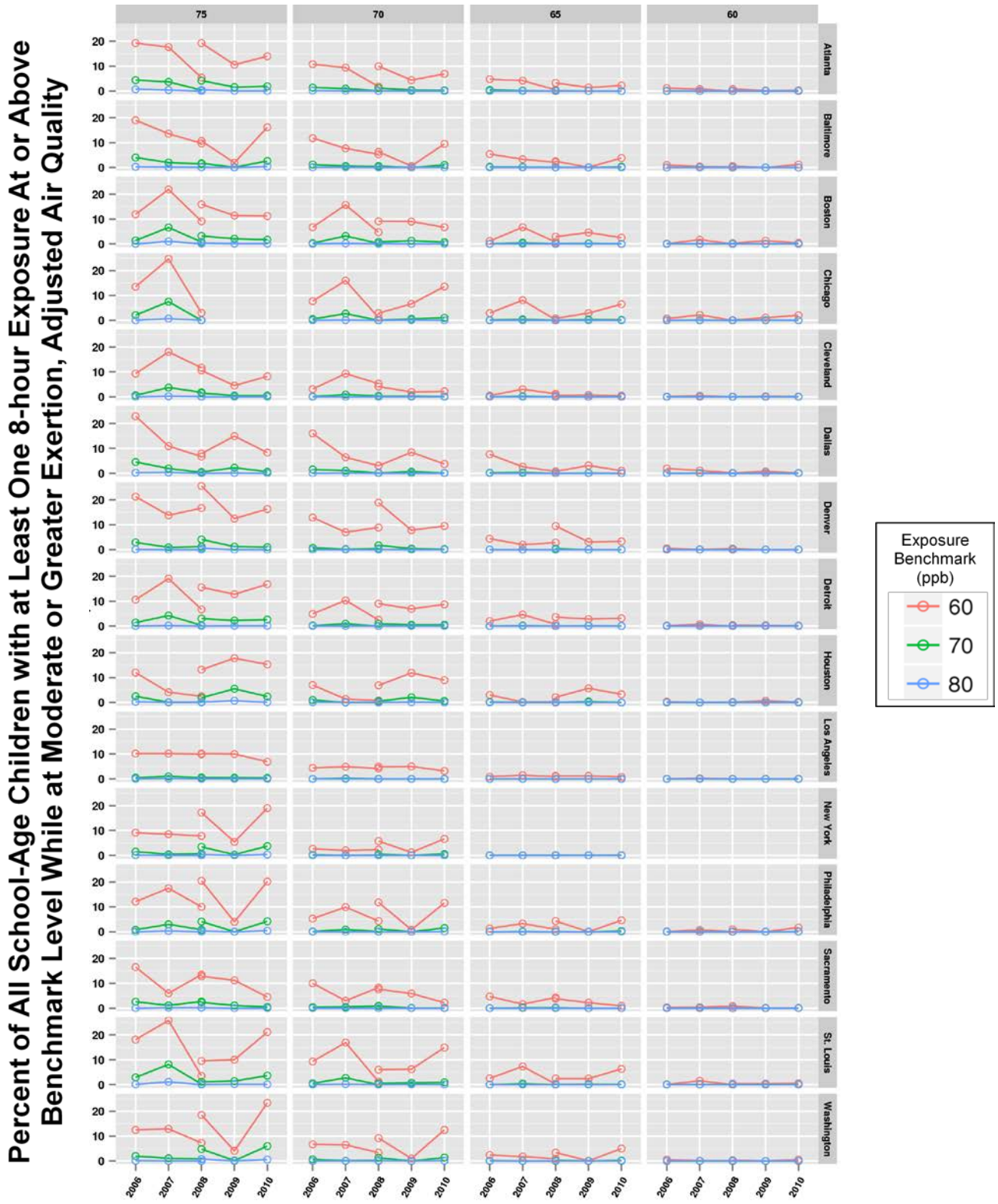


Figure 5-5 Percent of all school-age children with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards.

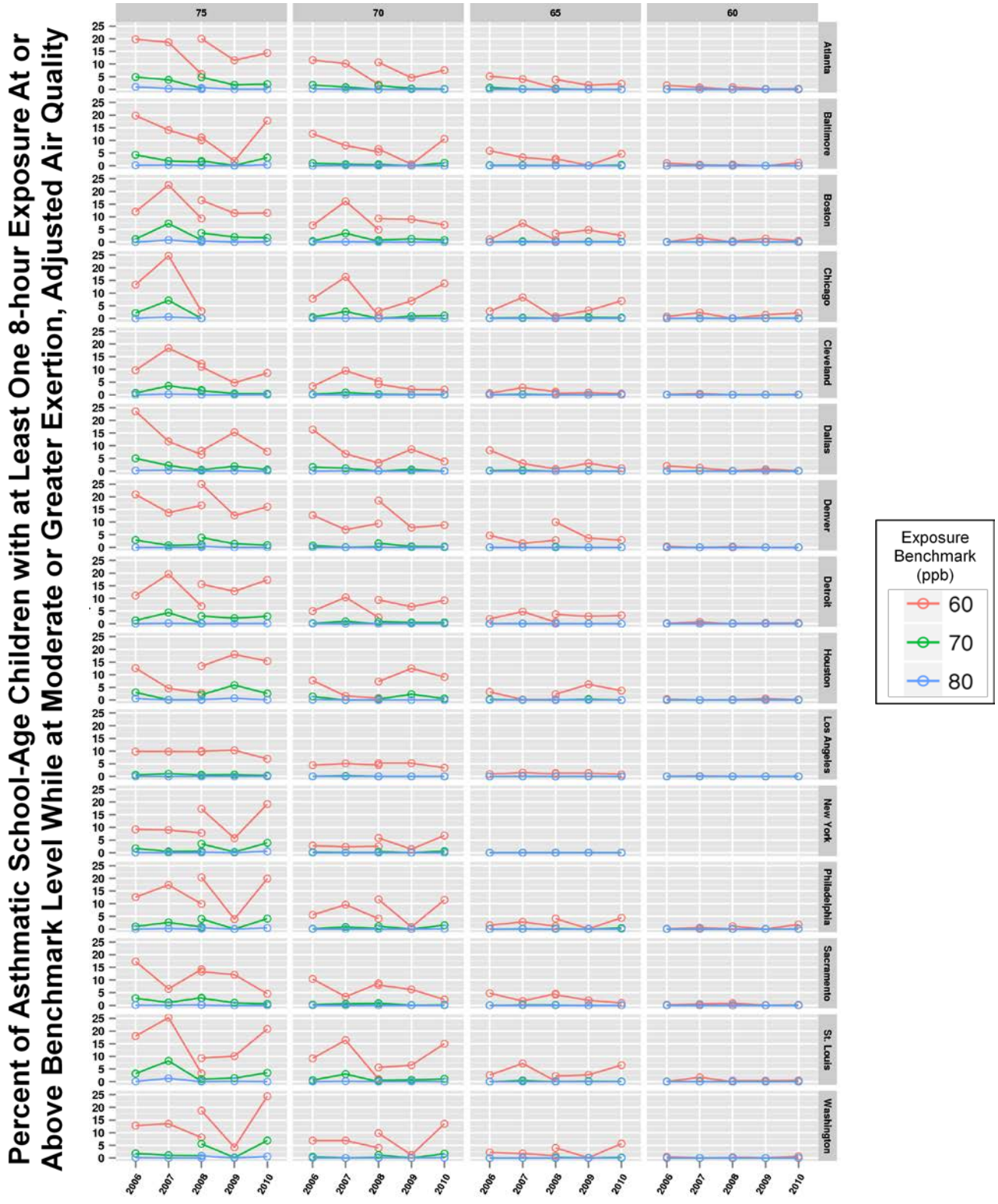


Figure 5-6 Percent of asthmatic school-age children with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards.

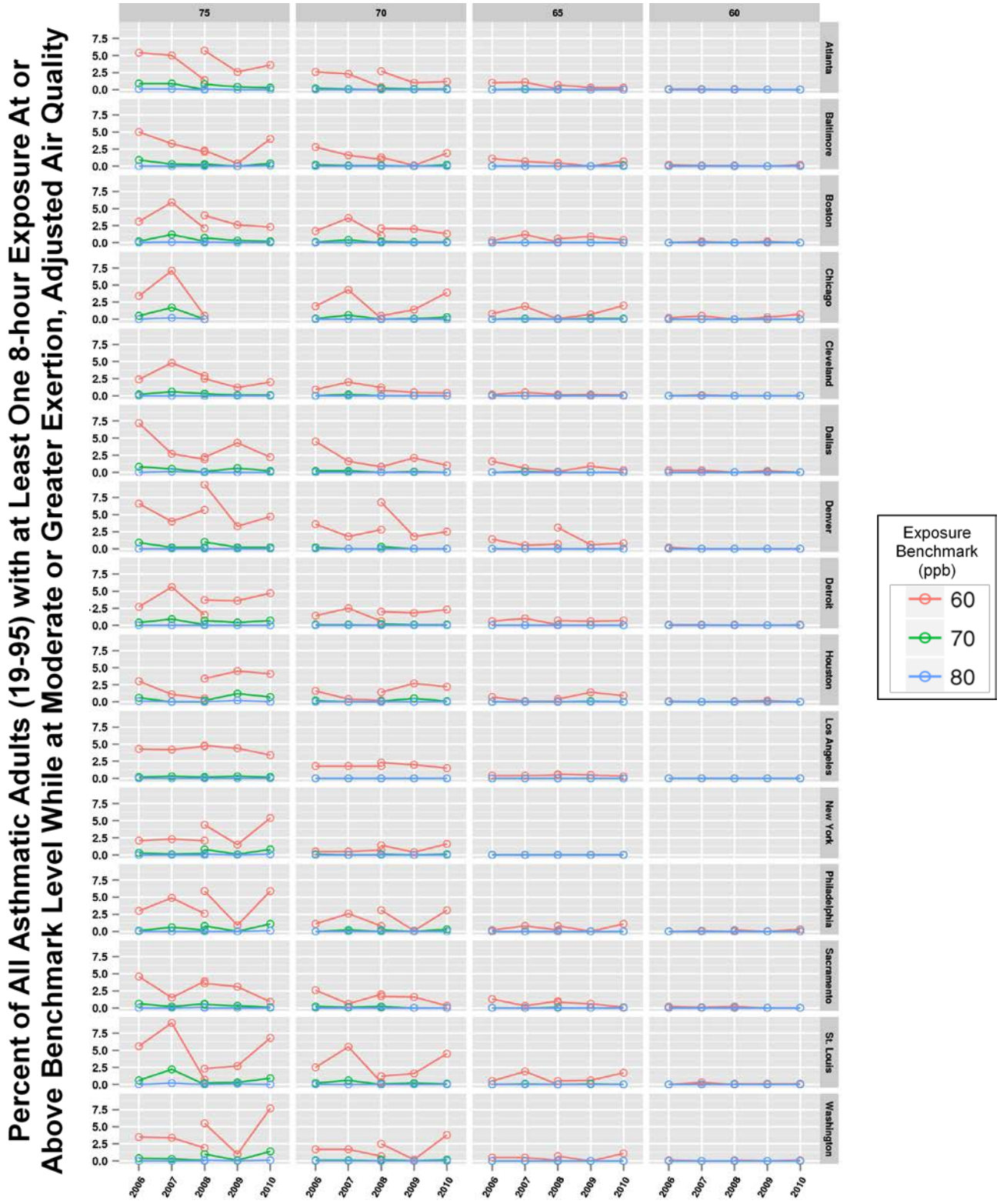


Figure 5-7 Percent of all asthmatic adults with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards.

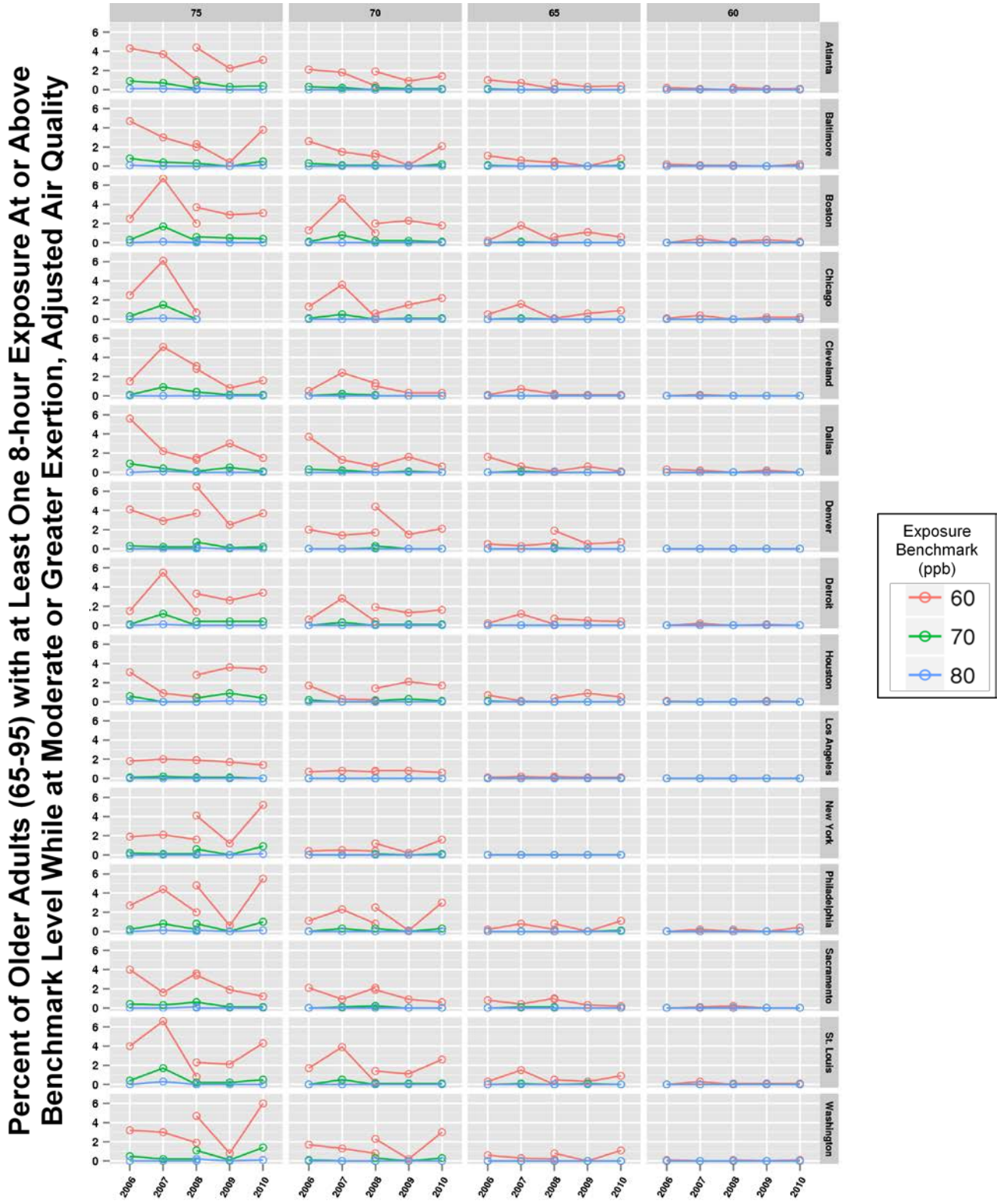


Figure 5-8 Percent of all older adults with at least one daily maximum 8-hr average O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards.

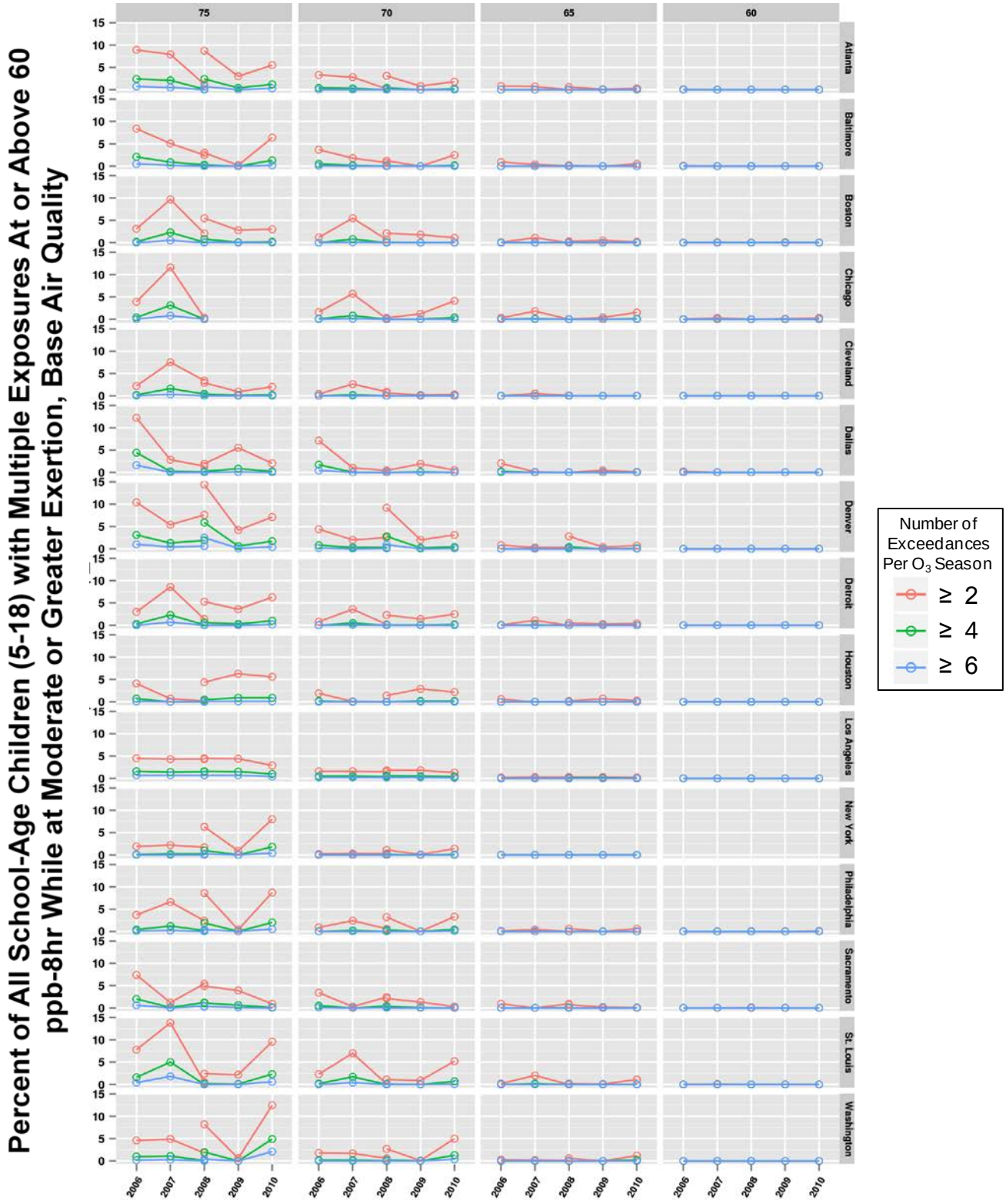


Figure 5-9 Percent of all school-age children with multiple daily maximum 8-hr average O_3 exposures at or above 60 ppb while at moderate or greater exertion, years 2006-2010, air quality adjusted to just meet the existing and potential alternative standards.

5.4 TARGETED EVALUATION OF EXPOSURE MODEL INPUT AND OUTPUT DATA

This section summarizes the results of several targeted evaluations intended to provide additional insights to APEX input data or approaches used to estimate exposures (CHAD data attributes and activity pattern evaluations, comparison of CHAD outdoor time data with ATUS, comparisons of asthmatic outdoor time expenditure and exertion levels to that of non-asthmatics), exposure results for additional exposure populations of interest (outdoor workers, school-age children during summers, impact of averting), and model performance evaluations (personal exposure measurements and independent ventilation rate estimates compared with APEX estimates). Detailed analysis results are provided in Appendix 5G.

5.4.1 ANALYSIS OF TIME-LOCATON-ACTIVITY DATA

While CHAD is the most comprehensive and relevant source of time-location-activity data available for use in our exposure modeling, there are a few limitations to the survey data contained therein, many of which are founded in the individual studies from which activity patterns were derived (Graham and McCurdy, 2004). CHAD is a collection of related survey data, though individual study attributes can range widely (e.g., survey participant ages, region or city of residence, time-of-year data collected). We note that many of the assumptions about use of these activity patterns in exposure modeling are strengthened by the manner in which they are used by APEX. This is done by focusing on selecting the most important individual attributes that contribute to variability in human behavior (e.g., age, sex, day-of-week, ambient temperature) and linking these attributes of simulated individuals to the population demographics of each census tract (see section 5.2.5) and the study area temperatures (section 5.2.4). Further, one key lifestyle attribute is also accounted for in generating longitudinal diary profiles by simulating both the intra- and interpersonal variability in time spent outdoors (section 5.2.5; Glen et al., 2008).

A few questions may arise as to the representativeness of the CHAD diaries to the simulated population. For example, the year of a particular survey study may differ from our simulated exposure population by as much as 30 years (i.e., some activity pattern data were generated in the 1980s). In addition, there are other personal attributes (e.g., ethnicity, income level, lifestyle factors[20]), health conditions (e.g., asthma, cardiovascular disease), and situational factors (e.g., availability of parks and recreation areas) that are not used in creating the simulated persons that could be influential in estimating exposures. Considering this, a number of

[20] Examples of such factors for adults could include married/unmarried, having infants or young children/no children. Lifestyle factors for children could include whether the child is active/non-active or whether or not there is time spent outdoors.

evaluations were performed to answer questions regarding important personal attributes used in generating simulated individuals and the general representativeness of the CHAD time-location-activity data. First though, we summarize the newly acquired activity pattern data now included in CHAD compared with data available and used in the 1st draft O_3 REA.

5.4.1.1 *General Evaluation of CHAD Study Data: Historical and Recently Acquired Data*

The number of diary days having complete information and used by APEX in the 2nd draft O_3 REA is 41,474 (Table 5-3). This is an increase of about 8,700 diaries currently used by APEX compared with what was used by APEX in the 1st Draft O_3 REA. Further, there have been eight new study data sets incorporated into CHAD and used in our current exposure assessment since the previous O_3 NAAQS review conducted in 2007, most of which were from recently conducted activity pattern studies (see Appendix 5B, Section 5B-4 for more information regarding these studies). The diary data included from these new studies have more than doubled the total activity pattern data used for 2007 O_3 exposure modeling and has increased the number of children's diaries by about a factor of five. Currently, the majority of diaries (54%) from CHAD are taken from surveys conducted in the past decade, while the pre-1990s diaries represent less than 15% of the total diaries available by APEX.

5.4.1.2 *Exposure-Relevant Personal Attributes Included in CHAD and APEX Simulated Individuals*

The survey participants whose diary data are within CHAD were asked a number of questions regarding their personal attributes. The number and type of attributes present for diaries in CHAD is driven largely by the original intent of the individual study. In our exposure assessment, we have strict requirements to simulate individuals using several personal attributes, namely age, sex, temperature (as a surrogate for seasonal variation in activity patterns), and day-of-week. These attributes are considered as important drivers influencing daily activity patterns (Graham and McCurdy, 2004) and when diaries do not have these particular attributes for a particular day, the diary day will not be used by APEX. We compared the representation of these and other attributes in the current CHAD used by APEX with that in the 1st draft O_3 REA and found strong similarities in the attribute distributions between both databases, suggesting little change in the overall composition of the database regarding these influential attributes.

While there may be other personal or situational attributes that affect daily time expenditure (e.g., socioeconomic status, occupation of an employed person), these attributes are typically not included in our assessment to generate simulated individuals simply because the response to the attribute is missing for most of the study participants/CHAD diary days. For example, income level is missing for about two-thirds of the CHAD diaries because either the original study did not have an income/occupation related survey question or perhaps the participant refused to answer the question if it were posed. If one were to select this personal

attribute in developing a simulated individual's activity pattern (among using any other attribute having missing responses), the pool of diaries available to simulate individuals may be extremely limited, likely leading to repetition of diaries used for individuals or groups of similar individuals and artificially reducing both intra- and inter-personal variability in time expenditure, or perhaps resulting in model simulation failure altogether. This is why personal attributes are carefully selected and prioritized according to both their prevalence in CHAD and whether the attribute has a known significant influence on activity patterns.

5.4.1.3 *Evaluation of Afternoon Time Spent Outdoors for CHAD and Survey Participants*

There have been questions raised regarding the representativeness of the diaries from studies conducted in the 1980s and whether there are any recognizable patterns in time expenditure in the CHAD diaries across the time period when data were collected. Because time spent outdoors is a significant factor influencing daily maximum 8-hr average O_3 exposures, we evaluated the current collection of CHAD diaries used by APEX for two metrics and considering two dimensions: outdoor participation rate (i.e., the percent of people who spent some time outdoors during their survey day) and the mean time spent outdoors for where the persons spent at least one minute outdoors or at least 2 hours outdoors. Because time spent outdoors is an important determinant for highly exposed individuals, we summarize the results here for the diaries having at least 2 hours of outdoor time here, while all other results are provided in Appendix 5G. CHAD diaries were stratified by five age groups (4-18, 19-34, 35-50, 51-64, 65+) and three decades (1980s, 1990s, and 2000s) using the year the particular activity pattern study was conducted. We note that CHAD is composed of primarily cross-sectional data (single diary days per person), thus the trend evaluated over the three decades is changes (if any) in participation rate and the time spent outdoors by the composite study population, not within individuals.

Regardless of decade and duration of time spent outdoors, children tended to have the highest outdoor participation rate when compared with the other age groups, while the oldest adults (aged 65 or greater) tend to have the lowest participation rate. The CHAD diaries from the 1980's studies for children ages 4-18 have the highest outdoor participation rate (50%) compared to other decades (35-40%) and all other age groups and decade of collection. When considering the pool of diaries available for this age group, these 1980's studies contribute to approximately 19% of diaries having two or more hours of time spent outdoors during the afternoon. This translates to a small effect on the overall outdoor participation rate for diary pools that would include these earlier studies (39% participation rate) compared to the participation rate excluding these studies (36% participation rate). In general, these outdoor participation rates are similar to the finding reported recently by Marino et al. (2012) of 37.5%, though estimated for pre-school age children. Thus, when considering participation in outdoor activities and the

representativeness of the CHAD study data from the 1980s, it is unlikely that use of these oldest diaries would strongly influence exposure model estimates.

There is variability in the amount of outdoor time evaluated over the three decades, with diaries from the 2000's studies exhibiting perhaps the lowest range of mean outdoor time (190-220 min/day) compared with the 1980's (210-240 min/day) and 1990's (212-258 min/day) studies, a trend perhaps most notable when considering the children's diaries (a decrease in time spent outdoors of about 30 minutes over the three decades). However, the coefficient of variation (COV) for each of the age groups and across all decades for the cross-sectional data was consistently about 40%, supporting a general conclusion of no appreciable differences in the mean time spent outdoors over the three decades of data collection. Thus, when considering all diaries having at least 2 hours of afternoon outdoors time and the representativeness of the CHAD study data from the 1980s, inclusion of these earlier diaries is also unlikely to have a strong adverse influence on exposure modeling outcomes. Though combined with the higher participation rate for these earlier diaries, exposures estimated using these diaries may be higher than when estimated when excluding these diaries from CHAD.

5.4.1.4 *Evaluation of Afternoon Time Spent Outdoors for ATUS Survey Participants*

We evaluated recent year (2002-2011) time expenditure data from the American Time Use Survey (ATUS) (US BLS, 2012). As was done with the CHAD data set, the purpose was to evaluate trends (if any) in outdoor time over the period of time data were collected. A few strengths of the ATUS data are (1) its recent and ongoing data collection efforts, (2) large sample size (totaling over 120,000 diary days), (3) national representativeness, and (4) that varying diary approaches would not be an influential or confounding factor in evaluating trends over time.

ATUS does however have a few noteworthy limitations when compared with the CHAD data: (1) there are no survey participants under 15 years of age, (2) time spent at home locations is neither distinguished as indoors or outdoors, and (3) missing or unknown location data can comprise a significant portion of a persons' day (on average, about 40% (George and McCurdy, 2009)). To overcome the limitation afforded by the ambiguous home location, we identified particular activity codes most likely to occur outdoors (e.g., participation in a sport) to better approximate each ATUS individual's outdoor time expenditure. Missing time was circumvented by our focused analysis: about 85% of missing time information occurs outside of the hours of interest here (i.e., before 12:00 PM and after 8:00 PM). Data were stratified by the same five age groups as was done for the CHAD data, though here the time trends were assessed over individual survey years.

When considering person-days having at least 2 hours of time spent outdoors, there were no clear trends over the nine year ATUS study period regarding either the participation rate or the mean time spent outdoors for any of the age groups. Consistent with CHAD, the participation

rate of children was greater than that of the other age groups. The range in ATUS diary outdoor participation rate (10-20%) for all age groups is lower than that observed for the CHAD data (generally between 20-40%), while the range in mean time spent outdoors (190-240 minutes per day) was similar to that of the CHAD data. The lower participation rate for ATUS participants is not surprising given the lack of distinction regarding time indoors and outdoors while at home for ATUS participants and possibly influenced in part by not having any activity patterns for children under 15 years old. Overall, results of the ATUS data analysis generally support the representativeness of the CHAD data, and while participation in outdoor activities calculated using ATUS diaries was less than CHAD diaries, ATUS survey methods obfuscate the strength of this finding.

5.4.1.5 *Evaluation of Outdoor Time and Exertion Level for Asthmatics and Non-Asthmatics in CHAD*

Due to limited number of CHAD diaries with survey requested health information, all CHAD diaries are assumed appropriate for any APEX simulated individual (i.e., whether asthmatic, non-asthmatic, or no compromising health condition was indicated), provided they concur with age, sex, temperature, and day-of-week selection criteria. In general, the assumption of modeling asthmatics similarly to healthy individuals (i.e., using the same time-location-activity profiles) is supported by the activity analyses reported by van Gent et al. (2007) and Santuz et al. (1997), though other researchers, for example, Ford et al., (2003), have shown significantly lower leisure time activity levels in asthmatics when compared with persons who have never had asthma. To provide additional support to the assumption that any CHAD diary day can be used to represent the asthmatic population regardless of the study participants' characterization of having asthma or not, we first compared participation in afternoon outdoor activities at elevated exertion levels among asthmatic, non-asthmatic, and unknown health status using the CHAD diaries. We then compared compatible CHAD diary days with literature reported outdoor time participation at varying activity levels.

In the first comparison, participation in afternoon outdoor activities for non-asthmatic children and adults in CHAD were found similar when compared with their respective asthmatic cohorts (both about 40-50%). Outdoor participation rate for persons having unknown asthma status, a smaller fraction of the total diaries, varied ±10% from that having known asthma status (children were higher, adults were lower). The amount of time spent outdoors by the persons that did so varied little across the two populations and three asthma categories. On average, CHAD diaries from children indicate approximately 2¼ hours of afternoon time is spent outdoors, 80% of which is at a moderate or greater exertion level, again regardless of their asthma status, known or unknown. Slightly less afternoon time is spent outdoors by adults when compared with children, and while their participation in moderate or greater exertion level activities is much less

(about 63%), there was little difference between asthmatic adults and non-asthmatic adults considering outdoor time or percent at moderate or greater exertion.

For the second comparison, the percentage of waking hours outdoors at varying activity levels for asthmatics reported in three independent asthma activity pattern studies (Shamoo et al., 1994; EPRI, 1988; EPRI 1992) were compared to CHAD diary days having similar personal attributes and stratified by asthma status. The range in the percent of waking hours outside at moderate activity level for CHAD diaries was similar to that estimated using the three independent literature sources (2-10%), however the range in percent of outdoor time associated with strenuous activities using the CHAD asthmatic diaries extends beyond that of asthmatic persons from the three independent studies by about a factor of two higher. At this time, the reason for this difference is unknown. Overall, given the above mentioned similarities in outdoor time, participation, and activity levels, use of a CHAD diary regardless of a persons' asthma condition is reasonably justified based on the available data analyzed.

5.4.2 Characterization of Factors Influencing High Exposures

We investigated the factors that influence persons experiencing the highest daily maximum 8-hr average exposures. These exposure results in six selected study areas, Atlanta, Boston, Denver, Houston, Philadelphia, and Sacramento, considering base air quality and air quality just meeting the existing standard were combined with each simulated individual's microenvironmental time expenditure during the afternoon hours (12:00 PM through 8:00 PM), times of day commonly when daily peak high O_3 concentrations occur. We first evaluated the relative contribution seven variables[21] had on the total explained variability in daily maximum 8-hr average exposures. We then evaluated the distribution of identified influential variables for simulated individuals with the highest exposures. And finally, we identified the microenvironmental locations highly exposed persons occupied and the activities performed within them, given that within an 8-hr time frame most persons would likely visit multiple locations and perform different activities.

When considering only person days having the highest daily maximum 8-hr average O_3 exposures at any of the six study areas and either air quality scenario and age groupings, collectively the main effects of ambient concentrations and outdoor time combined with their interaction similarly contribute to approximately 80% of the total explained variance results, suggesting that for highly exposed persons, the most important influential factors are time spent outdoors corresponding with high daily maximum 8-hr average ambient O_3 concentrations.

[21] The seven variables include the main effects of (1) daily maximum 8-hr ambient O_3, (2-4) afternoon time spent outdoors, near-roads, and inside vehicles, and (5) physical activity index (PAI), while also including interaction effects from (6) afternoon time outdoors by daily maximum 8-hr ambient concentration and (7) PAI by afternoon time outdoors.

The distributions of afternoon outdoor time and ambient concentration for highly exposed individuals were evaluated considering base air quality and air quality adjusted to just meeting the existing standard. As an example, exposure results in Boston indicated that for about half of the days, simulated school-age children experiencing high exposures spend about 240 minutes outdoors during the afternoon hours along with experiencing daily maximum 8-hr average ambient O_3 concentrations $\geq$ 75 ppb. In contrast when adjusting ambient concentrations to just meeting the existing standard, for about half of the days, simulated school-age children experiencing similar high exposures need to spend about 280 minutes outdoors during the afternoon hours along with experiencing daily maximum 8-hr average ambient O_3 concentrations $\geq$ 60 ppb. Simply put, under conditions of lower ambient concentrations, persons need to spend a significantly greater amount of time outdoors to experience similar exposures observed at higher ambient concentration conditions.

When considering these highly exposed children, on average about half of children's total afternoon time is spent outdoors on high exposure days, 40% is spent indoors, while only 10% of time is spent near-roads or inside motor vehicles. In general, greater than half of the time highly exposed children spent outdoors specifically involves performing a moderate or greater exertion level activity, such as a sporting activity. While apportionment of afternoon microenvironmental time was similar for highly exposed adults in other age groups considered (e.g., 19-35), important high exertion activities performed outdoors also included those associated with paid work and performing chores.

5.4.3 Exposure Results for Additional At-Risk Populations and Lifestages, Exposure Scenarios, and Air Quality Input Data Used

5.4.3.1 *Exposures Estimated for All School-age Children During Summer Months, Neither Attending School or Performing Paid Work*

As mentioned earlier in describing the longitudinal approach used in the main body of the exposure assessment, the sequence of activity diaries for all simulated individuals is determined by a user-selected profile variable of interest. In this assessment our longitudinal diary approach uses time spent outdoors to link together CHAD diary days, an attempt to appropriately balance intra- and inter-personal variability in that variable. For the primary exposure results, all available diaries were used in developing any one sample pool without restriction outside of the particular characteristics on interest in developing the pool (i.e., age, sex, day-of-week, temperature, time spent outdoors). In this targeted simulation in Detroit during three summer months of 2007 (June, July, and August), we restricted the diary pool of all school-age children to include only those diary days that did not have any time spent inside a school nor had time spent performing paid work during any day of the week. The results of this targeted simulation were compared to an identical simulation, only differing in that all CHAD diary days were used

i.e., including any diary day for persons having school time or paid work, and as was done for the main body of this exposure assessment.

Figure 5-10 indicates that when restricting the CHAD diary pool to include only those diaries having no time spent at school or performing paid work activities, there is about 1/3 or 33% increase in the number of all school-age children at or above the 60 ppb-8hr benchmark, a relationship also consistent across the alternative standards and when considering multiple exposures. A similar relationship was found for the other benchmarks (not shown, see Appendix 5-G). Clearly, based on the analysis results reported in section 5.4.2 regarding factors influencing those highly exposed, using only activity pattern data that do not include school or work-related events (which would likely occur more so indoors than outdoors) and sampling from a pool of diaries consistent with summer temperatures would increase the likelihood simulated individuals spend time outdoors and be exposed to concentrations at or above the selected benchmarks.

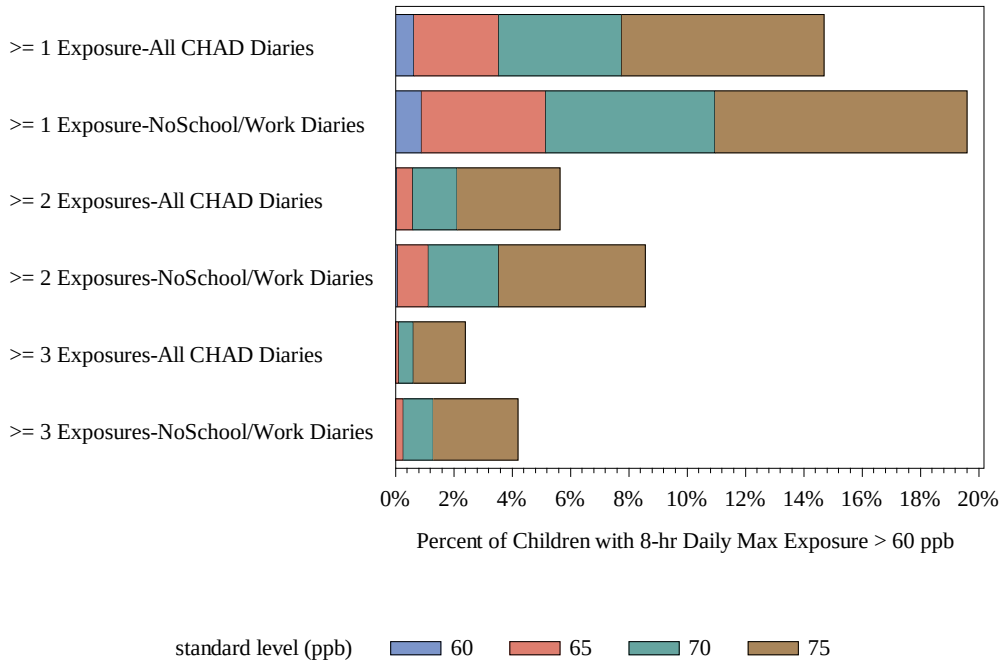


Figure 5-10 Comparison of the percent of all school-age children having daily maximum 8-hr average O_3 concentration at or above 60 ppb during June, July, and August in Detroit 2007: using any available CHAD diary ("All CHAD Diaries") or using CHAD diaries having no time spent in school or performing paid work ("No School/Work Diaries").

5.4.3.2 *Exposures Estimated for Outdoor Workers During Summer Months*

A targeted APEX simulation was performed for the Atlanta study area to simulate summertime exposures for two hypothetical outdoor worker study groups, persons between the age 19-35 and 36-55, using 2006 air quality just meeting the existing standard. To do this, both the daily and longitudinal activity patterns used by APEX were adjusted to best reflect patterns expected for outdoor workers (e.g., a standardized work schedule during weekdays) while also maintaining variability in those patterns across various occupation types. Briefly, the distribution of all employed persons' occupations was estimated using data provided by the U.S. Bureau of Labor and Statistics (US BLS, 2012b)[22] and linked with 144 occupation titles from the Occupational Information Network (O*NET)[23] identified as having one or more days per week where paid work was performed outdoors. These data were then aggregated to twelve broadly defined BLS occupation groups, generating a data set containing the number of days per week work time would be performed outdoors by that occupation group and properly weighted to reflect the population distribution of persons employed in each outdoor work group. Then, existing CHAD diary days reflecting outdoor paid work were identified, isolated and replicated to reflect this BLS/O*NET outdoor participation rate and occupation group frequencies. A 10,000 person simulation was performed by APEX using this adjusted CHAD activity pattern database designed to simulate outdoor workers and compared with exposure results generated from an identical APEX simulation of all employed persons, though differing by using the standard CHAD database and population-based modeling approach used in the main body REA. Details regarding the development of CHAD activity patterns used as input to simulate outdoor workers, as well as other settings and conditions for APEX is described in Appendix 5G.

Estimated exposures are presented in Figure 5-11 for one of two age study groups investigated (results for both age groups were similar) and considering either a longitudinal approach designed specifically to reflect an outdoor worker weekday schedule (left panel) or when using our general population-based modeling approach (right panel). The results indicate that when accounting for a structured schedule that includes repeated occurrences of time spent outdoors for a specified study group, all while simulated individuals are likely to be more consistently performing work tasks that may be at or above moderate or greater exertion levels, there are a greater percent of the study group experiences exposures at or above the selected health effect benchmark levels than that estimated using our general population-based modeling approach. Keep in mind outdoor workers are expected to experience more exposures at or above benchmark levels, though represent a fraction of the total employed population. It is possible

[22] U.S. employment data by SOC codes were obtained from: http://www.bls.gov/emp/#tables: Table 1.2 Employment by occupation, 2010 and projected 2020.
[23] Additional information is available at http://www.onetonline.org.

that, in using the general population-based approach along with the longitudinal algorithm that accounts for within and between variability in outdoor time, a number of outdoor workers are incidentally simulated and represent a significant portion of those who experienced exposures at or above benchmark levels. [24] However, the differences between exposures estimated for the two longitudinal approaches become much greater when considering the percent of persons experiencing multiple exposure days at or above benchmark levels, primarily when considering the 60 ppb-8hr benchmark level. For example, ≤ 2% of the general population-based exposure group was estimated to have two or more exposures at or above 60 ppb-8hr, while >17% of specifically simulated outdoor workers were estimated to experience exposures at or above that same level.

[24] In this outdoor worker exposure scenario, approximately 30% of our outdoor worker study group ages 19-55 were estimated to experience at least one exposure at or above 60 ppb-8hr while at moderate or greater exertion. Assuming outdoor workers constitute approximately 12% of the workforce (Appendix G, Table 5G-8), outdoor workers experiencing at least one exposure at or above 60 ppb-8hr could contribute 3.6% to a total exposed population (i.e., outdoor and non-outdoor workers). For the same air quality scenario and using the general population-based approach, we estimated 5-8% of a total employed study group (incidentally comprised of outdoor and non-outdoor workers) would experience exposures at or above the same benchmark, suggesting between 48-75% of persons experiencing exposures above the 60 ppb benchmark have similar activity pattern characteristics as outdoor workers.

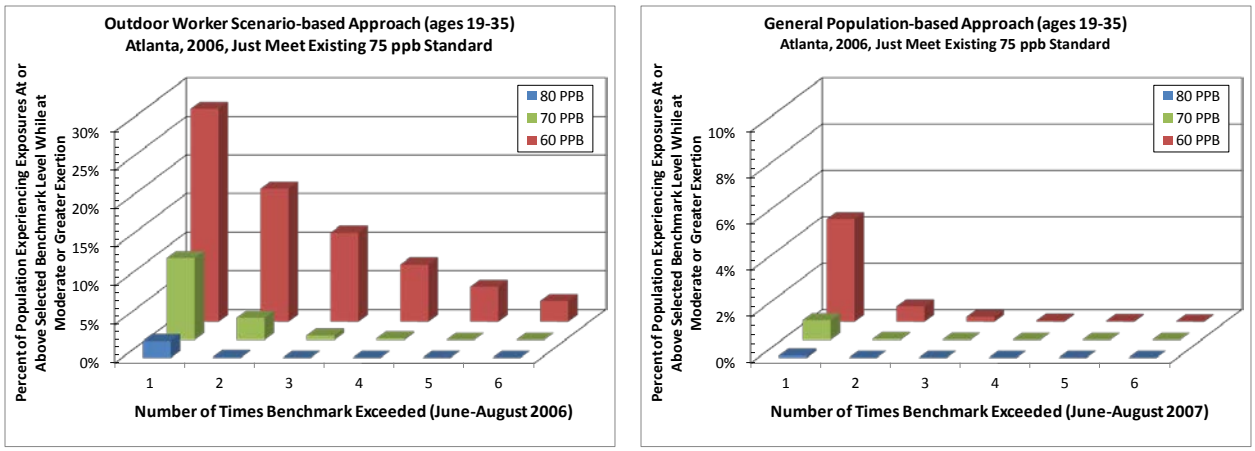


Figure 5-11 Percent of workers between ages 19-35 experiencing exposures at or above selected benchmark levels while at moderate or greater exertion using an outdoor worker approach (left panel) and a general population-based approach (right panel) considering air quality adjusted to just meet the existing standard in Atlanta, GA, Jun-Aug, 2006.

5.4.3.3 *Exposures Estimated for All School-age Children When Accounting for Averting Behavior*

A growing area of air pollution research involves evaluating the actions persons might perform in response to high O_3 concentration days (ISA, section 4.1.1). Most commonly termed *averting behaviors*, they can be broadly characterized as personal activities that either reduce pollutant emissions or limit personal exposure levels. The latter topic is of particular interest in this REA due to the potential negative impact it could have on O_3 concentration-response (C-R) functions used to estimate health risk and on time expenditure and activity exertion levels recorded in the CHAD diaries used by APEX to estimate O_3 exposures. To this end, we have performed an additional review of the available literature here beyond that summarized in the ISA to include several recent technical reports that collected and/or evaluated averting behavior data (Graham, 2012). The purpose was to generate a few reasonable quantitative approximations that allow us to better understand how averting behavior might affect time-location-activity patterns, and then simulate how such personal adjustments might affect our population exposure estimates.

Based on the elements evaluated in our literature review (i.e., air pollution awareness, prevalence and duration of averting response), we conclude that most people are aware of alert notification systems (in particular those persons having compromised health and reside in an urban area). We approximate that 30% of all asthmatics (or 15% of the general population) may reduce their outdoor activity level on alert days (e.g., KS DOH, 2006; McDermott et al., 2006; Wen et al., 2009; Zivin and Neidell, 2009) and that outdoor time/exertion during afternoon hours may be reduced by about 20-40 minutes in response to an air quality alert notification

(Bresnahan et al., 1997; Mansfield et.al, 2006, Neidell, 2010; Sexton, 2011). We used these literature derived estimates to generate an adjusted activity diary pool used by APEX to simulate a 2-day exposure period (August 1-August 2, 2007) in Detroit to approximate the effect averting may have on exceedances of exposure benchmarks.

When considering base air quality and our designed target to represent averting performed by the general population – 15.3 % of all simulated school-age children spent on average 44 minutes less time outdoors – resulting in approximately one percentage point or fewer children experienced exposures at or above any of the selected benchmark levels (Figure 5-12, left panel). When considering base air quality and our designed target to represent an averting response by the population of asthmatics – 30.3% of simulated asthmatic school-age children spent on average 44 minutes less time outdoors – resulting in approximately two percentage points or fewer experienced exposures at or above any of the selected benchmark levels (Figure 5-12, right panel).

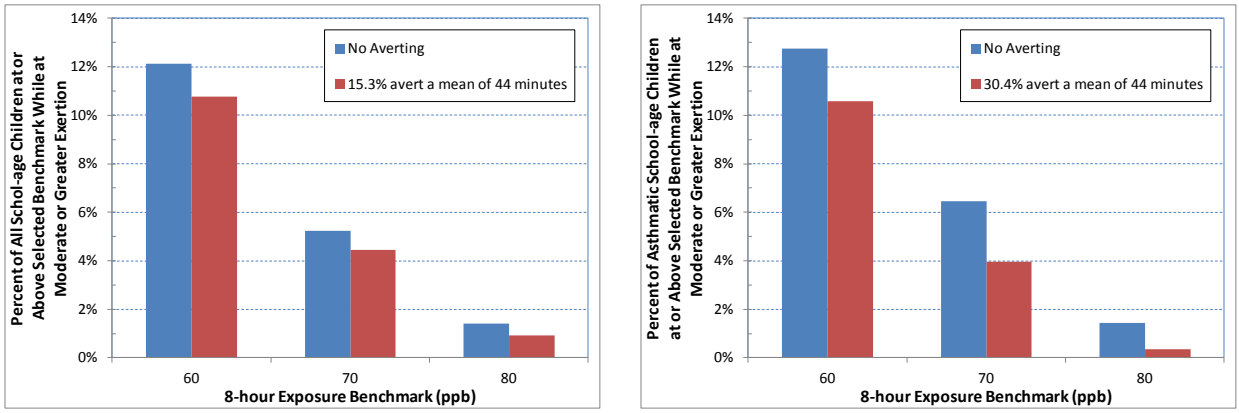


Figure 5-12 Percent of all school-age children (left panel) and asthmatic school-age children (right panel) having daily maximum 8-hr average O_3 concentration at or above benchmark levels during a 2-day simulation in Detroit, base air quality, August 1-2, 2007. Red bars indicate exposure results when considering effect of averting.

5.4.3.4 *Comparison of APEX Estimated Exposures Using Three Different Base Case Air Quality Data Sets: AQS, VNA, and EVNA*

For this exposure assessment, we elected to use a modeling approach to estimate the ambient input concentration field and better account for spatial gradients that may exist (Chapter 4). To support the selection of VNA, we compared exposure results separately generated using ambient monitor (AQS), eVNA, and VNA as input to APEX for three study areas: Atlanta, Detroit, and Philadelphia. All APEX settings were generally consistent with the simulations discussed previously, though the air quality data differed in that the year selected was 2005 (based on the available CMAQ data) and that a 4 Km grid was used to define the spatial area for this evaluation rather than census tracts. Daily maximum 8-hr average exposures were estimated

for asthmatic school-age children residing in the same census tracts comprising each air quality domain and summarized in Figure 5-13.

Exposure results for all three air quality input data sets were very comparable, with a few notable differences. Using AQS monitor concentration data tended to result in a 1-3% greater percent of asthmatic school-age children at or above each of the selected benchmark levels when compared with exposures estimated using VNA concentrations. While the VNA concentrations are based on the AQS monitor data, the approach generates a concentration gradient with distance from areas of known concentration that are typically less than the observed values, thus yielding fewer persons exposed to the highest concentrations. Using the eVNA approach to generate ambient concentrations tended to result in 2-5% greater percent of asthmatic school-age children at or above each of the selected benchmark levels when compared with exposures estimated using either the AQS or VNA approaches. This is because at times, the eVNA approach estimated high concentrations in areas where no observations were present, based on modeling which captures gradients in O_3 that may result from nearby sources (see Chapter 4).

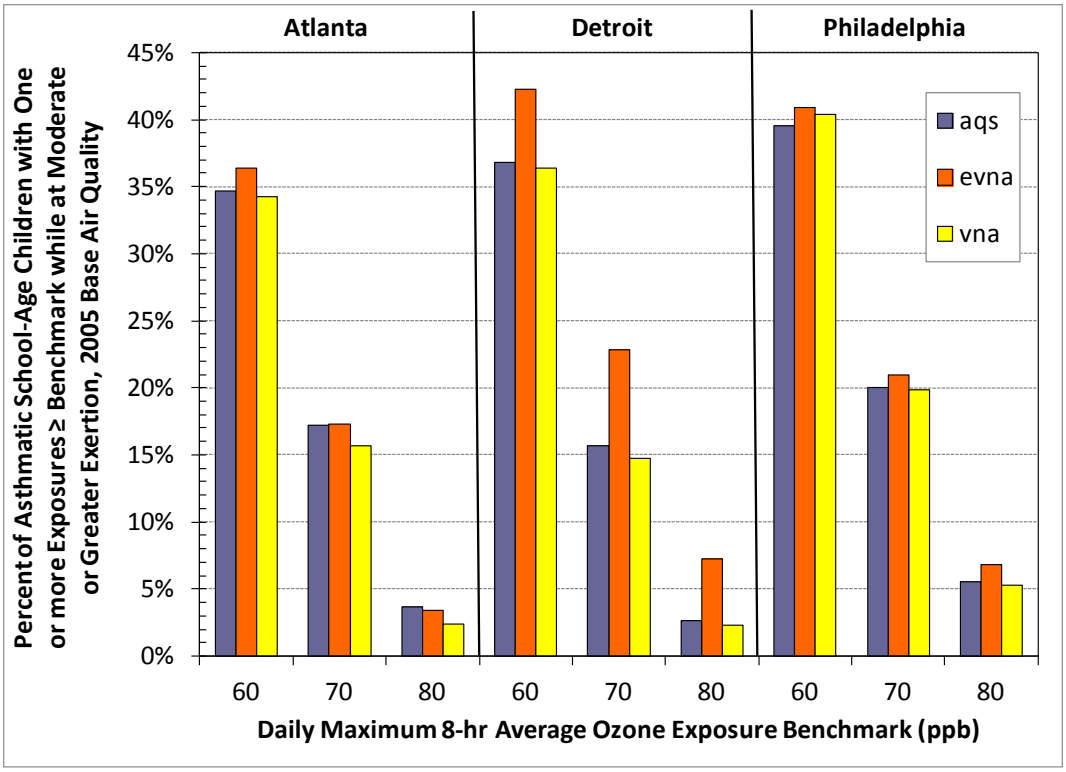


Figure 5-13 Comparison of APEX exposure results generated for three study areas (Atlanta, Detroit, and Philadelphia) using three different 2005 air quality input data sets: AQS, VNA, and eVNA.

5.4.3.5 *Comparison of APEX Estimated Exposures Using Two Different Adjusted Air Quality Data Sets: Quadratic Rollback and HDDM*

We elected to use an air quality modeling based approach rather than the previously used statistical approach to adjust air quality to just meet the current and alternative standard levels (Chapter 4). To support the selection of the HDDM approach, we compared exposure results for the scenario of just meeting the existing standard, separately generated using air quality inputs obtained using the quadratic rollback and HDDM method to adjust air quality for the Atlanta study area. All APEX settings were generally consistent with the simulations discussed previously, though both the air quality data sets used in this comparison differed from that done in the main exposure results above in that only the ambient monitor locations were used to define the air districts and assumed a 30 km radius of influence, as was done for the first draft REA. Daily maximum 8-hr average exposures were estimated for asthmatic school-age children in census tracts within 30 km of each air district and summarized in Figure 5-14.

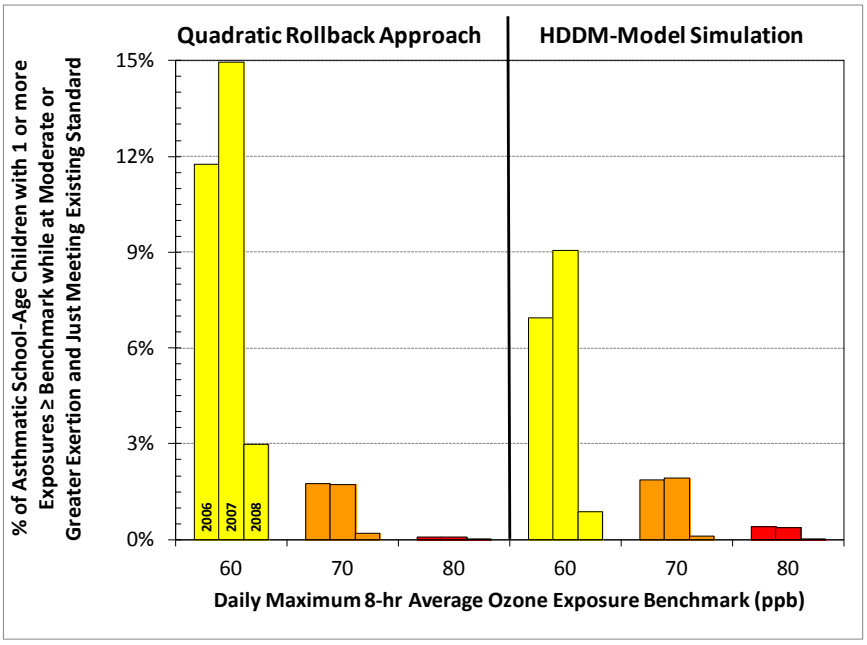


Figure 5-14 Comparison of exposure results generated by APEX using two different air quality adjustment approaches to just meet the existing standard in Atlanta: quadratic rollback (left panel) and HDDM (right panel).

The quadratic adjusted air quality resulted in slightly fewer percent of asthmatic school-age children exposed at or above the highest benchmark (80 ppb-8hr) when compared with exposures estimated using the HDDM model simulation approach, though a significantly greater percent of asthmatic school-age children were exposed to the lowest benchmark (60 ppb-8hr) using the quadratic approach. This is because the quadratic approach generally targets the highest

concentrations for adjustment, while the HDDM approach accounts for changes across the full concentration distribution to meet the adjusted concentration level of interest.

5.4.4 Limited Performance Evaluations

5.4.4.1 *Personal Exposure Comparisons*

A new evaluation of APEX was performed using a subset of personal O_3 exposure measurements obtained from the Detroit Exposure and Aerosol Research Study (DEARS) (Meng et. al, 2012). For five consecutive days, personal O_3 outdoor concentrations along with daily time-location activity diaries were collected from 36 adult study participants in Wayne County Michigan during July and August 2006. An APEX simulation was performed considering these same geographic and temporal features, followed with the sub-setting of APEX output data according to important personal attributes of the DEARS study participants (5-day collection study periods, age/sex distributions, outdoor time, ambient concentrations, and air exchange rate). A comparison sample was generated randomly from the complete simulation, selecting for 50 APEX simulated individuals.

For both data sets and considering the two output variables separately (outdoor time and daily exposure), the median daily values for each study participant were ranked, then plotted along with each individual's corresponding minimum and maximum value using each individual's 5 person-days of data (Figure 5-15). In spite of the distinct matching of influential personal attributes, over 50% of APEX simulated individuals had median daily O_3 exposure concentrations above 10 ppb, while only 3% of DEARS participants' median values exceeded 10 ppb. The reason(s) for this difference is being investigated.

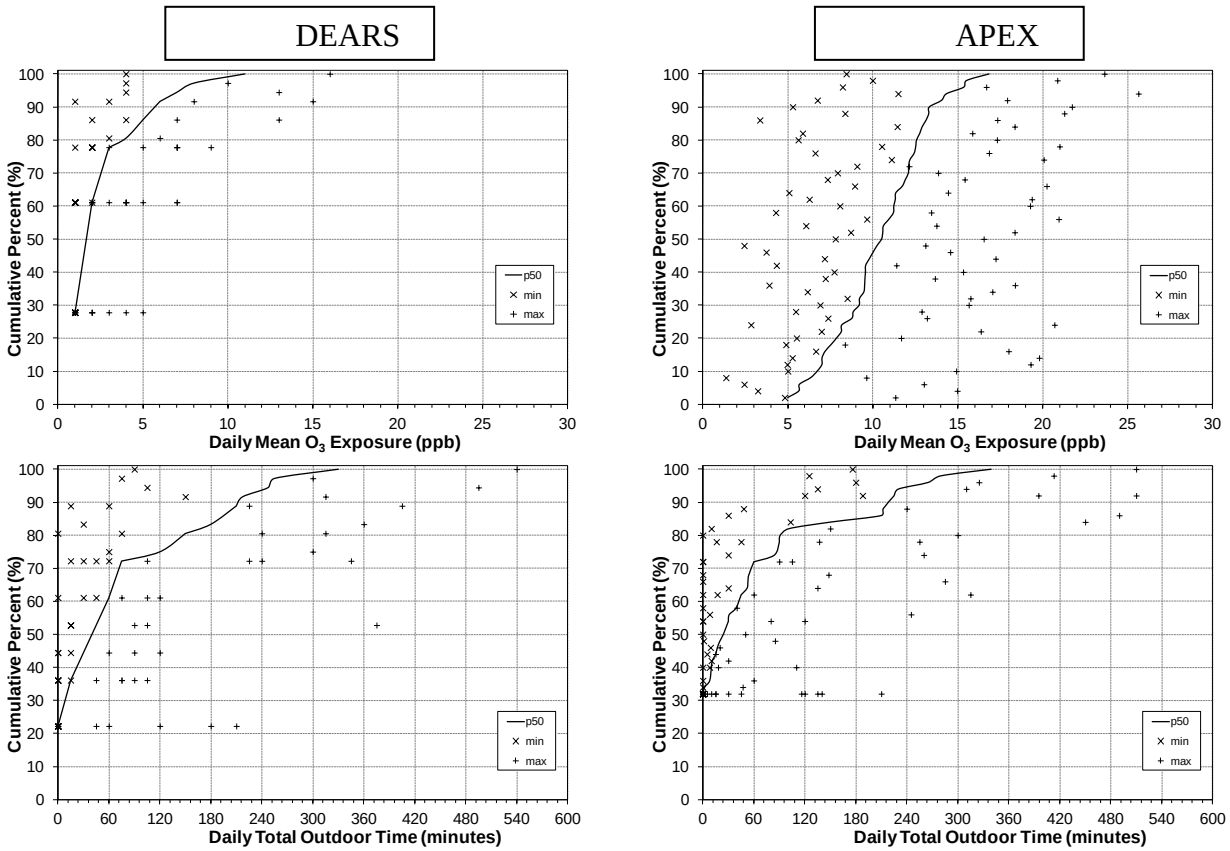


Figure 5-15 Distribution of daily average O_3 exposures (top panels) and daily afternoon outdoor time (bottom panels) and for DEARS study participants (left panels) and APEX simulated individuals (right panels) in Wayne County, MI, July-August 2006.

APEX modeled exposures have previously been compared with personal exposure measurements for O_3 (US EPA, 2007b). Briefly, APEX O_3 simulation results were compared with 6-day personal O_3 concentration measurements for children ages 7-12 (Xue et al., 2004; Geyh et al., 2000). Two separate areas of San Bernardino County were surveyed: urban Upland CA, and the combined small mountain towns of Lake Arrowhead, Crestline, and Running Springs, CA. Available ambient monitoring data for these locations during the same study years (1995-1996) were used as the air quality input to APEX. APEX predicted personal exposures, averaged similarly across a 6-day period, matched reasonably well for much of the concentration distribution considering both locations, but tended to underestimate exposures at the upper percentiles of the distribution. The average difference between the 6-day means was less than 1 ppb, with a range of -11 ppb to +8 ppb, though predicted upper bounds for a few averaged exposures having higher exposure concentrations were under-predicted by up to 24 ppb (e.g., Figure 5-16). In addition, modeled exposure concentration variability was less than that observed in the personal exposure measurements. At the time of analysis, these differences were proposed

to be largely driven by under-estimation of the spatial variability of the outdoor concentrations used by APEX (US EPA, 2007b).

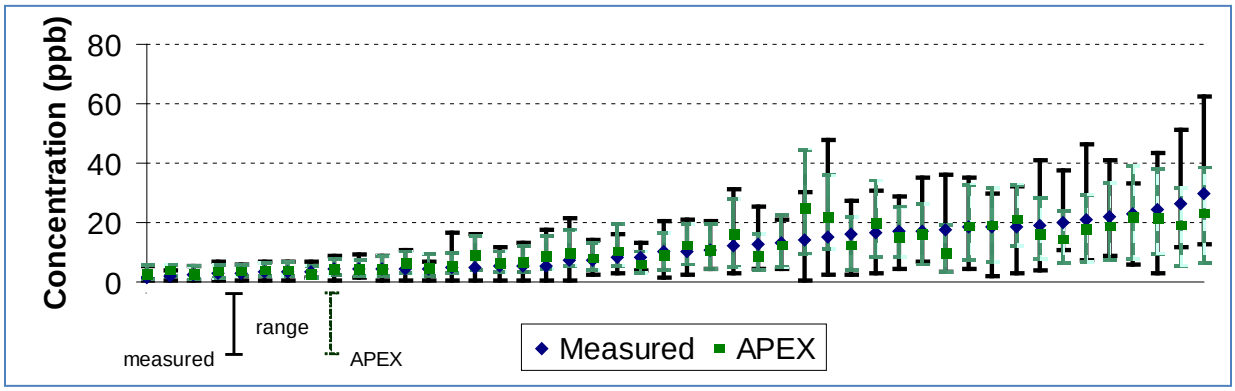


Figure 5-16 Means (and range) of 6-day average personal O_3 exposures, measured and modeled (APEX), Upland Ca. Obtained from Figure 8-22 of US EPA (2007b).

5.4.4.2 *Ventilation Rate Comparisons*

The algorithm used by APEX to estimate minute-by-minute ventilation rate serves as the basis for recent updates to the ventilation rate distributions provided in EPAs Exposure Factors Handbook (U.S. EPA, 2009b; US EPA, 2011). During the development of the ventilation distributions for EPA at that time, two peer-reviewed studies were identified as providing somewhat relevant measurement data to evaluate the APEX energy expenditure and ventilation algorithm (see Graham, 2009 for additional comparison details). The results of this evaluation are summarized below.

Briefly, Brochu et al. (2006a,b) presents data for ventilation rates derived from tracking doubly-labeled water (DLW) consumption/elimination to estimate energy expenditure in healthy normal-weight males and females, ages from 1 month to 96 years (n=1,252). Estimates of energy expended were combined with a fixed oxygen uptake factor (H=0.21) and using a fixed ventilatory equivalent (VQ)[25] of 27. The DLW measurement period ranged from 7-21 days, resulting in time-averaged metrics that may in some instances provide reasonable estimates for a mean daily ventilation rate, but not useful for estimating variability in an individual's ventilation rate over shorter time periods (as is needed by APEX). Further, while DLW is considered by some as a 'gold standard' for measuring energy expenditure, this characterization would not necessarily be directly transferable to approximations that use this measured value (i.e., ventilation rate in Brochu et al. (2006a,b) is a calculated value, not measured). Reported

[25] The ventilatory equivalent (VQ) is the ventilation rate (VE) divided by the oxygen consumption rate (VO_2)

ventilation rates are daily averages for several age groupings (e.g. ages 1 to < 2, 2 to < 5, 5 to < 7, etc.) along with derived percentiles, each assuming the existence of normally distributed data.

A 14-day APEX simulation was performed (i.e., the median of 7-21 days for the DLW measurement study) to estimate daily ventilation rates for comparison with the time-averaged Brochu et al (2006a) data. Twenty-five thousand persons were simulated by APEX to generate a reasonable number of persons within each year of age and other potential categorical variables (e.g., 100-200, although a few older age groups resulted in having fewer persons). It is important when comparing the two types of data for them to be similar as possible, particularly since age and body mass are important influential variables in both estimation methods. A total of 9,613 normal-weight individuals were simulated by APEX and used for the following analysis. Multi-day ventilation rates were averaged across the 14-day simulation period, yielding a mean daily ventilation rate for each person to best represent the DLW time averaging done by Brochu et al. (2006a).

Figure 5-17 compares the APEX simulated individuals body mass normalized mean daily ventilation rates with those reported by Brochu et al. (2006a; Table 2, page 684) for several age groupings of normal-weight individuals. The two largest differences appear for children of both sexes less than age 10 (i.e., Brochu et. al (2006a) estimates are systematically lower than APEX estimates) and for ages 16-33 (i.e., APEX estimates are lower than Brochu et al (2006a). Body mass normalized ventilation rates also appear to be slightly higher using APEX when considering persons above age 64 and for both sexes.

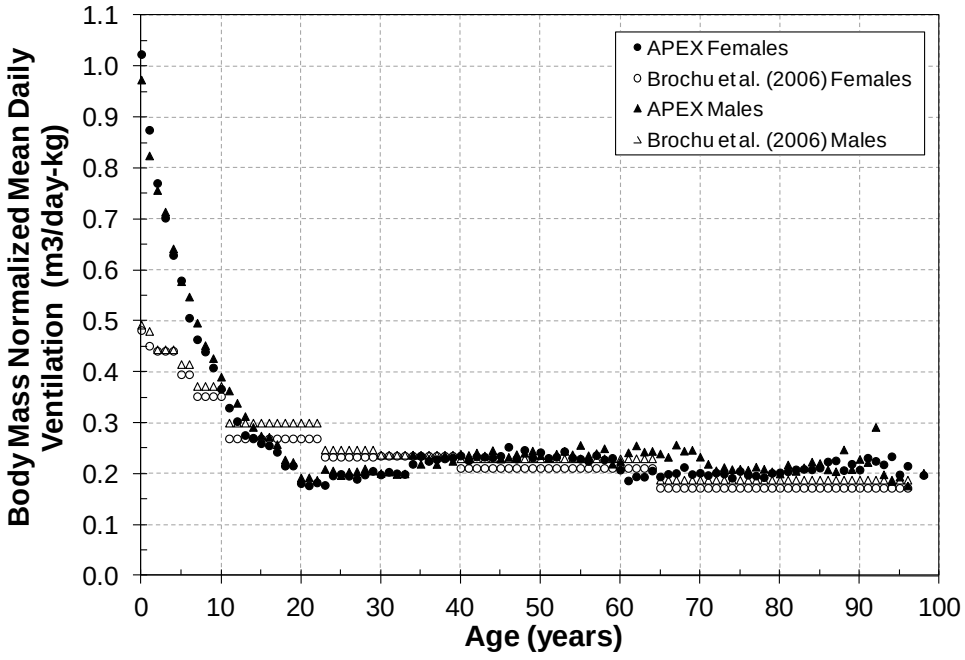


Figure 5-17 Comparison of body mass normalized mean daily ventilation rates estimated by APEX (closed symbols) and by Brochu et al., 2006 (open symbols).

One principal issue identified by us as potentially responsible for some of the above differences in ventilation estimates is in the VQ used by Brochu et al. (2006a). A single value of 27 was used in estimating ventilation rates for both children and adults, however it is widely recognized that while a VQ of 27 may be a reasonable approximation for estimating mean ventilation rates of adults, it is not appropriate for use in estimating mean ventilation rates in children. With this in mind, the Brochu et al. (2006a) ventilation estimates were modified here using the VQ estimates offered by Arcus-Arth and Blaisdell (2007). Figure 5-18 illustrates the comparison of APEX body mass normalized mean daily ventilation rates with that of Brochu et al. (2006a) corrected ventilation estimates. The body mass normalized ventilation estimates for school-age children are more similar to those generated by APEX when correcting the Brochu et al (2006a) VQ parameter. Thus, mean ventilation rates generated by APEX are reasonably correlated with independent measures from the Brochu et al. (2006a, b) estimates, particularly when correcting the Brochu et al (2006a) ventilation estimates for children using a more appropriate estimate of VQ for children.

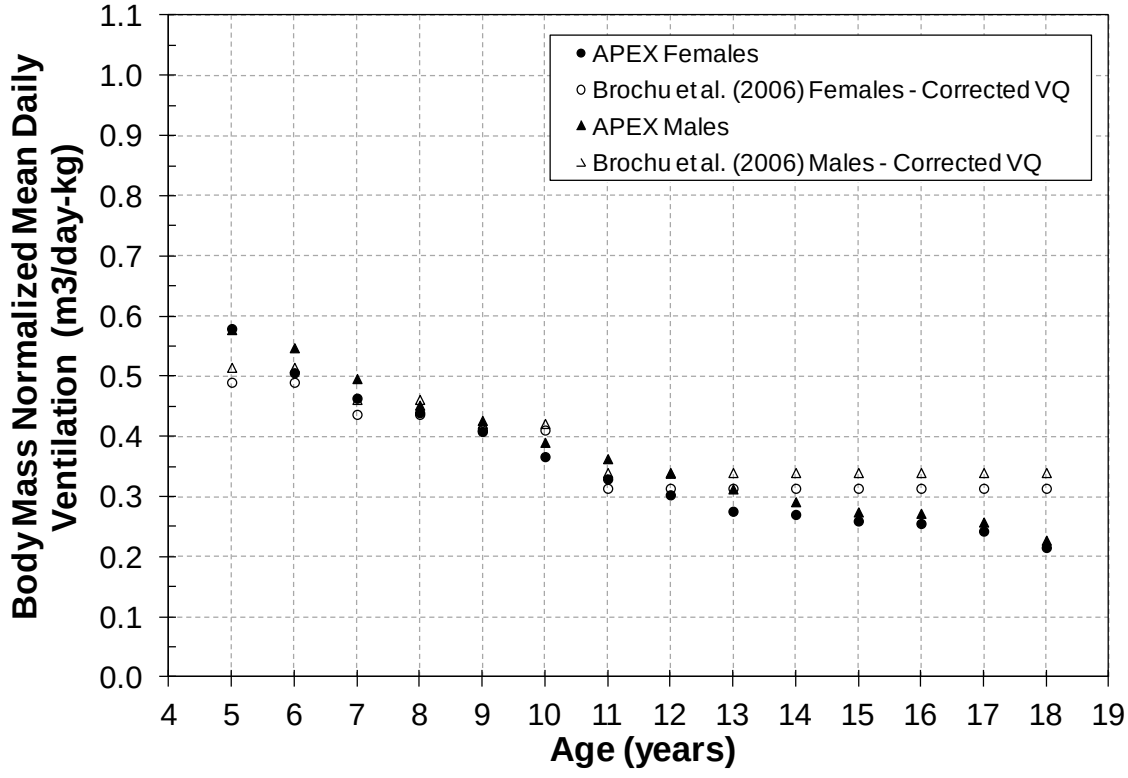


Figure 5-18 Comparison of body mass normalized mean daily ventilation rates in male and female school-age children (5-18) when correcting Brochu et al. (2006a) results with child appropriate VQ estimates.

In a second study identified for comparison with APEX estimates, Arcus-Arth and Blaisdell (2007) provide ventilation estimates for children <19 years of age using energy intake (EI, or calories consumed) and body mass data provided by the USDA's Continuing Survey of Food Intake for Individuals (CSFII; USDA, 2000). Two-day daily average EIs were combined with a values of H (i.e., 0.22 for infants, 0.21 for non infants) and VQ (i.e., 33.5 for children 0-8, 30.6 for boys 9-18, 31.5 for girls 9-18 years old). Again, time-averaging of the data may provide reasonable estimates of a daily mean, but offer no variability in ventilation estimates for shorter durations. Furthermore, data for both sexes are combined and reported by age, with stratified results by sex reported only for aggregated age groups (males and females, 9-18 years old).

A 2-day model simulation was performed by APEX to generate ventilation estimates for children to compare with results of Arcus-Arth and Blaisdell (2007).[26] APEX ventilation estimates were time-averaged to generate mean daily values, and since the data reported in Arcus-Arth and Blaisdell (2007) were not separated by sex (outside of broad age categories), the APEX estimates were also combined by sex to provide a comparable mean estimate for each year of age (5-18). Body mass was also not used as a categorical variable in Arcus-Arth and Blaisdell (2007), therefore all APEX simulated individuals were used, regardless of whether they could be classified as overweight or of normal weight. In addition, daily ventilation rates for a few age groups of children were obtained from Tables 3 and 4 of Brochu et al. (2006a), though considering both estimates for normal and overweight individuals (there were no combined data available). The Brochu et al. (2006a) results have been corrected for VQ as noted above using VQ estimates of Arcus-Arth and Blaisdell (2007) and added for comparison.

Figure 5-19 illustrates ventilation rate estimates from the APEX simulation, along with associated data for school-age children (ages 5-18) obtained from the two publications. Daily mean ventilation estimates are quite similar at each year of age, with slightly higher estimates by Arcus-Arth and Blaisdell (2007) at ages 9 and above, particularly when compared with APEX ventilation estimates. Ventilation estimates are remarkably similar for school-age children for all three sources of data, particularly when considering the differences in the type of input data used and the varied approaches of APEX, Brochu et al. (2006a), and Arcus-Arth and Blaisdell (2007). This overall agreement suggests reasonable confidence can be conferred to the algorithm used by APEX to estimate at a minimum, daily mean ventilation rates.

[26] Table III, page 103 of Arcus-Arth and Blaisdell (2007) provided body mass normalized ventilation rates.

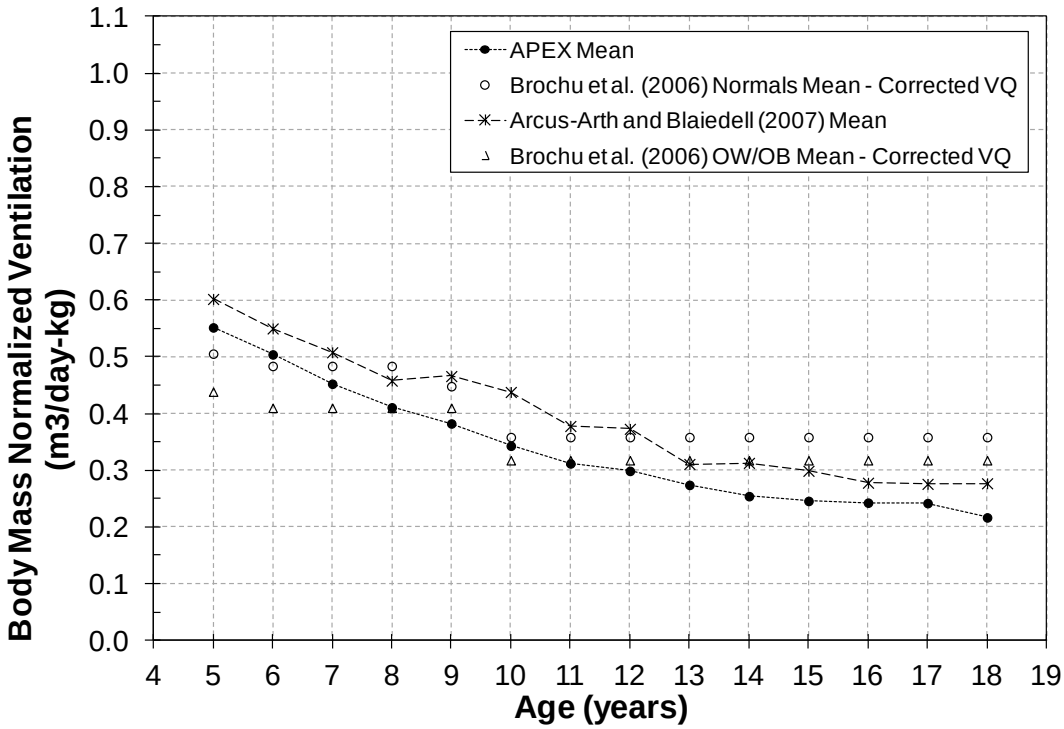


Figure 5-19 Comparison of body mass normalized daily mean ventilation rates in school-age children (5-18) estimated using APEX and literature reported values.

5.4.4.3 *Evaluation of Longitudinal Profile Methodology*

We evaluated the APEX approach used for linking together cross-sectional activity pattern diaries to generate longitudinal profiles for our simulated individuals (Appendix 5G, Section 5G-3). Of particular interest were how well variability in outdoor participation rate and the amount of time expended were represented in our population-based exposure simulations. Our goal in developing the most reasonable longitudinal profiles is to capture expected, important features of population activity patterns, i.e., there is correlation within an individual's day-to-day activity patterns (though neither exactly repeated nor entirely random for individuals) and variability across the modeled study group in day-to-day activity patterns (i.e., not every simulated individual in the study group does the same activity on the same day).

The simulated longitudinal profiles indicate the method for linking together cross-sectional diaries generates a diverse mixture of persons having variable, though expected, activity patterns: A small fraction of the simulated population spend a limited amount of afternoon time outdoors and occurring at a low frequency across an O_3 season, a small fraction consistently spends a greater amount (> 2 hours) of time outdoors and occurring at greater frequency (e.g., 4/5 days per week), while the remaining simulated individuals fall somewhere in between regarding participation and total time. While we are not aware of a population database

available to compare with these simulated results, we are comfortable with the method performance in representing the intended variability in longitudinal activity patterns (see section 5G-3 for details).

5.5 VARIABILITY AND UNCERTAINTY

An important issue associated with any population exposure or risk assessment is the characterization of variability and uncertainty. *Variability* refers to the inherent heterogeneity in a population or variable of interest (e.g., residential air exchange rates). The degree of variability cannot be reduced through further research, only better characterized with additional measurement. *Uncertainty* refers to the lack of knowledge regarding the values of model input variables (i.e., *parameter uncertainty*), the physical systems or relationships used (i.e., use of input variables to estimate exposure or risk or *model uncertainty*), and in specifying the scenario that is consistent with purpose of the assessment (i.e., *scenario uncertainty*). Uncertainty is, ideally, reduced to the maximum extent possible through improved measurement of key parameters and iterative model refinement. The approaches used to assess variability and to characterize uncertainty in this REA are discussed in the following two sections. Each section also contains a concise summary of the identified components contributing to uncertainty and how each source may affect the estimated exposures.

5.5.1 TREATMENT OF VARIABILITY

The purpose for addressing variability in this REA is to ensure that the estimates of exposure and risk reflect the variability of ambient O_3 concentrations, population and lifestage characteristics, associated O_3 exposure and dose, and potential health risk across the study area and for the simulated at-risk study groups. In this REA, there are several algorithms that account for variability of input data when generating the number of estimated benchmark exceedances or health risk outputs. For example, variability may arise from differences in the population residing within census tracts (e.g., age distribution) and the activities that may affect population and lifestage exposure to O_3 (e.g., time spent inside vehicles, time performing moderate or greater exertion level activities outdoors). A complete range of potential exposure levels and associated risk estimates can be generated when appropriately addressing variability in exposure and risk assessments; note however that the range of values obtained would be within the constraints of the input parameters, algorithms, or modeling system used, not necessarily the complete range of the true exposure or risk values.

Where possible, we identified and incorporated the observed variability in input data sets to estimate model parameters within the exposure assessment rather than employing standard default assumptions and/or using point estimates to describe model inputs. The details regarding

variability distributions used in data inputs are described in Appendix 5B. To the extent possible given the data available for the assessment, we accounted for variability within the exposure modeling. APEX has been designed to account for variability in some of the input data, including the physiological variables that are important inputs to determining ventilation rates. As a result, APEX addresses much of the variability in factors that affect human exposure. Important sources of the variability accounted for in this analysis are summarized in Appendix 5D.

5.5.2 CHARACTERIZATION OF UNCERTAINTY

While it may be possible to capture a range of exposure or risk values by accounting for variability inherent to influential factors, the true exposure or risk for any given individual within a study area is unknown, though can be estimated. To characterize health risks, exposure and risk assessors commonly use an iterative process of gathering data, developing models, and estimating exposures and risks, given the goals of the assessment, scale of the assessment performed, and limitations of the input data available. However, significant uncertainty often remains and emphasis is then placed on characterizing the nature of that uncertainty and its impact on exposure and risk estimates.

The REA's for the previous O_3, NO_2, SO_2, and CO NAAQS reviews each presented a characterization of uncertainty of exposure modeling (Langstaff, 2007; US EPA 2008, 2009a, 2010). The qualitative approach used in this and other REAs is described by WHO (2008). Briefly, we identified the key aspects of the assessment approach that may contribute to uncertainty in the exposure and risk estimates and provided the rationale for their inclusion. Then, we characterized the *magnitude* and *direction* of the influence on the assessment results for each of these identified sources of uncertainty. Consistent with the WHO (2008) guidance, staff scaled the overall impact of the uncertainty by considering the degree of uncertainty as implied by the relationship between the source of uncertainty and the exposure concentrations. A qualitative characterization of low, moderate, and high was assigned to the magnitude of influence and knowledge base uncertainty descriptors, using quantitative observations relating to understanding the uncertainty, where possible. A summary of the key findings of those prior characterizations that are most relevant to the current O_3 exposure assessment are provided in Table 5-6.

Table 5-6 Characterization of Key Uncertainties in Historical and Current APEX Exposure Assessments

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	Element	Direction	Magnitude			
Ambient Monitoring Concentrations	Database Quality	Both	Low	Low	All ambient pollutant measurements available from AQS are both comprehensive and subject to quality control.	Yes. No further characterization needed.
	Instrument Measurement Error	Over	Low	Low	Mean bias estimated as 1.2% (CV of 4.4%). See Table 2 and Figure 6 of Langstaff (2007).	Yes. No further characterization needed.
	Missing Data Substitution Method	Both	Low	Low	Overall completeness of data yield negligible mean bias (~0) along with an estimated standard deviation of 4 ppb when replacing missing values. See Table 3 of Langstaff (2007).	Yes. No further characterization needed.
	Temporal Representation	Both	Low	Low	Appropriately uses 1-hr time-series of O_3 concentrations for 5 years. No missing data for any hour input to APEX.	Yes. No further characterization needed.
	Spatial Representation: Large Scale	Both	Low	Low	Tens of monitors used in each study area.	Yes. No further characterization needed.
	Spatial Representation: Neighborhood Scale (1)	Both	Low	Low	Spatial interpolation using jackknife method (removal of a single monitor) yielded generally unbiased observed/predicted ratios (mean 1.06), having an estimated standard deviation of 0.2. Langstaff (2007).	Yes. For the uncertainties characterized, the historical rating is appropriate if and when using ambient monitor data alone to represent air quality surface. However in this 2nd draft REA, local-scale air quality was estimated using VNA (see below).
	Spatial Representation: Neighborhood Scale (2)	Over	Low	Low	When reducing the APEX radius setting from an unlimited value (actual value used) to 10 km (i.e., the tendency would be to more accurately represent exposure), a smaller fraction (1-3 percentage points) of population exceeds benchmark levels. See Figures 7 – 9 of Langstaff (2007).	

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	Element	Direction	Magnitude			
	Spatial Representation: Local Scale VNA estimates	Both	Low	Low - Moderate	Scenario-based evaluation in three study areas indicated small differences in exposure results when comparing ambient monitor data or statistically interpolated concentrations to 4 Km grid as an input to APEX (Figure 5-13). General dependencies of the approaches used could lead to observed lack of distinction in exposure results.	Yes. Newly evaluated.
	Spatial Representation: Vertical Profile	Both	Moderate	Moderate	Differences between ground-level (0-3 meters) and building rooftop sited (25 meters) monitor concentrations can be significant. Most importantly, use of higher elevation monitors would tend to overestimate ground-level exposures (i.e., persons outdoors).	Yes. Given judged impact to exposure, additional characterization is possibly warranted.
Adjustment of Air Quality to Simulate Just Meeting the Current Standard	Quadratic Approach	Both	Low - Moderate	Moderate	Variable differences (e.g., none to a factor of two or three) in the estimated number of persons exposed across study areas when using differing 3-year roll-back periods for a single year of air quality (Langstaff, 2007).	Yes. Uncertainty in the approach has resulted in use of HDDM approach.
	HDDM Simulation Approach	Both	Low - Moderate	Low - Moderate	Expected patterns in both air quality and exposure result from HDDM/emissions reduction approach (full distribution affected rather than only upper percentiles, Figure 5-14). Variable differences remain (e.g., none to a factor of two or three) in the estimated percent of persons exposed across study areas when using differing 3-year roll-back periods for 2008 air quality (Figures 5-5 to 5-9). New York study area could not be simulated to just meet 60 and 55 ppb alternative standards.	Yes. Newly evaluated.
APEX: General Input Databases	Population Demographics and Commuting (US Census)	Under	Low	Low	Comprehensive and subject to quality control. Differences in 2000 data versus modeled years (2006-2010) are likely small when estimating percent of population exposed.	Yes. No further characterization needed.

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	Element	Direction	Magnitude			
	Activity Patterns (CHAD)	Both	Low - Moderate	Low-Moderate	Comprehensive and subject to quality control. Significantly increased number of diaries used to estimate exposure from prior review and 1st draft REA for this review (Table 5-3). Thoroughly evaluated trends and patterns in historical data – no major issues noted with use of historical data to represent current patterns (Figures 5G-1 and 5G-2). Compared outdoor participation and time with ATUS data base – CHAD participation is higher than ATUS, likely due to ATUS survey methods. Activity data for asthmatics generally similar to non-asthmatics (Tables 5G2-to 5G-5). Remaining uncertainty with other influential factors that cannot be accounted for (e.g., SES, region/local outdoor participation rates)	Yes. Newly evaluated.
	Meteorological (NWS)	Both	Low	Low	Comprehensive and subject to quality control, few missing values. Limited application in selecting CHAD diaries and AERs.	Yes. No further characterization needed.
	Poverty Status (US Census) Weighted Asthma Prevalence (CDC)	Both	Low	Low	Data used are from a peer-reviewed quality controlled source. Application accounts for variability in most important influential variables (age, sex, region, poverty) though possible that variability in microscale prevalence not entirely represented.	New. Could possibly use further characterization, though typically available local prevalence rates are not well stratified by influential variables.

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	**Element**	**Direction**	**Magnitude**			
APEX: Microenvironmental Concentrations	Outdoor Near-Road and Vehicular: Proximity Factors	Both	Low	Low-Moderate	Uncertainty in mean value used approximated as 15 percentage points. See Figure 10 and Table 7 of Langstaff (2007). May be of greater importance in certain study areas or under varying conditions, though even with this mean difference, in-vehicle penetration/decay decreases exposures and hence importance of in-vehicle microenvironments.	Yes. No further characterization needed.
	Indoor: Near-Road	Over	Low	Low	Expected reduction in O_3 for persons residing near roads not modeled here, but when included, there is a small reduction (~3%) in the number of persons experiencing exposure above benchmark levels (Langstaff, 2007).	Yes. No further characterization needed.
	Indoor: Air Exchange Rates	Both	Low	Moderate	Uncertainty due to random sampling variation via bootstrap distribution analysis indicated the AER GM and GSD uncertainty for a given study area tends range to at most from fitted ±1.0 GM and ± 0.5 GSD hr^{-1}. Non-representativeness remains an important issue as city-to-city variability can be wide ranging (GM/GSD pairs can vary by factors of 2-3) and data available for city-specific evaluation are limited (Langstaff, 2007). Also, indoor exposures are estimated as not important to daily maximum 8-hr average O_3 exposure.	Yes. No further characterization needed.

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	Element	Direction	Magnitude			
	Indoor: A/C Prevalence (AHS)	Both	Low	Low	Comprehensive and subject to quality control, estimated 95th percentile confidence bounds range from a few to just over ten percentage points, though some cities use older year data (Table 9 of Langstaff, 2007). Note, variable indicates presence/absence not actual use. Also, indoor exposures are estimated here as limited in importance to daily maximum 8-hr average exposures and sensitivity analyses in NO_2 REA (in-vehicle was most influential exposure ME) concluded indoor prevalence variable was of limited importance.	Yes. No further characterization needed.
	Indoor: Removal Rate	Both	Low	Low	Greatest uncertainty in the input distribution regarded representativeness, though estimated as unbiased but correct to within 10% (Langstaff, 2007).	Yes. No further characterization needed.
	Vehicular: Penetration Factors	Both	Low	Moderate	Input distribution is from an older measurement study though consistent with recent, albeit limited data.	Yes. No further characterization needed.

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	Element	Direction	Magnitude			
APEX: Simulated Activity Profiles	Longitudinal Profiles	Under	Low - Moderate	Moderate	Depending on the longitudinal profile method selected, the number of persons experiencing multiple exposure events at or above a selected level could differ by about 15 to 50% (see Appendix B, Attachment 4 of NO_2 REA). Long-term diary profiles (i.e., monthly, annual) do not exist for a population, limiting the evaluation. The general population-based modeling approach used for main body REA results does not assign rigid schedules, for example explicitly representing a 5-day work week for employed persons. However, when considering such scheduling (e.g., outdoor workers or all children spending entire summer season not in-school), estimated exposures are greater than when not considering rigid weekly/seasonal schedules. For our hypothetical outdoor worker scenario, the number of multiday exposures at or above benchmark levels was primarily affected (though mainly the 60 ppb level, Figure 5-11), while both percent of children experiencing single and multiday exposures were increased by about 30% when simulating a rigid schedule (Figure 5-10).	Yes. Newly evaluated.
	Commuting	Both	Low	Moderate	New method used in this assessment is designed to link Census commute distances with CHAD vehicle drive times. Considered an improvement over the former approach that did not match distance and time. While vehicle time accounted for through diary selection, not rigidly scheduled. However, In-vehicle exposures are not important drivers for persons exceeding benchmark levels (section 5.3.2).	Yes. Newly evaluated.

Sources of Uncertainty		Historical Uncertainty Characterization				Is rating appropriate for current APEX O_3 exposure assessment?
		Influence of Uncertainty on Exposure/Intake Dose Estimates		Knowledge-base Uncertainty	Comments	
Category	**Element**	**Direction**	**Magnitude**			
	At-Risk Population and Lifestages	Both	Low	Low – Moderate	An updated evaluation shows activity patterns of asthmatics are similar to that of non-asthmatics (section 5.3.1, Tables 5G-2 to 5G-5).	Yes. Newly evaluated.
APEX: Physiological Processes	Body Mass (NHANES)	Unknown	Low	Low	Comprehensive and subject to quality control, though older (1999-2004) than current simulated population, possible small regional variation is not represented by national data.	Yes. No further characterization needed.
	NVO_{2max}	Unknown	Low	Low	Upper bound control for unrealistic activity levels rarely used by model, thus likely not very influential.	Yes. No further characterization needed.
	RMR	Unknown	Low	Low	Approach from older literature (Schofield, 1985), used in ventilation equation. Note ventilation rate estimates are reasonable.	Newly identified. May need additional characterization.
	METS distributions	Over	Low - Moderate	Low - Moderate	APEX estimated daily mean METs range from about 0.1 to 0.2 units (between about 5-10%) higher than independent literature reported values (Table 15 of Langstaff, 2007). Shorter-term values are of greater importance in this assessment.	Yes. Given judged impact to exposure, additional characterization is needed.
	Ventilation rates	Over	Low - Moderate	Low - Moderate	APEX estimated daily ventilation rates can be greater (2-3 m^3/day) than literature reported measurement values (Table 25 of Langstaff, 2007), though if accounting for measurement bias this minimizes the discrepancy (Graham and McCurdy, 2005; see Figures 5-18 and 5-19). Also, a shorter-term comparison (hours rather than daily), while more informative, cannot be performed due to lack of data.	Yes. Additional characterization would be warranted if minute or hourly ventilation rate data were available.
Exposure Benchmark Level	EVR characterization of moderate or greater exertion	Over	Moderate	Low - Moderate	Given that the EVR serves as a cut point for selecting persons performing at moderate or greater exertion and is a lower bound value (~5th percentile), the simulated number of persons achieving this level of exercise is possibly overestimated.	Newly identified. May need additional characterization.

5.6 KEY OBSERVATIONS

Two additional tables are provided to additionally summarize the exposure results across all study areas and years of air quality data: Table 5-7 contains the percent of all school-age children experiencing at least one exposure at or above the three exposure benchmark levels, while Table 5-8 contains the percent of all school-age children experiencing at least two exposures at or above the three exposure benchmark levels, with both tables considering results associated with each of the adjusted air quality scenarios.[27] Two descriptive statistics are provided from the exposure results for each study area: the mean percent of persons exposed in each study area averaged across the 5 years simulated and the maximum percent of persons exposed in each study area, representing the worst year of air quality simulated. Figure 5-20 illustrates the estimated mean and maximum percent of all school-age children exposed for each study area when considering the 60 ppb-8hr benchmark and adjusted air quality scenarios, and using the data provided in Table 5-7 and Table 5-8.

Presented below are key observations resulting from the O_3 exposure analysis:

- **General**: The estimated percent of any study group exposed at least once at or above the selected benchmark levels were highest considering the base air quality though percent exposed varied by study area, year, and benchmark level (Appendix 5F). Very few persons within any study group (all are estimated to be < 0.3%) experienced any benchmark exceedances when considering an alternative standard level of 55 ppb-8hr (data not shown).
- **Study Group:** The percent of all school-age children exposed at or above the selected benchmark levels across all study areas, years, and air quality scenarios were similar to exposures for asthmatic school-age children (e.g., Figure 5-5 and Figure 5-6, respectively) with both of these study groups having consistently higher percent of persons exposed than that estimated for asthmatic adults and all older adults (Figure 5-7 and Figure 5-8, respectively), generally by about a factor of three or more. The percent of all older adults at or above any benchmark level tended to be only a few percentage points or less when compared with corresponding benchmark exceedances for asthmatic adults.
- **80 ppb-8hr Exposure Benchmark:** In general, less than 1% of any study group, including all school-age children and any study area, was exposed at least once at or above the highest exposure benchmark, 80 ppb-8hr, when considering the existing

[27] The maximum sample size is 6 years based on years simulated, and for a few instances varied based on available air quality (e.g., Chicago does not have 3 years simulated for just meeting the current standard during 2008-2010 period because air quality was below the current standard, thus the total sample size for this study area is 3.

standard air quality scenario (Table 5-7). When considering a standard level of 70 ppb-8hr, ≤ 0.2% of any study group and any study area was exposed at least once at or above that same benchmark.

- **70 ppb-8hr Exposure Benchmark:** Less than 10% of any study group, including all school-age children and any study area, was exposed at least once at or above an exposure benchmark of 70 ppb-8hr, when considering the existing standard air quality scenario (Table 5-7). When considering a standard level of 70 ppb-8hr, ≤ 3.5% of any study group and in any study area was exposed at least once at or above that same benchmark. A standard level of 65 ppb-8hr is estimated to reduce the percent of persons at or above an exposure benchmark of 70 ppb-8hr to ≤0.5% of any study group and in any study area.
- **60 ppb-8hr Exposure Benchmark:** In general, no more than 26% of any study group in any study area was exposed at least once at or above the lowest exposure benchmark, 60 ppb-8hr, when considering the existing standard air quality scenario (Table 5-7, Figure 5-20). When considering a standard level of 70 ppb-8hr, < 20% of any study group in any study area was exposed at least once at or above that same benchmark. A standard level of 65 ppb-8hr is estimated to reduce the percent of persons at or above an exposure benchmark of 60 ppb-8hr to ≤ 10% of any study group and study area.
- **Multi-day Benchmark Exceedances:** When considering air quality adjusted to just meet the existing standard, multi-day exposure benchmark exceedances are largely limited to two or more exceedances at the 60 ppb-8hr benchmark, all occurring for < 15% of any study group in any study area (e.g., Table 5-8, Figure 5-9). There were no persons estimated to experience any multi-day exposures at or above 80 ppb-8hr for any study group in any study area, while ≤ 2.2% of persons were estimated to experience two or more exposures at or above 70 ppb-8hr, each considering any adjusted air quality scenario.
- **Targeted Data Evaluations:** Afternoon time spent outdoors, along with ambient O_3 concentrations are the most influential factors when considering those persons highest exposed. There is no apparent temporal trend in the amount of outdoor time or participation rate when comparing historical CHAD diaries (1980s studies) to recently collected diary data (2000s studies); regardless, majority of CHAD data are from studies conducted since 2000. Use of activity pattern data from non-asthmatics to represent asthmatics appears reasonably justified based on an evaluation indicating their having similar outdoor time expenditure and attaining similar activity levels. APEX estimated daily exposures are somewhat comparable to personal exposure measurements; however, both over- and under-estimations occurred to varying degrees (Figure 5-15; Figure 5-16).

APEX estimated ventilation rates were comparable to literature provided estimates, particularly those of school-age children (Figure 5-19).

- **Targeted Exposure Scenarios:** When considering a modeling approach that more rigidly schedules longitudinal time location activity patterns compared with the standard longitudinal approach used by APEX, a greater percent of persons experience at least one or more exposures at or above benchmark levels. For example, an APEX model simulation using only summer time (no school) CHAD diary days for non-working school-age children generated approximately 30% more persons at or above exposure benchmark levels compared with exposures estimated using our population-based modeling approach (Figure 5-10). When accounting for a fraction of the population to avert in response to a bad air quality day, approximately 1-2 percentage point fewer persons experienced exposures at or above benchmark levels compared with exposures estimated using our population based modeling approach (Figure 5-12).

Table 5-7 Mean and Maximum Percent of all School-age Children Estimated to Experience at Least One Daily Maximum 8-hr Average Exposure to O_3 at or Above Selected Health Benchmark Levels

Study Area	Adjusted Air Quality Scenario	Percent of All School-Age Children Experiencing At Least One Exposure At or Above Selected Benchmark Level[1]					
		60 ppb-8hr		70 ppb-8hr		80 ppb-8hr	
		mean	max	mean	max	mean	max
Atlanta	75	14.8	19.3	2.8	4.4	0.3	0.7
	70	7.5	10.8	0.7	1.4	0.1	0.2
	65	2.9	4.8	0.2	0.5	0	0
Baltimore	75	12.2	19.0	2.0	4.0	0.2	0.4
	70	7.1	11.8	0.7	1.2	0.1	0.1
	65	3.0	5.4	0.2	0.3	0	0
Boston	75	13.8	21.9	2.8	6.6	0.3	1.0
	70	9.0	15.7	1.2	3.2	0.1	0.2
	65	3.4	6.7	0.2	0.5	0	0
Chicago	75	13.7	24.7	3.2	7.5	0.2	0.7
	70	9.2	16.0	1.0	2.7	0	0.1
	65	4.2	8.1	0.2	0.4	0	0
Cleveland	75	10.2	18	1.4	3.7	0.1	0.2
	70	4.2	9.3	0.3	0.9	0	0
	65	1.1	3.0	0.1	0.2	0	0
Dallas	75	12.9	22.9	1.9	4.5	0.1	0.3
	70	7.5	16.0	0.6	1.5	0	0.1
	65	3.0	7.6	0.1	0.3	0	0
Denver	75	17.0	25.6	1.7	4.1	0.1	0.5
	70	10.2	18.9	0.5	1.7	0	0.1
	65	3.8	9.5	0.1	0.4	0	0
Detroit	75	14.1	19.1	2.4	4.2	0.1	0.2
	70	7.3	10.3	0.5	0.9	0	0
	65	2.9	4.6	0.1	0.2	0	0
Houston	75	11.4	17.8	2.3	5.5	0.3	0.7
	70	6.6	11.9	0.8	2.1	0	0.1
	65	2.7	5.7	0.1	0.4	0	0
Los Angeles	75	9.5	10.2	0.6	1.0	0	0.1
	70	4.4	5.0	0.1	0.2	0	0
	65	1.1	1.5	0	0	0	0
New York	75	10.9	19.0	1.6	3.7	0.1	0.3
	70	3.3	6.6	0.2	0.5	0	0
	65	0	0.1	0	0	0	0
Philadelphia	75	13.8	20.5	2.1	4.2	0.2	0.4
	70	7.1	11.8	0.6	1.5	0	0.1
	65	2.4	4.6	0.1	0.3	0	0
Sacramento	75	10.3	16.5	1.6	2.7	0.1	0.2
	70	5.8	10.0	0.4	0.9	0	0
	65	2.7	4.7	0.1	0.2	0	0
St. Louis	75	16.3	25.8	3.3	8.1	0.3	1.1
	70	10.2	16.9	1.0	2.7	0.1	0.2
	65	3.9	7.3	0.1	0.4	0	0
Washington	75	13.2	23.4	2.4	6.0	0.2	0.8
	70	6.6	12.5	0.6	1.4	0	0.1
	65	2.3	5.0	0.1	0.2	0	0

[1] The mean is the arithmetic average of the estimated percent of all school-age children exposed across 2006-2010 year air quality; max is the highest estimated percent of all school-age children exposed in a year.

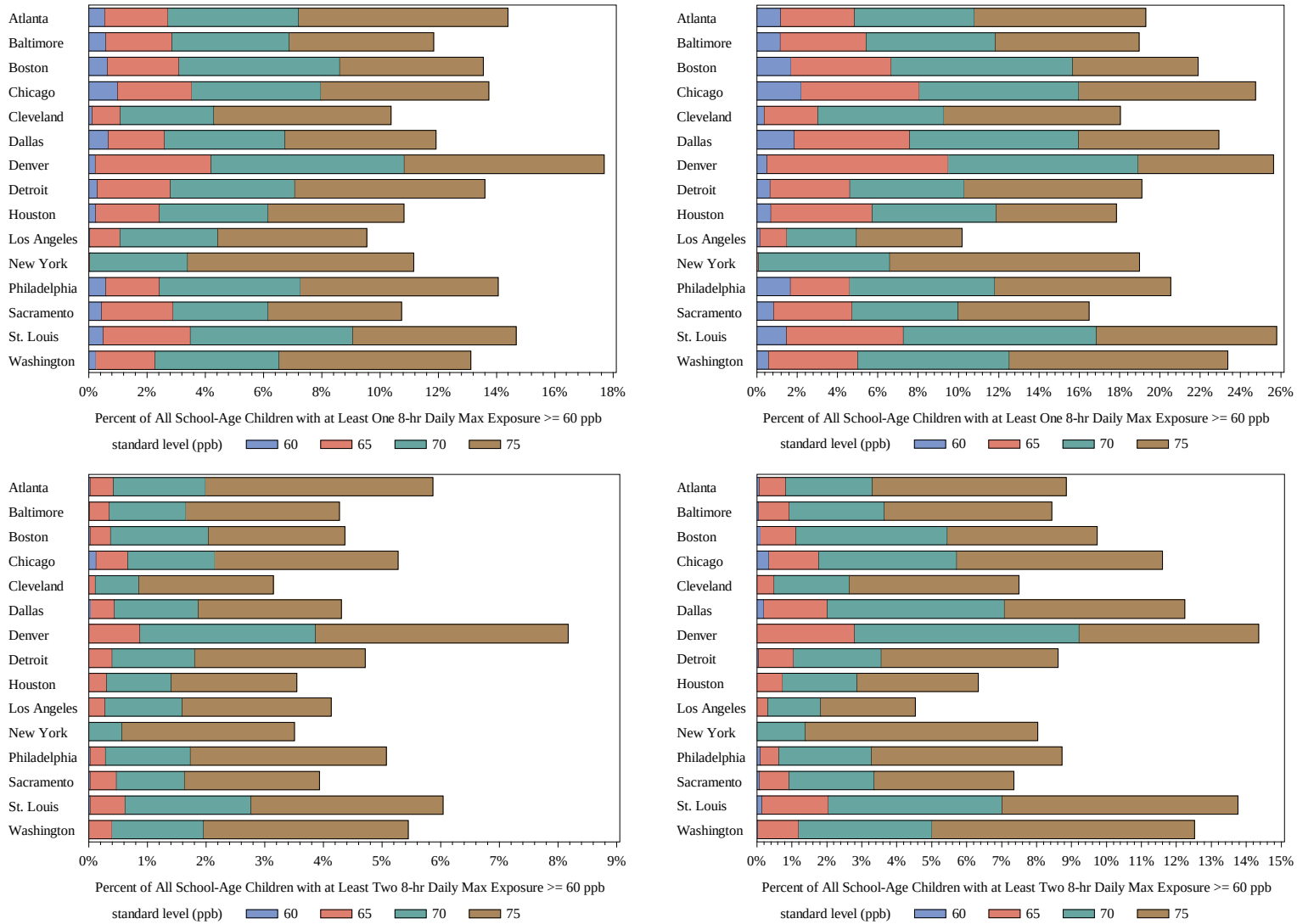


Figure 5-20 Incremental increases in percent of all school-age children exposed to O_3 at or above 60 ppb-8hr for each study area, year 2006-2010 air quality. Average percent (left panels), maximum percent (right panels), at least one exposure (top panels), at least two exposures (bottom panels) per year.

Table 5-8 Mean and Maximum Percent of All School-age Children Estimated to Experience at Least Two Daily Maximum 8-hr Average Exposures to O_3 At or Above Selected Health Benchmark Levels

Study Area	Adjusted Air Quality Scenario	Percent of All School-Age Children Experiencing At Least Two Exposures At or Above Selected Benchmark Level[1]					
		60 ppb-8hr		70 ppb-8hr		80 ppb-8hr	
		mean	max	mean	max	mean	max
Atlanta	75	6.0	8.9	0.4	0.7	0	0
	70	2.1	3.3	0	0.1	0	0
	65	0.4	0.8	0	0	0	0
Baltimore	75	4.6	8.4	0.2	0.5	0	0
	70	1.8	3.7	0	0.1	0	0
	65	0.4	0.9	0	0	0	0
Boston	75	4.5	9.7	0.3	1.1	0	0
	70	2.2	5.5	0.1	0.4	0	0
	65	0.4	1.1	0	0	0	0
Chicago	75	5.3	11.6	0.5	1.3	0	0
	70	2.5	5.7	0.1	0.2	0	0
	65	0.8	1.8	0	0	0	0
Cleveland	75	3.1	7.5	0.1	0.5	0	0
	70	0.9	2.6	0	0	0	0
	65	0.1	0.5	0	0	0	0
Dallas	75	4.8	12.2	0.2	0.8	0	0
	70	2.2	7.1	0	0.1	0	0
	65	0.5	2.0	0	0	0	0
Denver	75	7.6	14.4	0.2	0.4	0	0
	70	3.5	9.2	0	0.1	0	0
	65	0.7	2.8	0	0	0	0
Detroit	75	5.0	8.6	0.3	0.8	0	0
	70	1.9	3.6	0	0.1	0	0
	65	0.4	1.1	0	0	0	0
Houston	75	3.8	6.3	0.2	0.6	0	0
	70	1.5	2.9	0	0.1	0	0
	65	0.3	0.7	0	0	0	0
Los Angeles	75	4.1	4.5	0.1	0.1	0	0
	70	1.6	1.8	0	0	0	0
	65	0.3	0.3	0	0	0	0
New York	75	3.4	8.0	0.1	0.4	0	0
	70	0.5	1.4	0	0	0	0
	65	0	0	0	0	0	0
Philadelphia	75	5.0	8.7	0.2	0.5	0	0
	70	1.7	3.3	0	0.1	0	0
	65	0.3	0.6	0	0	0	0
Sacramento	75	3.7	7.4	0.2	0.5	0	0
	70	1.5	3.4	0	0.1	0	0
	65	0.4	0.9	0	0	0	0
St. Louis	75	7.0	13.8	0.6	2.2	0	0.1
	70	3.2	7.0	0.1	0.3	0	0
	65	0.7	2.0	0	0	0	0
Washington	75	5.5	12.5	0.4	1.4	0	0
	70	2.0	5.0	0	0.1	0	0
	65	0.4	1.2	0	0	0	0

[1] The mean is the arithmetic average of the estimated percent of all school-age children exposed across 2006-2010 year air quality; max is the highest estimated percent of all school-age children exposed in a year.

5.7 REFERENCES

Ainsworth, B. E.; W.L. Haskell; A. S. Leon; D. R. Jacobs Jr. and H. J. Montoye; J. F. Sallis and R. S. Paffenbarger Jr. 1993. "Compendium of Physical Activities: Classification of Energy Costs of Human Physical Activities." *Medicine & Science in Sports Exercise,* 25:71-80.

Arcus-Arth, A. and J. Blaisdell. 2007. "Statistical Distributions of Daily Breathing Rates for Narrow Age Groups of Infants and Children." *Risk Analysis,* 27(1):97-110.

Bresnahan, B. W.; M. Dickie and S. Gerking. 1997. "Averting Behavior and Air Pollution." *Land Economics*, 73:340-357.

Brochu, P.; J.F. Ducre-Robitaille and J. Brodeur. 2006a. "Physiological Daily Inhalation Rates for Free-living Individuals Aged 1 Month to 96 Years, Using Data from Doubly Labeled Water Measurements: A Proposal for Air Quality Criteria, Standard Calculations and Health Risk Assessment." *Human and Ecological Risk Assessment,* 12:675-701.

Brochu, P.; J. F. Ducre-Robitaille and J. Brodeur. 2006b. "Supplemental Material for Physiological Daily Inhalation Rates for Free-living Individuals Aged 1 Month to 96 Years, Using Data from Doubly Labeled Water Measurements: A Proposal for Air Quality Criteria, Standard Calculations and Health Risk Assessment." *Human and Ecological Risk Assessment*, 12:1-12.

Burmaster, D.E. 1998. "LogNormal Distributions for Skin Area as a Function of Body Weight." *Risk Analysis,* 18(1):27-32.

EPRI. 1988. *A Study of Activity Patterns Among a Group of Los Angeles Asthmatics*. Research Project 940-5. Electric Power Research Institute. Prepared by Roth Associates, November 1988.

EPRI. 1992. *A Survey of Daily Asthmatic Activity Patterns in Cincinnati*. TR-101396. Research Project 940-05. Electric Power Research Institute. Prepared by Roth Associates. November 1992.

Ford, E. S.; G. W. Heath; D. M. Mannino and S. C. Redd. 2003. "Leisure-time Physical Activity Patterns Among U.S. Adults with Asthma." *Chest,* 124:432-437.

George, B. J. and T. McCurdy. 2009. "Investigating the American Time Use Survey from an Exposure Modeling Perspective." *Journal of Exposure Science and Environmental Epidemiology*, 21:92-105.

Geyh, A. S.; J. Xue, H. Özkaynak and J. D. Spengler. 2000. "The Harvard Southern California Chronic Ozone Exposure Study: Assessing Ozone Exposure of Grade-school-age Children in Two Southern California Communities." *Environmental Health Perspectives,* 108:265-270.

Glen, G.; L. Smith; K. Isaacs; T. McCurdy and J. Langstaff. 2008. "A New Method of Longitudinal Diary Assembly for Human Exposure Modeling." *Journal of Exposure Science and Environmental Epidemiology,* 18:299-311.

Graham, S. E. and T. McCurdy. 2004. "Developing Meaningful Cohorts for Human Exposure Models." *Journal of Exposure Analysis and Environmental Epidemiology*, 14:23-43.

Graham, S. E. and T. McCurdy. 2005. *Revised Ventilation Rate (VE) Equations for Use in Inhalation-oriented Exposure Models*. Washington, DC: EPA Office of Air and Radiation. (EPA document number EPA/600/X-05/008, Appendix A). Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=202543>.

Graham, S. 2009. *Response to Peer-review Comments on Appendix A,* prepared by S. Graham (U.S. EPA). Response is Appendix D of US EPA (2009b). Research Triangle Park, NC: EPA Office of Air and Radiation. Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=202543>.

Graham, S. E. 2012. *Comprehensive Review of Published Averting Behavior Studies and Available Technical Documents.* Memo to Bryan Hubbell, Group Leader, Risk and Benefits Group, Office of Air Quality Planning and Standards. (EPA Docket # EPA-HQ-OAR-2008-0699-0085).

Isaacs, K. and L. Smith. 2005. *New Values for Physiological Parameters for the Exposure Model Input File Physiology*, text memorandum submitted to the U.S. Environmental Protection Agency under EPA Contract EP-D-05-065. NERL WA 10. Alion Science and Technology. Found in US EPA . (2009). Risk and Exposure Assessment to Support the Review of the SO_2 Primary National Ambient Air Quality Standard. (EPA document number EPA-452/R-09-007, August 2009). Available at: <http://www.epa.gov/ttn/naaqs/standards/so2/data/200908SO2REAFinalReport.pdf>.

Isaacs, K.; G. Glen; T. McCurdy and L. Smith. 2008. "Modeling Energy Expenditure and Oxygen Consumption in Human Exposure Models: Accounting for Fatigue and EPOC." *Journal of Exposure Science and Environmental Epidemiology*, 18:289–298.

Kansas Department of Health and Environment. 2006. *Environmental Factors, Outdoor Air Quality, and Activity Level*: *Results from 2005 Kansas Behavioral Risk Factor Surveillance System*. Office of Health Promotion, KDOH, available at: <http://www.kdheks.gov/brfss/PDF/cste_report_final.pdf>.

Langstaff, J. E. 2007. *Analysis of Uncertainty in Ozone Population Exposure Modeling*, OAQPS Staff Memorandum to Ozone NAAQS Review, January 31. Washington, DC: Office of Air Radiation. (EPA docket number OAR-2005-0172). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_cr_td.html>.

Lioy, P.J. 1990. "The Analysis of Total Human Exposure for Exposure Assessment: A Multi-discipline Science for Examining Human Contact with Contaminants." *Environmental Science and Technology*, 24: 938-945.

Mansfield, C.; F. R. Johnson and G. Van Houtven. 2006. "The Missing Piece: Averting Behavior for Children's Ozone Exposures." *Resource Energy Economics*, 28:215-228.

Marino, A. J.; E. N. Fletcher; R. C. Whitaker and S. E. Anderson. 2012. "Amount and Environmental Predictors of Outdoor Playtime at Home and School: Across-sectional

Analysis of a National Sample of Preschool-aged Children Attending Head Start." *Health & Place,* 18: 1224-1230.

McCurdy, T. 2000. "Conceptual Basis for Multi-route Intake Dose Modeling Using an Energy Expenditure Approach." *Journal of Exposure Analysis and Environmental Epidemiology,* 10:1-12.

McCurdy, T.; G. Glen; L. Smith and Y. Lakkadi. 2000. "The National Exposure Research Laboratory's Consolidated Human Activity Database." *Journal of Exposure Analysis and Environmental Epidemiology,* 10:566-578.

McDermott, M.; R. Srivastava and S. Croskell. 2006. "Awareness of and Compliance with Air Pollution Advisories: A Comparison of Parents of Asthmatics with Other Parents." *Journal of Asthma,* 43:235-239.

McDonnell, W. F.; H. R. Kehrl; S. Abdul-Salaam; P. J. Ives; L. J. Folinsbee; R. B. Devlin; J. J. O'Neil and D. H. Horstman. 1991. "Respiratory Response of Humans Exposed to Low Levels of Ozone for 6.6 Hours." *Archives of Environmental Health,* 46(3):145-150.

Meng, Q.; R. Williams and J. P. Pinto. (2012). "Determinants of the Associations Between Ambient Concentrations and Personal Exposures to $PM_{2.5}$, NO_2, and O_3 During DEARS." *Atmospheric Environment,* 63:109-116.

Montoye, H. J.; H. C. G. Kemper; W. H. N. Saris and R. A. Washburn. 1996. *Measuring Physical Activity and Energy Expenditure*. Human Kinetics: Champaign, IL.

National Research Council. 1991. *Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities*. Washington, DC: National Academy of Sciences.

Neidell, M. 2010. "Air Quality Warnings and Outdoor Activities: Evidence from Southern California Using a Regression Discontinuity Approach Design." *Journal of Epidemiology Community Health,* 64:921-926.

Santuz, P.; E. Baraldi; M. Filippone and F. Zacchello. 1997. "Exercise Performance in Children with Asthma: Is it Different from that of Healthy Controls?" *European Respiratory Journal,* 10:1254-1260.

Schofield, W. N. 1985. "Predicting Basal Metabolic Rate, New Standards, and Review of Previous Work." *Human Nutrition - Clinical Nutrition,* 39C(S1):5-41.

Sexton, A. L. (2011). "Responses to Air Quality Alerts: Do Americans Spend Less Time Outdoors?" Available at: <http://www.apec.umn.edu/prod/groups/cfans/@pub/@cfans/@apec/documents/asset/cfans_asset_365645.pdf>.

Shamoo, D. A.; W. S. Linn; R. C. Peng; J. C. Solomon; T. L. Webb; J. D. Hackney and H. Hong. 1994. "Time-activity Patterns and Diurnal Variation of Respiratory Status in a Panel of Asthmatics: Implications for Short-term Air Pollution Effects." *Journal of Exposure Analysis and Environmental Epidemiology,* 4(2):133-148.

U.S. Bureau of Labor Statistics. 2012. *American Time Use Survey User's Guide*. "Understanding ATUS 2003 to 2011," August 2012. Data and documentation available at: <http://www.bls.gov/tus/>.

U.S. Census Bureau. 2007. "2000 Census of Population and Housing. Summary File 3 (SF3) Technical Documentation," available at: <http://www.census.gov/prod/cen2000/doc/sf3.pdf>. Individual SF3 files '30' (for income/poverty variables pct50) for each state were downloaded from: <http://www2.census.gov/census_2000/datasets/Summary_File_3/>.

U.S. Department of Agriculture. 2000. "Continuing Survey of Food Intake by Individuals (CSFII) 1994-96," 1998. Agricultural Research Service, CD-ROM, available at: <http://www.ars.usda.gov/Services/docs.htm?docid=14531>.

U.S. Environmental Protection Agency. 1986. *Air Quality Criteria for Ozone and Other Photochemical Oxidants*. Research Triangle Park, NC: Office of Health and Environmental Assessment, Environmental Criteria and Assessment Office. (EPA document number EPA-600/8-84-020aF-eF). Available from the National Technical Information Service, Springfield, VA. (NTIS publication number PB87-142949).

U.S. EPA. 1996a. *Review of National Ambient Air Quality Standards for Ozone: Assessment of Scientific and Technical Information - OAQPS Staff Paper*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA/452/R-96-007). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_pr_sp.html>.

U.S. EPA. 1996b. *Air Quality Criteria for Ozone and Related Photochemical Oxidants*. Research Triangle Park, NC: EPA Office of Research and Development, National Center for Environmental Assessment. (EPA document number EPA/600/P-93/004aF-cF). Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=2831>.

U.S. EPA. 2002. *Consolidated Human Activities Database (CHAD) Users Guide*. Database and documentation available at: <http://www.epa.gov/chadnet1/>.

U.S. EPA. 2005. *Guidance on Selecting Age Groups for Monitoring and Assessing Childhood Exposures to Environmental Contaminants*. (EPA document number EPA/630/P-03/003F). Available at: <http://www.epa.gov/raf/publications/pdfs/AGEGROUPS.PDF>.

U.S. EPA. 2006. *Air Quality Criteria for Ozone and Related Photochemical Oxidants (Final)*. Research Triangle Park, NC: EPA National Center for Environmental Assessment. (EPA document number EPA/600/R-05/004aF-cF). Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=149923>.

U.S. EPA. 2007a. *Review of National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information - OAQPS Staff Paper*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-07-007). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/data/2007_07_ozone_staff_paper.pdf>.

U.S. EPA. 2007b. *Ozone Population Exposure Analysis for Selected Urban Areas*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_O_3_cr_td.html>.

U.S. EPA. 2008. *Risk and Exposure Assessment to Support the Review of the NO_2 Primary National Ambient Air Quality Standard*. Washington, DC: EPA Office of Air and Radiation. (EPA document number EPA-452/R-08-008a0, November). Available at: <http://www.epa.gov/ttn/naaqs/standards/nox/data/20081121_NO2_REA_final.pdf>.

U.S. EPA. 2009a. *Risk and Exposure Assessment to Support the Review of the SO_2 Primary National Ambient Air Quality Standard*. (EPA document number EPA-452/R-09-007, August). Available at: <http://www.epa.gov/ttn/naaqs/standards/so2/data/200908SO2REAFinalReport.pdf>.

U.S. EPA. 2009b. *Metabolically Derived Human Ventilation Rates: A Revised Approach Based Upon Oxygen Consumption Rates*. Washington, DC: EPA National Center for Environmental Assessment. (EPA document number EPA/600/R-06/129F). Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=202543>.

U.S. EPA. 2010. *Quantitative Risk and Exposure Assessment for Carbon Monoxide – Amended*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-10-009, July). Available at: <http://www.epa.gov/ttn/naaqs/standards/co/data/CO-REA-Amended-July2010.pdf>.

U.S. EPA. 2011. *Exposure Factors Handbook*, 2011 edition. Washington, DC: EPA National Center for Environmental Assessment. (EPA document number EPA/600/R-09/052F). Available at: <http://www.epa.gov/ncea/efh/pdfs/efh-complete.pdf>.

U.S. EPA. 2012a. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume I: User's Guide*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001a). Available at: <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. (2012b. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume II: Technical Support Document*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/B-12-001b). Available at: <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2013. *Integrated Science Assessment of Ozone and Related Photochemical Oxidants*. Research Triangle Park, NC: EPA National Center for Environmental Assessment. (EPA document number EPA 600/R-10/076F). Available at: <http://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=247492#Download>.

van Gent, R.; K. van der Ent; L. E. M. van Essen-Zandvliet; M. M. Rovers; J. L. L. Kimpen; G. de Meer and P. H. C. Klijn. 2007. "No Difference in Physical Activity in (Un)diagnosed Asthma and Healthy Controls." *Pediatric Pulmonology*, 42:1018-1023.

Wen, X. J.; L. Balluz and A. Mokdad. 2009. "Association Between Media Alerts of Air Quality Index and Change of Outdoor Activity Among Adult Asthma in Six States," BRFSS, 2005. *Journal of Community Health*, 34:40-46.

Whitfield, R.; W. Biller; M. Jusko and J. Keisler. 1996. *A Probabilistic Assessment of Health Risks Associated with Short- and Long-Term Exposure to Tropospheric Ozone*. Argonne, IL: Argonne National Laboratory.

World Health Organization. 2008. *Uncertainty and Data Quality in Exposure Assessment."*Part 1: Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment." Available at: <http://www.who.int/ipcs/publications/methods/harmonization/exposure_assessment.pdf.

Xue, J.; T. McCurdy; J. Spengler and H. Özkaynak. 2004. "Understanding Variability in Time Spent in Selected Locations for 7-12 Year Old Children." *Journal of Exposure Analysis and Environmental Epidemiology*,14(3):222-33.

Zivin, J. G. and M. Neidell. 2009. "Days of Haze: Environmental Information Disclosure and Intertemporal Avoidance Behavior." *Journal of Environmental Economic Management*, 58(2):119-128.

6 CHARACTERIZATION OF HEALTH RISKS BASED ON CONTROLLED HUMAN EXPOSURE STUDIES

6.1 INTRODUCTION

This chapter presents information regarding the methods and results for a controlled human exposure-based O_3 (O_3) health risk assessment that builds upon the methodology used in the assessment conducted as part of the O_3 NAAQS review completed in 2008 and also introduces a new method for estimating risk. In the previous review, EPA conducted a health risk assessment that produced risk estimates for the number and percent of school-aged children, asthmatic school-aged children, and the general population experiencing lung function decrements associated with O_3 exposures for 12 urban areas, where lung function is measured as forced expiratory volume in one second (FEV_1). That portion of the risk assessment was based on exposure-response relationships developed from analysis of data from several controlled human exposure studies which were combined with population-level exposure distributions developed for children and adults. Risk estimates for lung function decrements were developed for recent air quality levels and for just meeting the existing 8-hour standard and several alternative 8-hour standards. The methodological approach followed in the last risk assessment and risk estimates resulting from that assessment are described in the 2007 Staff Paper (U.S. EPA, 2007a).

The goals of the current O_3 risk assessment are to provide estimates of the number and percentage of persons that would experience adverse respiratory effects associated with recent O_3 levels and with meeting the existing and potential alternative O_3 standards in specific urban areas; and to develop a better understanding of the influence of various inputs and assumptions on the risk estimates. The current assessment includes estimates of risks of lung function decrements in school-aged children (ages 5 to 18), asthmatic school-aged children, and the adult population (19 and above). We recognize that there are many sources of uncertainty in the inputs and approach used in this portion of the health risk assessment which make the specific estimates uncertain, however, we have sufficient confidence in the magnitude and direction of the estimates provided by the assessment for it to serve as a useful input to decisions on the adequacy of the O_3 standard and risk reductions associated with alternative standards.

We are estimating lung function risk using two methodologies in this review. The primary results are based on a new model that estimates FEV_1 responses for individuals associated with short-term exposures to O_3 (McDonnell, Stewart, and Smith, 2012). We refer to this model as the McDonnell-Stewart-Smith (MSS) model. We also provide estimates following the methodology used in previous reviews which provides population level estimates of the

percent and number of people at risk. We refer to this model as the Exposure-Response (E-R) model used in previous reviews. Both of these models are implemented in the air pollution exposure model APEX (EPA, 2012b,c). Following this introductory section, this chapter discusses the scope of the controlled human exposure study based risk assessment, describes the risk models, and provides key results from the assessment. The results of sensitivity analyses are reported and key uncertainties are identified and summarized. More detailed descriptions of several parts of the analyses are included in appendices that accompany the REA.

6.1.1 Development of Approach for Current Risk Assessment

The lung function risk assessment described in this chapter builds upon the methodology and lessons learned from the risk assessment work conducted for previous reviews (EPA, 1996, 2007a). The current risk assessment also is based on the information evaluated in the ISA (EPA, 2013a). The general approach used in the current risk assessment was described in the Scope and Methods Plan for Health Risk and Exposure (EPA, 2011), that was released to the CASAC and general public in April 2011 for review and comment and which was the subject of a consultation with the CASAC O_3 Panel in May 2011. The first draft REA was reviewed by CASAC in September 2012. The approach used in the current risk assessment reflects consideration of the comments offered by CASAC members and the public on the Scope and Methods Plan and the first draft REA.

Controlled human exposure studies involve volunteers who are exposed while engaged in different exercise regimens to specified levels of O_3 under controlled conditions for specified amounts of time. For the current health risk assessment, we are using probabilistic exposure-response relationships based on analysis of individual data that describe the relationship between measures of personal exposure to O_3 and measures of lung function recorded in the studies. Therefore, a risk assessment based on exposure-response relationships derived from controlled human exposure study data requires estimates of personal exposure to ambient O_3. Because data on personal hourly exposures to O_3 of ambient origin are not available, estimates of personal exposures to varying ambient concentrations are derived through exposure modeling, as described in Chapter 5. While the quantitative risk assessment based on controlled human exposure studies addresses only lung function responses, it is important to note that other respiratory responses have been found to be related to O_3 exposures in these types of studies, including increased lung inflammation, increased respiratory symptoms, increased airway responsiveness, and impaired host defenses. Sufficient information is not available to quantitatively model these other endpoints. Section 6.2 of the ISA provides a discussion of these additional health endpoints which are an important part of the overall characterization of risks associated with ambient O_3 exposures.

6.1.2 Comparison of Controlled Human Exposure- and Epidemiologic-based Risk Assessments

In contrast to the **exposure-response** relationships derived from controlled human exposure studies, epidemiological studies provide estimated **concentration-response** relationships based on data collected in real world community settings. The assessment of health risk based on epidemiological studies is the subject of Chapter 7. The characteristics that are relevant to carrying out a risk assessment based on controlled human exposure studies versus one based on epidemiology studies can be summarized as follows:

- The relevant controlled human exposure studies in the ISA provide data that can be used to estimate exposure-response functions, and therefore a risk assessment based on these studies requires as input (modeled) personal exposures to ambient O_3. The relevant epidemiological studies in the ISA provide concentration-response functions, and, therefore, a risk assessment based on these studies requires as input (actual monitored or adjusted based on monitored) ambient O_3 concentrations, and personal exposures are not required as inputs to the assessment.
- Epidemiological studies are carried out in specific real world locations (e.g., specific urban areas). To minimize extrapolation uncertainty, a risk assessment based on epidemiological studies is best performed in locations where the studies took place. Controlled human exposure studies, carried out in laboratory settings, are generally not specific to any particular real world location. A risk assessment based on controlled human exposure studies can therefore appropriately be carried out for any location for which there are adequate air quality and other data on which to base the modeling of personal exposures.
- To derive estimates of risk from concentration-response relationships estimated in epidemiological studies, it is usually necessary to have estimates of the baseline incidences of the health effects involved. Such baseline incidence estimates are not needed in a controlled human exposure studies-based risk assessment.

6.2 SCOPE OF LUNG FUNCTION HEALTH RISK ASSESSMENT

The current controlled human exposure-based O_3 health risk assessment is one approach used to estimate risks associated with exposure to ambient O_3 in a number of urban areas selected to illustrate the public health impacts of this pollutant. The short-term exposure related health endpoints selected for this portion of the O_3 health risk assessment include those for which the ISA concludes that the evidence as a whole supports the general conclusion that O_3, acting alone and/or in combination with other components in the ambient air pollution mix is causal or likely to be causally related to the endpoint.

In the 2007 O_3 NAAQS review, the controlled human exposure-based health risk assessment involved developing risk estimates for lung function decrements (≥ 10, ≥ 15, and $\geq 20\%$ changes in FEV_1) in school-aged children (ages 5 to 18 years old). The strong emphasis

on children reflects the finding of previous O_3 NAAQS reviews that children are an important at-risk group. Due to the increased amount of time spent outdoors engaged in relatively high levels of physical activity (which increases intake), school-aged children as a group are particularly at risk for experiencing O_3-related health effects.

Outdoor workers and other adults who engage in moderate exertion for prolonged periods or heavy exertion for shorter periods during the day also are clearly at risk for experiencing similar lung function responses when exposed to elevated ambient O_3 concentrations. In this second draft REA, we focus the quantitative risk assessment for lung function decrements on all and asthmatic school-aged children (ages 5-18), and the adult population (ages 19 and above).

For the second draft assessment, lung function risks are estimated for 15 cities, Atlanta, Baltimore, Boston, Chicago, Cleveland, Dallas, Denver, Detroit, Houston, Los Angeles, New York, Philadelphia, Sacramento, St. Louis, and Washington, DC.

6.2.1 Selection of Health Endpoints

The ISA identifies several responses to short-term O_3 exposure that have been evaluated in controlled human exposure studies (US EPA, 2013, sections 6.2.1.1, 6.2.2.1, 6.2.3.1, and 6.3.1). These include decreased inspiratory capacity; decreased forced vital capacity (FVC) and forced expiratory volume in one second (FEV_1); mild bronchoconstriction; rapid, shallow breathing patterns during exercise; symptoms of cough and pain on deep inspiration (PDI); increased airway responsiveness; and pulmonary inflammation. Such studies provide direct evidence of relationships between short-term O_3 exposure and an array of respiratory-related effects, however, there are only sufficient exposure-response data at different concentrations to develop quantitative risk estimates for O_3-related decrements in FEV_1. Other responses to O_3 which may be equally or more important then FEV_1 decrements (e.g., inflammation) do not necessarily correlate with FEV_1 responses (ISA, section 6.2.3.1) and this risk assessment is not able to address these other responses.

As stated in the 2006 Criteria Document (Table 8-3, p.8-68) for adults with lung disease, even moderate functional responses (e.g., FEV_1 decrements $\geq$ 10% but < 20%) would likely interfere with normal activities for many individuals, and would likely result in more frequent medication use. In a recent letter to the Administrator, the CASAC O_3 Panel stated that "'Clinically relevant' effects are decrements > 10%, a decrease in lung function considered clinically relevant by the American Thoracic Society" (Samet, 2011, p.2). The CASAC O_3 Panel also stated that:

> a 10% decrement in FEV_1 can lead to respiratory symptoms, especially in individuals with pre-existing pulmonary or cardiac disease. For example, people with chronic obstructive pulmonary disease have decreased ventilatory

> reserve (i.e., decreased baseline FEV_1) such that a ≥ 10% decrement could lead to moderate to severe respiratory symptoms (Samet, 2011, p.7).

This is consistent with the most recent official statement of the American Thoracic Society on what constitutes an adverse lung function health effect of air pollution:

> The committee recommends that a small, transient loss of lung function, by itself, should not automatically be designated as adverse. In drawing the distinction between adverse and nonadverse reversible effects, this committee recommended that reversible loss of lung function in combination with the presence of symptoms should be considered adverse (ATS, 2000, p.672).

For this lung function risk assessment, a focus on the mid- to upper-end of the range of moderate levels of functional responses and higher (FEV_1 decrements ≥ 15%) is appropriate for estimating potentially adverse lung function decrements in active healthy adults, while for people with asthma or lung disease, a focus on moderate functional responses (FEV_1 decrements down to 10%) may be appropriate.

6.2.2 Approach for Estimating Health Risk Based on Controlled Human Exposure Studies

The major components of the health risk assessment based on data from controlled human exposure studies are illustrated in Figure 3-3 in Chapter 3. As shown in this figure, under this portion of the risk assessment, exposure estimates for a number of different air quality scenarios (i.e., recent year of air quality, just meeting the existing 8-hour and alternative standards) are combined with probabilistic exposure-response relationships derived from the controlled human exposure studies to develop risk estimates associated with recent air quality and after simulating just meeting the existing and alternative standards. The health effect included in this portion of the risk assessment is lung function decrement, as measured by changes in FEV_1. The population risk estimates for a given lung function decrement (e.g., ≥ 15% reduction in FEV_1) are estimates of the expected number of people who will experience that lung function decrement, the number of times that people experience repeated occurrences of given lung function decrements, and the number of occurrences (person-days) of the given lung function decrement. The air quality and exposure analysis components that are integral to this portion of the risk assessment are discussed in Chapters 4 and 5.

We used two approaches to estimate health risk. As done for the risk assessment conducted during the previous O_3 NAAQS review, a Bayesian Markov Chain Monte Carlo approach was used to develop probabilistic exposure-response functions. These functions were then applied to the APEX estimated population distribution of 8-hour maximum exposures for persons at or above moderate exertion (≥ 13 L/min-m^2 body surface area) to estimate the number

of persons expected to experience lung function decrements. The primary approach, based on the McDonnell-Stewart-Smith FEV_1 model, uses the time-series of O_3 exposure and corresponding ventilation rates for each APEX simulated individual to estimate their personal time-series of FEV_1 reductions, selecting the daily maximum reduction for each person. A key difference between these approaches is that the previous method estimates a population distribution of FEV_1 reductions, where the MSS model estimates FEV_1 reductions at the individual level. Each of these approaches is discussed in detail below.

6.2.3 Controlled Human Exposure Studies

Modeling of risks of lung function decrements as a function of exposures to O_3 is based on application of results from controlled human exposure studies. As discussed in Chapter 6 of the ISA (EPA, 2013a), there is a significant body of controlled human exposure studies reporting lung function decrements and respiratory symptoms in adults associated with 1- to 8-hour exposures to O_3. In the ISA sections on controlled human exposure (Sections 6.2.1.1, 6.2.2.1, 6.2.3.1, and 6.3.1) over 140 references to human clinical studies are reported.

6.2.3.1 Life Stages

Consistent with the approach used in the previous O_3 NAAQS review and lacking a significant body of controlled human exposure studies on children, we judge that it is reasonable to estimate exposure-response relationships for lung function decrements associated with O_3 exposures in children 5-18 years old based on data from young adult subjects (18-35 years old). As discussed in the ISA (EPA, 2013a), findings from clinical studies for children and summer camp field studies of children 7-17 years old in at least six different locations in the U.S. and Canada found lung function decrements in healthy children similar to those observed in healthy young adults exposed to O_3 under controlled chamber conditions. There are fewer studies of young children than adolescents to draw upon, which may add to uncertainties in the modeling. Additional uncertainties are likely introduced since the lungs and airways of children are developing, while development is complete in adults (Dietert et al., 2000). The primary period of alveolar development is from birth to around eight years of age, but there is evidence for continued development through adolescence. The adult number of alveoli is reached by 2–3 years of age and the size and surface area of the alveoli increase until after adolescence (Hislop, 2002; Narayanan et al., 2012).

Lung function responses to O_3 exposure for adults older than 18 decrease with age until around age 55, when responses are minimal. “Children, adolescents, and young adults appear, on average, to have nearly equivalent spirometric responses to O_3, but have greater responses than middle-aged and older adults when similarly exposed to O_3” (ISA p. 6-21). “In healthy individuals, the fastest rate of decline in O_3 responsiveness appears between the ages of 18 and

35 years (Passannante et al., 1998; Seal et al., 1996), more so for females then males (Hazucha et al., 2003). During the middle age period (35-55 years), O_3 sensitivity continues to decline, but at a much lower rate. Beyond this age (>55 years), acute O_3 exposure elicits minimal spirometric changes" (ISA p. 6-23).

6.2.3.2 **Asthma**

There have been several controlled human exposure studies of the effects of O_3 on asthmatic subjects, going back to 1978 (Linn et al., 1978). In reference to these studies, the ISA states that "[b]ased on studies reviewed in the 1996 and 2006 O_3 AQCDs, asthmatic subjects appear to be at least as sensitive to acute effects of O_3 as healthy nonasthmatic subjects" (ISA p. 6-20). Studies published since the 2006 O_3 AQCD do not alter this conclusion (ISA, p. 6-20 to 6-21). In the 2010 O_3 NAAQS proposal (75 FR 2969-2972), EPA describes the evidence that people with asthma are as sensitive as, if not more sensitive than, normal subjects in manifesting O_3-induced pulmonary function decrements.

In reference to epidemiologic studies, the ISA states that "[t]he evidence supporting associations between short-term increases in ambient O_3 concentration and increases in respiratory symptoms in children with asthma is derived mostly from examination of 1-h max, 8-h max, or 8-h avg O_3 concentrations and a large body of single-region or single-city studies. The few available U.S. multicity studies produced less consistent associations." (ISA, p. 6-101 to 6-102). "Although recent studies contributed mixed evidence, the collective body of evidence supports associations between increases in ambient O_3 concentration and increased asthma medication use in children" (ISA, p. 6-109).

6.2.3.3 **Ethnicity**

There are two controlled human exposure studies that have assessed differences in lung function responses comparing ethnic groups (ISA, p. 6-23 to 6-24). Both of these studies show greater FEV_1 decrements in blacks than whites, however, epidemiologic studies were less supportive of this difference in response. The data available are insufficient to quantify any differences that might exist due to the limited number of studies and a lack of consistency between disciplines.

6.2.3.4 **Body Mass Index**

Some studies have found greater FEV_1 decrements to be associated with increasing BMI. BMI was included in some of the models of McDonnell et al. (2012); however, the BMI terms were found to be statistically insignificant, indicating that the effect of BMI on FEV_1 in the presence of O_3 is likely to be small, within the range of BMIs of the subjects studied.

6.2.3.5 **Outdoor Workers**

Although there are no controlled human exposure studies that have had specifically outdoor workers as subjects, the studies are applicable to outdoor workers: the 6.6-hour experimental protocol was intended to simulate the performance of heavy physical labor for a full workday (ISA, p. 6-9).

6.2.3.6 **Variability of Responses**

Responses to O_3 exposure are variable within the population, even within cohorts of similar people (e.g., healthy young adult white males) (ISA, p. 6-16 to 6-20). Factors which contribute to interindividual variability include health status, body mass index, age, sex, race/ethnicity, and the intrinsic responsiveness of individuals. Other factors which contribute to the variability of responses include the duration and concentration of O_3 exposure, the level of exercise and breathing rate, attenuation due to repeated exposures, and co-exposures with other pollutants. For specific individuals, lung function responses tend to be reproducible over a period of several months.

6.2.4 The McDonnell-Stewart-Smith (MSS) Model

In this review, EPA is investigating the use of a new model that estimates FEV_1 responses for individuals associated with short-term exposures to O_3 (McDonnell, Stewart, and Smith, 2007; McDonnell, Stewart, and Smith, 2010). This is a fundamentally different approach than the previous approach, for which the exposure-response function is at a population level, not an individual level. This model was developed using the controlled human exposure data described in Section 6.2.5 as well as incorporating several additional data sets from studies using shorter exposure durations and different exertion levels and breathing rates. These data were from 15 controlled human O_3 exposure studies that included exposure of 541 volunteers (ages 18 to 35[1]) on a total of 864 occasions. These data are described in McDonnell et al. (1997). Schelegle et al. (2009) found that there appears to be a delay in response when modeling FEV_1 decrements as a function of cumulative dose and estimated a threshold associated with the delay. McDonnell et al. (2012) refit their 2010 model using data from eight additional studies with 201 subjects and incorporating a threshold parameter into the model. Their threshold parameter allows for modeling a delay in response until cumulative dose rate (taking into account decreases over time according to first order reaction kinetics) reaches a threshold value and is found by McDonnell et al. (2012) to slightly improve model fit. That latest model is the model described

[1] The ages in these studies range from 18 years 1 month to 35 years 1 month.

here and is the model used in this risk assessment. The threshold is not a concentration threshold and does not preclude responses at low concentration exposures.

Schelegle et al. (2012) have also developed a 2-compartment model for predicting FEV_1 decrements (ISA, p. 6-15,16). Their model is similar to the MSS model in that it accounts for the effects of cumulative dose coupled with an exponential decay and also has a threshold, below which response is delayed. The primary difference between this model and the MSS model is that in the Schelegle et al. model the net cumulative dose is multiplied by an individual's responsiveness coefficient to obtain a predicted FEV_1 decrement, whereas in the MSS model the FEV_1 decrement increases as a sigmoid-shaped function of the net cumulative dose rate. Also, the Schelegle et al. model's threshold is based on cumulative intake dose (the integral of concentration x volume inhaled), where the MSS model's threshold is based on net cumulative dose rate (taking into account the first order decay). A direct comparison of the results of these two models has not been performed.

The MSS model is conceptually a two-compartment model. The cumulative amount of exposure to O_3 (exposure concentration times ventilation rate, loosely speaking a measure of dose rate) is modeled in the first compartment and modified by an exponential decay factor to yield an intermediate quantity **X**. The response (lung function decrement) of the individual to **X** is modeled in the second compartment (Figure 6-1). The threshold parameter imposes the constraint that there is no response while the value of **X** is below the threshold value.

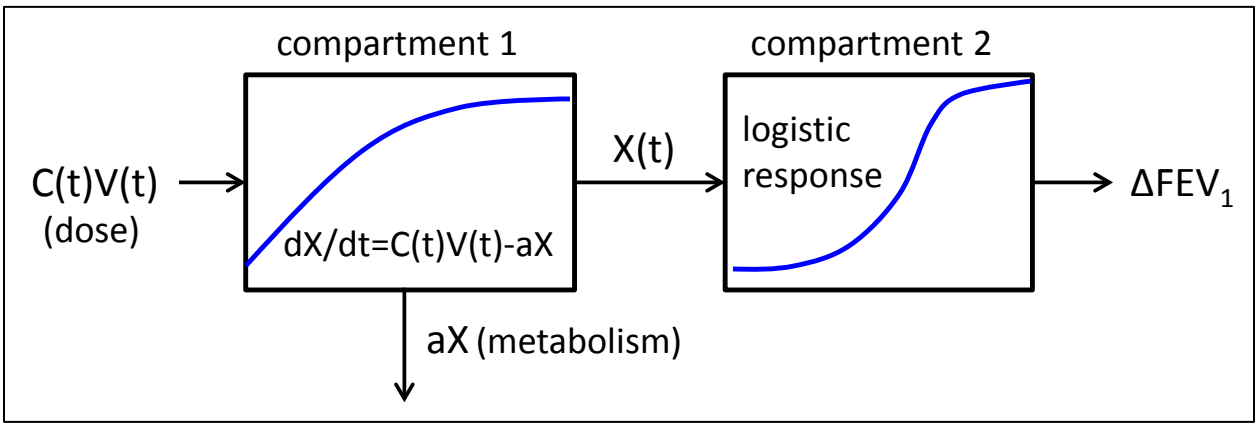


Figure 6-1. Two-Compartment Model

C is exposure concentration, **V** is ventilation rate, **t** is time, **X** is an intermediate quantity, **a** is a decay constant. Adapted from Figure 1 in McDonnell et al. (1999).

X is given by the solution of the differential equation (6-1):

$$\frac{dX}{dt} = C(t)V(t)^{\beta 6} - \beta_5 X(t) \qquad \text{(Equation 6-1)}$$

X(*t*) increases with "dose" ($C \cdot V^{\beta 6}$) over time for an individual and allows for removal of O_3 with a half-life of $1/\beta_5$ through the 2nd term in equation (6-1). In APEX, because the exposure concentration, exertion level, and ventilation rate are constant over an event, this equation has an analytic solution for each event ("events" in APEX are intervals of constant activity and concentration, where an individual is in one microenvironment, and range in duration from one to 60 minutes):

$$X(t) = X(t_o)\, e^{-\beta 5(t-t0)} + \frac{C(t)}{\beta_5} V(t)^{\beta 6}(1 - e^{-\beta 5(t-t0)}) \qquad \text{(Equation 6-2)}$$

This model calculates the FEV_1 decrement due to O_3 exposure (compartment 2) as:

$$\%\Delta FEV1_{ijk} = e^{Ui}[\beta_1 + \beta_2(Age_{ik} - \bar{A})]\left\{\frac{1}{1 + \beta_4\, e^{-\beta 3\, Tijk}} - \frac{1}{1 + \beta_4}\right\} + \varepsilon_{ijk} \qquad \text{(Equation 6-3)}$$

where $T_{ijk} = \max\{0, \mathbf{X}_{ijk} - \beta_9\}$. β_9 is a threshold parameter which allows **X** to increase up to the threshold before the median response is allowed to exceed zero.

The variables in the above equations are defined as:

The indices *i,j,k* refer to the *i*th subject at the *j*th time for the *k*th experiment for that subject,

C(t) is the O_3 exposure concentration at time *t* (ppm) during the event,

V(t) = *VE(t)*/*BSA* is the ventilation rate normalized by body surface area at time *t* (L/min-m^2),

VE(t) is the expired minute volume at time *t* (L min^{-1}),

BSA is the body surface area (m^2),

t is the time (minutes), t_0 is the time at the start of the event,

Age_{ik} is age in years of the *i*th subject in the *k*th study,

$\bar{A}$ is an age parameter (taken to be the approximate mean age of the clinical study subjects in the McDonnell, Stewart, and Smith 2007 ($\bar{A}$=25), 2010 ($\bar{A}$=25), and 2012 ($\bar{A}$=23.8) papers),

U_i is a subject-level random effect (between-individual variability not otherwise captured by the model), and

ε_{ijk} is a variability term, which includes measurement error and intra-individual variability not otherwise captured by the model.

The βs and the variances of the $\{U_i\}$ and $\{\varepsilon_{ijk}\}$ are fitted model parameters (see McDonnell, et al. (2007, 2010, and 2012) for details). In APEX, values of U_i and ε_{ijk} are drawn from Gaussian distributions with mean zero and variances var(U) and var(ε), constrained to be within ±2 standard deviations from the means. The values of U_i are chosen once for each individual and remain constant for individuals throughout the simulation. The ε_{ijk} are sampled daily for each individual. The best fit values (based on maximum likelihood) for these parameters are listed in Table 6-1. The values in parentheses are standard errors of the estimates (given here to two significant digits; the values in the papers are given to up to five significant digits). Although some of the parameters are quite different in the three models in Table 6-1, the predictions of these three models are similar. The relative influences of the parameters are discussed in Section 6.5.1.

Table 6-1. Estimated Parameters in the MSS Models

Model	β1	β2	β3	β4	β5	β6	β9	var(U)	var(E)
2007[1], 2010[2]	9.9047 (0.61)	–0.4106 (0.11)	0.0164 (0.0030)	46.9397 (7.3)	0.003748 (0.00027)	0.9123 (0.054)		0.835 (0.080)	13.8279 (0.36)
2012[3]	9.8057 (0.74)	–0.1907 (0.28)	0.01839 (0.0051)	65.826 (12)	0.003191 (0.00021)	0.8753 (0.086)	0	0.9449 (0.083)	17.120 (1.2)
2012T[4]	10.916 (0.84)	–0.2104 (0.31)	0.01506 (0.0033)	13.497 (4.7)	0.003221 (0.00021)	0.8839 (0.065)	59.284 (10)	0.9373 (0.082)	17.0816 (1.2)

The values in parentheses are standard errors of the estimates.
[1] McDonnell, Stewart, and Smith (2007). $\bar{A} = 25$.
[2] McDonnell, Stewart, and Smith (2010). $\bar{A} = 25$.
[3] McDonnell, et al. (2012). $\bar{A} = 23.8$. No-threshold.
[4] McDonnell, et al. (2012). $\bar{A} = 23.8$. Threshold.

We are using this model to estimate lung function decrements for people ages 5 and older. However this model was developed using only data from individuals aged 18 to 35 and the age adjustment term $[\beta_1 + \beta_2 (Age_{ijk} - \bar{A})]$ in the model is not appropriate for all ages. In addition to this age term, the effects of age are also taken into account through the dependence of ventilation rate and body surface area on age. The APEX estimates of lung function risk for different age groups are also influenced by the time spent outdoors and the activities engaged in by those groups, which vary by age (see Appendix 6-E).

Clinical studies data for children which could be used to fit the model for children are not available at this time. In the absence of data, we are extending the model to ages 5 to 18 by holding the age term constant at the age 18 level. Since the response increases as age decreases in the range 18 to 35, this trend may extend into ages of children, in which case the responses of

children could be underestimated. However, the slope of the age term in the MSS model is estimated based on data for ages 18 to 35 and does not capture differences in age trend within this range; in particular, we don't know at what age the response peaks, which could be above or below 18. The evidence from clinical studies indicates that the responsiveness of children to O_3 is about the same as for young adults (ISA, 2012, p. 6-21). This suggests that the age term for children should not be higher than the age term for young adults.

Because the responses to O_3 decline from age 18 until around age 55 and for ages older than 55 the response are minimal, we let the age term for ages 35 to 55 linearly decrease to zero and set it to zero for ages > 55.

> "In healthy individuals, the fastest rate of decline in O_3 responsiveness appears between the ages of 18 and 35 years …. During the middle age period (35-55 years), O_3 sensitivity continues to decline, but at a much lower rate. Beyond this age (>55 years), acute O_3 exposure elicits minimal spirometric changes." (ISA, 2012, p. 6-23)

In order to extend the age term to ages outside the range of ages the MSS model is based on (ages 18-35), we parameterize the age term by $[\beta_1 + \beta_2(\alpha_1 \text{ Age} + \alpha_2)]$, for different ranges of ages (α_1 and α_2 depend on age), requiring that these terms match at each boundary to form a piecewise linear continuous function of age. The foregoing assumptions result in the following values of α_1 and α_2 for four age ranges (Table 6-2).

Table 6-2. Age Term Parameters for Application of the 2012 MSS Threshold Model to All Ages

Age Range	β_1	β_2	α_1	α_2
5 – 17	10.916	–0.2104	0	–5.8
18 – 35	10.916	–0.2104	1	–23.8
36 – 55	10.916	–0.2104	2.0341	–59.994
> 55	0	0	0	0

The lung function decrements estimated by the MSS (2010) model for a particular case are illustrated in Figure 6-2 and Figure 6-3. Figure 6-2 shows the predictions of the MSS model for 20-year old individuals with a (typical) body surface area (BSA) of 2 m^2 and a target ventilation rate of 40 L/min (moderate exertion) and an O_3 exposure level of 100 ppb, under the conditions of a typical 6.6-hour clinical study. Subjects alternated 50 minutes of moderate exercise with 10 minutes of rest for the first three hours, with the exercise occurring first. For the next 35 minutes (lunch), subjects continued exposure at rest. For the remaining three hours of the exposure period, subjects again alternated 50 minutes of exercise with 10 minutes of rest. The

inter-individual variability predicted by this model is depicted by the boxplots in this figure. The predictions for the median individual over time are given by the line. Minute-by-minute predictions for the median individual for an exposure level of 100 ppb are shown in Figure 6-3. The stairstep response results from the pattern of exercise and rest during the experiment.

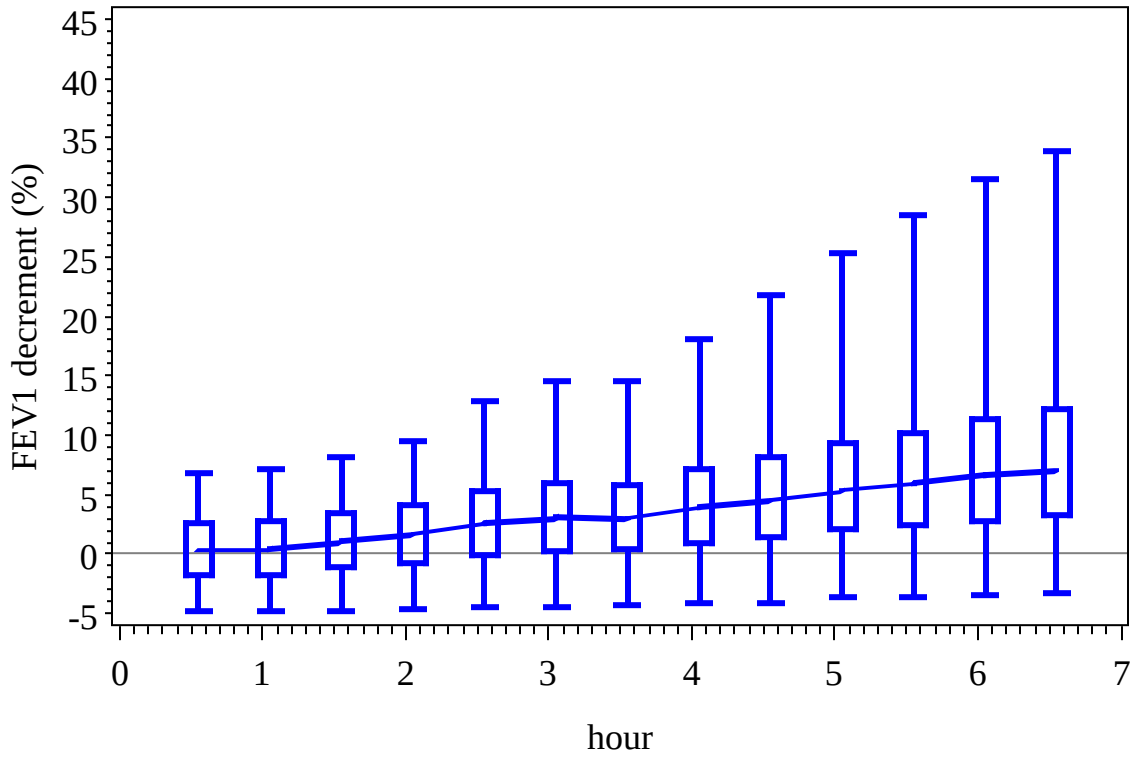


Figure 6-2. Distribution of Responses (Lung Function Decrements in FEV_1) Predicted by the MSS Model for 20-Year Old Individuals. Exposure to 100 ppb O_3 at Moderate Exercise (40 L/min, BSA=2 m^2) Under the Conditions of a Typical 6.6-hour Clinical Study.

The bottom and top edges of the boxes are at the 25th and 75th percentiles. The center horizontal line is drawn at the 50th percentile (median). The whiskers are at the 1st and 99th percentiles.

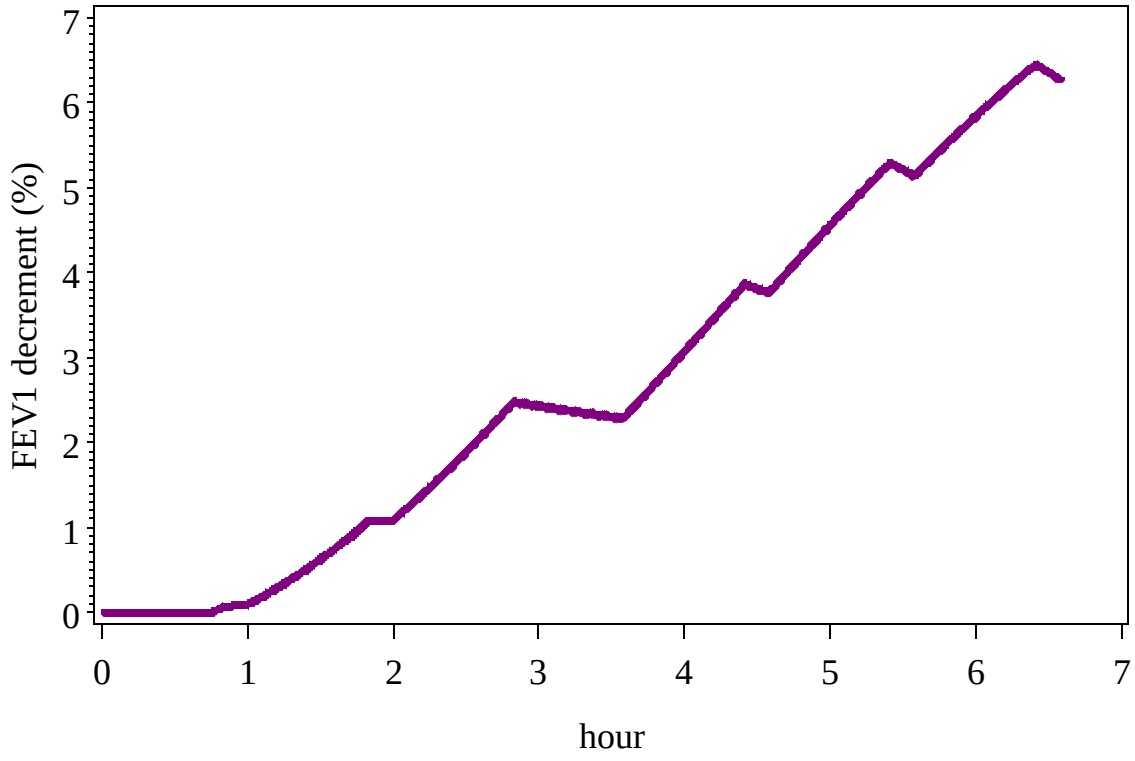


Figure 6-3. Median Response (Lung Function Decrements in FEV_1) Predicted by the MSS Model for 20-Year Old Individuals. Exposure to 100 ppb O_3 at Moderate Exercise (40 L/min, BSA=2 m^2) Under the Conditions of a Typical 6.6-hour Clinical Study.

Figure 6-4 and Figure 6-5 illustrate the threshold effect based on McDonnell et al. (2012). Figure 6-4 is a graph of the median response for a population of 20-year old individuals over a 6.6-hour time period. The exposure concentration is a constant 100 ppb over this time period, while the individuals are exercising from hour 1 to hour 3 and at rest otherwise. There is a 30-minute delay in response due to the threshold; without the threshold, the response starts increasing when exercise starts. Figure 6-5 shows the corresponding probability of a response (FEV_1 decrement) $\geq$ 10% over the time period for the two models. There is very little difference in response between the threshold and non-threshold models.

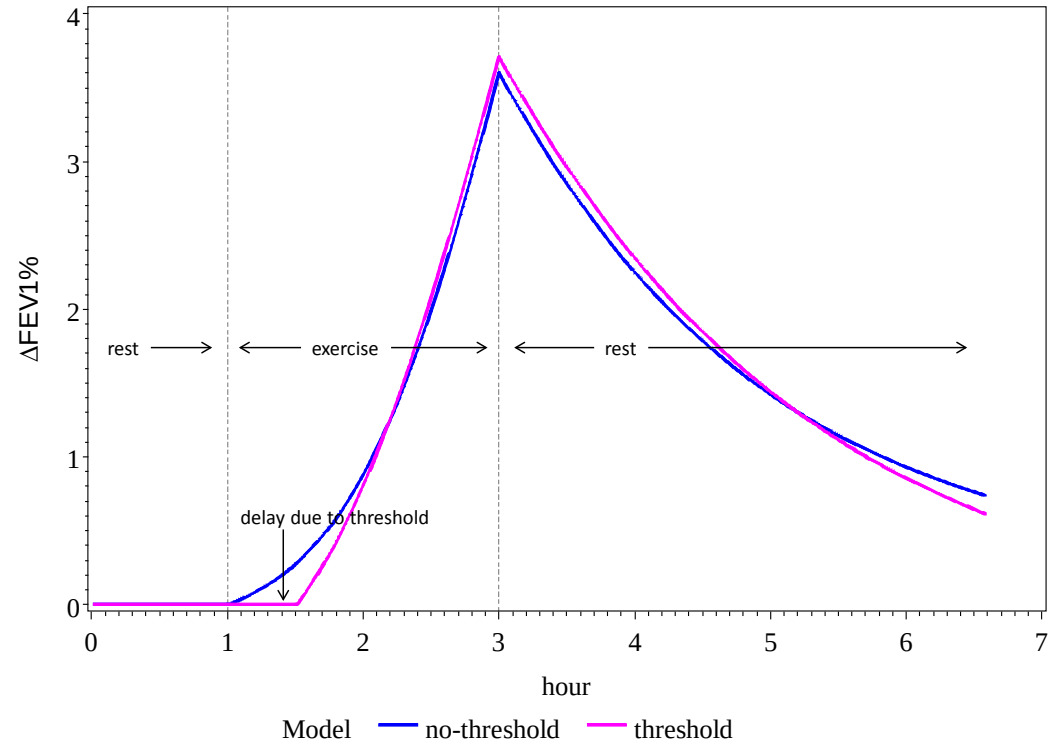


Figure 6-4. Median Response (FEV_1 Decrements) Predicted by the MSS Threshold and Non-Threshold Models for 20-Year Old Individuals, Constant 100 ppb O_3 Exposure, 2 Hours Heavy Exercise (30 L/min-m^2 BSA).

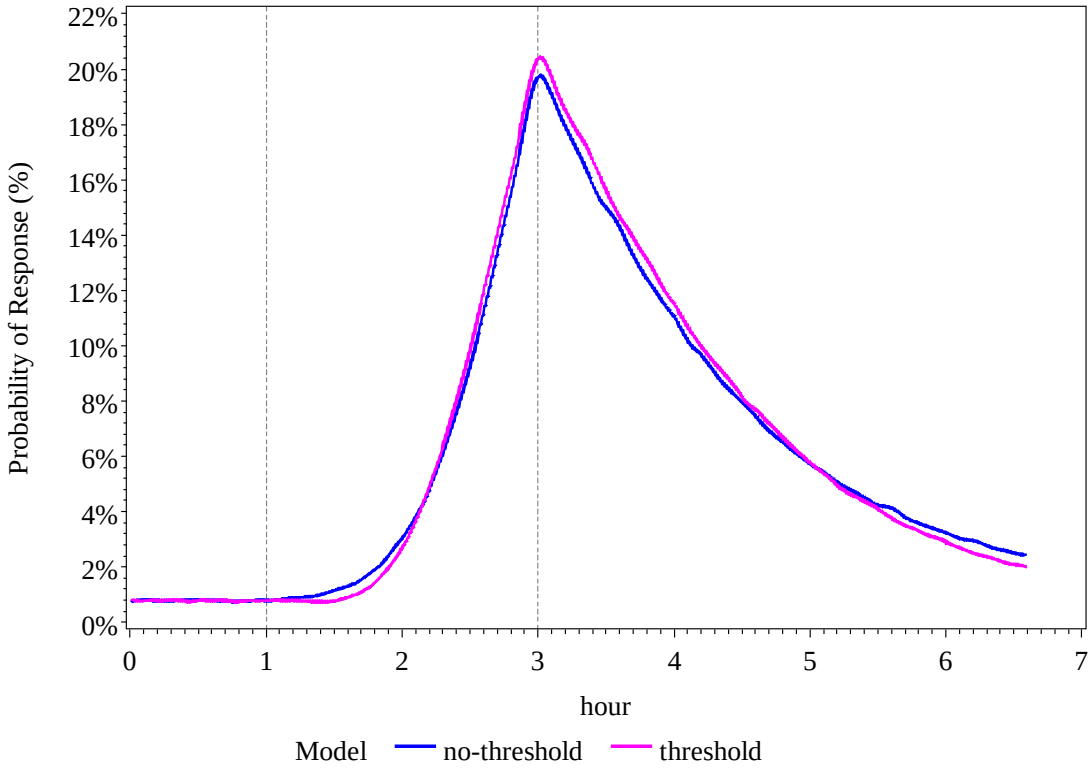


Figure 6-5. Probability of Response ≥ 10% Predicted by the MSS Threshold and Non-Threshold Models for 20-Year Old Individuals, Constant 100 ppb O_3 Exposure, 2 Hours Heavy Exercise (30 L/min-m^2 BSA).

6.2.5 The Exposure-Response Function Approach Used in Prior Reviews

As described in section 3.1.2 of the 2007 Risk Assessment Technical Support Document (EPA, 2007b), a Bayesian Markov Chain Monte Carlo approach (Lunn et al., 2012) was used to estimate probabilistic exposure-response relationships for lung function decrements associated with 8-hour O_3 exposures occurring at moderate exertion. In the previous review, summary data from the Folinsbee et al. (1988), Horstman et al. (1990), McDonnell et al. (1991), and Adams (2002, 2003, 2006) studies were combined to estimate exposure-response relationships for 8-hour exposures at moderate exertion for each of the three measures of lung function decrement ($\geq 10, \geq 15, \geq 20\%$ decrements in FEV_1). In this second draft REA we have updated this exposure-response function with the results from two additional studies (Kim et al., 2011; Schelegle et al., 2009). The controlled human exposure study data were corrected for the effect of exercise in clean filtered air on an individual basis to remove any systematic bias that might be present in the data attributable to an exercise effect (ISA, Section 6.2.1.1). This is done by subtracting the FEV_1 decrement in filtered air from the FEV_1 decrement (at the same time point) during exposure to O_3. An example of this calculation is given in Appendix 6-D.

Table 6-3 presents a summary of the study-specific results based on correcting all individual responses for the effect on lung function decrements of exercise in clean air.

Table 6-3. Study-specific O_3 Exposure-response Data for Lung Function Decrements Based on Correcting Individual Responses for the Effect on Lung Function of Exercise in Clean Air, Ages 18-35

Study, Grouped by Average O_3 Exposure	Protocol	Number Exposed	Number of Responses[a]		
			$\Delta FEV_1 \geq 10\%$	$\Delta FEV_1 \geq 15\%$	$\Delta FEV_1 \geq 20\%$
0.04 ppm O_3					
Adams et al. (2002)	Square-wave, face mask	30	2 (2)	0 (0)	0 (0)
Adams et al. (2006)	Triangular	30	0 (0)	0 (0)	0 (0)
0.06 ppm O_3					
Adams et al. (2006)	Square-wave	30	2 (2)	0 (0)	0 (0)
	Triangular	30	2 (2)	2 (2)	0 (0)
Kim et al. (2011)	Square-wave	59	3 (6)	1 (3)	0 (0)
Schelegle et al. (2009)	Variable levels (0.06 ppm avg)	31	4 (8)	2 (3)	1 (1)
0.07 ppm O_3					
Schelegle et al. (2009)	Variable levels (0.07 ppm avg)	31	6 (12)	3 (7)	2 (3)
0.08 ppm O_3					
Adams et al. (2002)	Square-wave, face mask	30	6 (6)	5 (5)	2 (2)

Table 6-3. Study-specific O_3 Exposure-response Data for Lung Function Decrements Based on Correcting Individual Responses for the Effect on Lung Function of Exercise in Clean Air, Ages 18-35

Study, Grouped by Average O_3 Exposure	Protocol	Number Exposed	Number of Responses[a]		
			$\Delta FEV_1 \geq 10\%$	$\Delta FEV_1 \geq 15\%$	$\Delta FEV_1 \geq 20\%$
Adams et al. (2003)	Square-wave, chamber	30	6 (6)	2 (2)	1 (1)
	Square-wave, face mask	30	5 (5)	2 (2)	2 (2)
	Variable levels (0.08 ppm avg), chamber	30	6 (6)	1 (1)	1 (1)
	Variable levels (0.08 ppm avg), face mask	30	5 (5)	1 (1)	1 (1)
Adams et al. (2006)	Square-wave	30	7 (7)	2 (2)	1 (1)
	Triangular	30	9 (9)	3 (3)	1 (1)
F-H-M[b]	Square-wave	60	17 (19)	11 (14)	8 (8)
Kim et al. (2011)	Square-wave	30	4 (6)	1 (1)	0 (0)
Schelegle et al. (2009)	Variable levels (0.08 ppm avg)	31	10 (15)	5 (8)	4 (6)
0.087 ppm O_3					
Schelegle et al. (2009)	Variable levels (0.087 ppm avg)	31	14 (17)	10 (12)	7 (9)
0.1 ppm O_3					
F-H-M[b]	Square-wave	32	13 (13)	11 (12)	6 (9)
0.12 ppm O_3					
Adams et al. (2002)	Square-wave, chamber	30	17 (17)	12 (12)	10 (10)
	Square-wave, face mask	30	21 (21)	13 (13)	7 (7)
F-H-M[b]	Square-wave	30	18 (19)	15 (15)	10 (10)

a. The first number in each cell is the number of responses based on post-exposure decrements in FEV_1 (i.e., we used only the last FEV_1 measurement and the pre-exposure FEV_1 to obtain a single percentage change in FEV_1 for each subject in each experiment). The numbers in parentheses are the numbers of responses based on maximum FEV_1 decrements. Specifically, when there were multiple FEV_1 measurements after the beginning of the exposure, we calculated multiple FEV_1 percentage changes for each subject in each experiment and used the maximum change when calculating the numbers of responses greater than 10%, 15%, and 20%.

b. F-H-M combines data from Folinsbee et al. (1988), Horstman et al. (1990), and McDonnell et al. (1991).

For the risk assessment conducted during the 2007 O_3 NAAQS review (EPA, 2007b), EPA considered both linear and logistic functional forms in estimating the exposure-response relationship and chose a 90 percent logistic/10 percent piecewise-linear split using a Bayesian Markov Chain Monte Carlo approach. This Bayesian estimation approach incorporates both model uncertainty and uncertainty due to sampling variability.

For each of the three measures of lung function decrement, EPA assumed a 90 percent probability that the exposure-response function has the following 3-parameter logistic form:[2]

$$y(x;\alpha,\beta,\gamma) = \frac{\alpha * e^{\gamma}(1-e^{\beta x})}{(1+e^{\gamma})(1+e^{\beta x+\gamma})}, \qquad \textbf{(Equation 6-4)}$$

where x denotes the O_3 concentration (in ppm) to which the individual is exposed, y denotes the corresponding response (decrement in $FEV_1 \geq 10\%$, $\geq 15\%$ or $\geq 20\%$), and α, β, and γ are the three parameters whose values are estimated.

We assumed a 10 percent probability that the exposure-response function has the following linear with threshold (hockey stick) form:

$$y(x;\alpha,\beta) = \begin{cases} \alpha + \beta x, & for\ \alpha + \beta x > 0 \\ 0, & for\ \alpha + \beta x < 0 \end{cases} \qquad \textbf{(Equation 6-5)}$$

We assumed that the number of responses, S, out of N subjects exposed to a given concentration, x, has a binomial distribution with response probability given by Eq (6-4) with 90 percent probability and response probability given by Eq (6-5) with 10 percent probability. In the 2007 review, we also considered 80/20 and 50/50 probabilities for the logistic and hockey stick forms, and ran those as sensitivity analyses. We performed those analyses with the updated data and found that for each of the three exposure-response curves, the 90/10 mix has smaller error in fit (weighted RMSE) than the other two combinations of probabilities, and we are using only that function in this review.

In some of the controlled human exposure studies, subjects were exposed to a given O_3 concentration more than once – for example, using a constant (square-wave) exposure pattern in one protocol and a variable (triangular) exposure pattern in another protocol. However, because there were insufficient data to estimate subject-specific response probabilities, we assumed a single response probability (for a given definition of response) for all individuals and treated the repeated exposures for a single subject as independent exposures in the binomial distribution.

For each of the two functional forms (logistic and linear), we derived a Bayesian posterior distribution using this binomial likelihood function in combination with prior distributions for each of the unknown parameters (Box and Tiao, 1973). We assumed lognormal priors with maximum likelihood estimates of the means and variances for the parameters of the logistic function, and normal priors, similarly with maximum likelihood estimates for the means and variances, for the parameters of the linear function. For each of the two functional forms

[2] The 3-parameter logistic function is a special case of the 4-parameter logistic, in which the function is forced to go through the origin, so that the probability of response to 0.0 ppm is 0.

considered, we used 1,000 iterations as the "burn-in" period[3] followed by 9,000 iterations for the estimation. Each iteration corresponds to a set of values for the parameters of the (logistic or linear) exposure-response function. We combined the 9,000 sets of values from the logistic model runs with the last 1,000 sets of values from the linear model runs to get a single combined distribution of 10,000 sets of values reflecting the 90 percent/10 percent assumption stated above. WinBUGS version 1.4.3 was used for these analyses (WinBUGS; Lunn et al., 2012).

For any O_3 concentration, *x*, we can derive the nth percentile response value, for any n, by evaluating the exposure-response function at *x* using each of the 10,000 sets of parameter values (9,000 of which were for a logistic model and 1,000 of which were for a linear model). The resulting 2.5^{th} percentile, median (50th percentile), and 97.5^{th} percentile exposure-response functions for changes in $FEV_1 \geq 10\%$ are shown in Figure 6-6, along with the response data to which they were fit. The corresponding exposure-response functions for changes in $FEV_1 \geq 15\%$ and $\geq 20\%$ are shown in Appendix 6-A. The values of the functions are also provided in Appendix 6-A.

[3] Markov chain Monte Carlo (MCMC) simulations require an initial adaptive "burn-in" set of iterations, which are not used. This allows the MCMC sampling to stabilize.

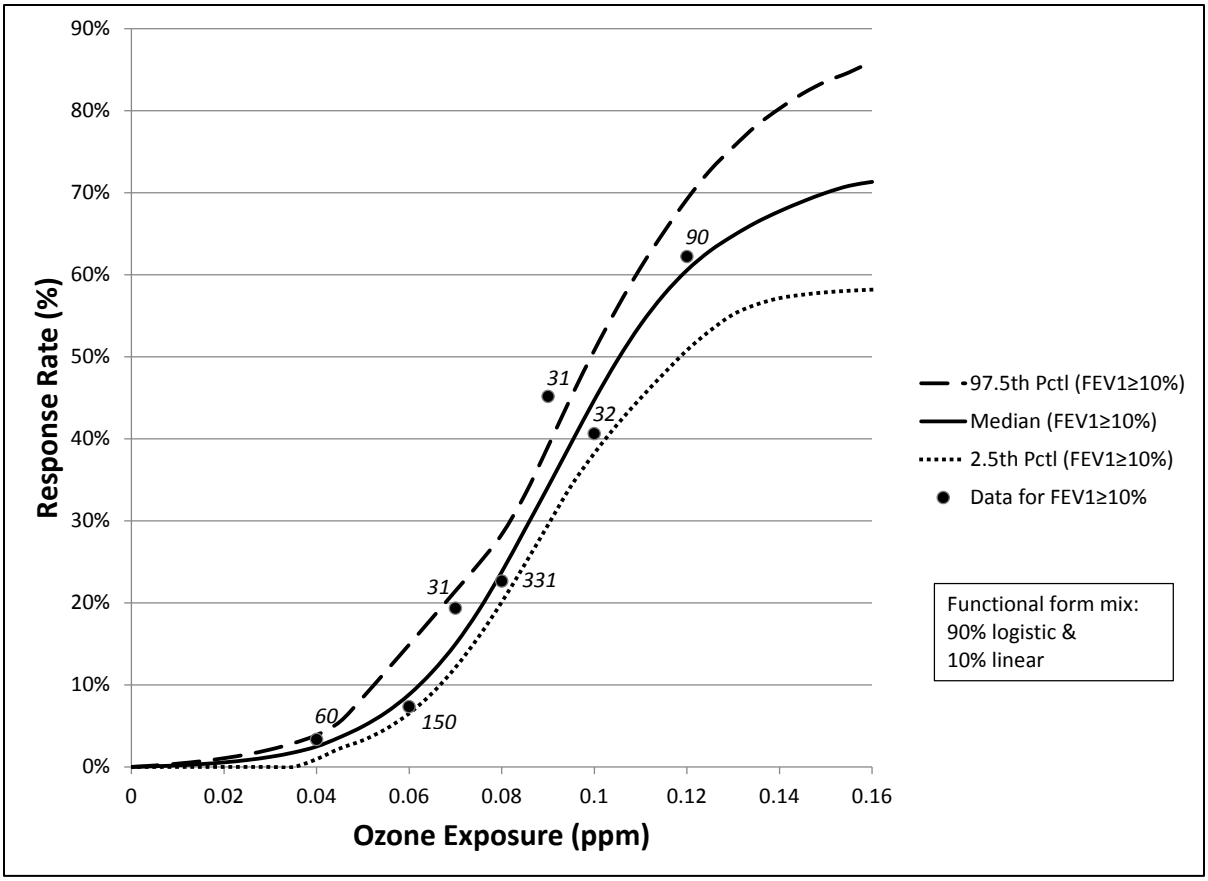


Figure 6-6. Probabilistic Exposure-Response Relationships for FEV_1 Decrements ≥ 10% for 8-Hour Exposures At Moderate Exertion, Ages 18-35. Values associated with data points are the number of subject-exposures at each exposure concentration.

The population risk is estimated by multiplying the expected risk by the number of people in the relevant population, as shown in Equation 6-6 below. The risk (i.e., expected fractional response rate) for the kth fractile, R_k is estimated as:

$$R_k = \sum_{j=1}^{N} P_j x(RR_k \mid e_j) \qquad \text{(Equation 6-6)}$$

where:

e_j = (the midpoint of) the jth category of personal exposure to O_3;

P_j = the fraction of the population having personal exposures to O_3 concentration of e_j ppm;

$RR_k \mid e_j$ = k-fractile response rate at O_3 exposure concentration e_j;

N = number of intervals (categories) of O_3 personal exposure concentration.

Exposure estimates used in this portion of the risk assessment were obtained from APEX for each of the fifteen urban areas and the five air quality scenarios. Chapter 5 provides additional details about the inputs and methodology used to estimate population exposure in the urban areas. Exposure estimates for all and asthmatic school-aged children (ages 5 to 18) were combined with probabilistic exposure-response relationships for lung function decrements associated with 8-hour exposure while engaged in moderate exertion. Individuals engaged in activities that resulted in an average equivalent (BSA-normalized) ventilation rate (EVR) for the 8-hour period at or above 13 L/min-m^2 BSA were included in the exposure estimates for 8-hour moderate or greater exertion. This range was selected based on the EVRs for the group of subjects in the controlled human exposure studies that were the basis for the exposure-response relationships used in this portion of the risk assessment.

6.3 O_3 RISK ESTIMATES

This section provides lung function risk estimates associated with several air quality scenarios: five recent years of air quality as represented by 2006 to 2010 monitoring data, and air quality in those years after simulating just meeting the existing O_3 standard and alternative standard levels of 0.070, 0.065, and 0.060 ppm. The risk measures presented here are the percents of the population estimated to experience lung function responses greater then 10, 15, and 20%, one or more times or six or more times during an O_3 season, for three age groups: school-aged children (ages 5-18), young adults (ages 19-35) and adults ages 36-55. Results for adults older than 55 are not presented since the responses for this age group are estimated to be minimal. People with multiple events with large lung function decreases are more at risk than those with only one such event during the O_3 season. Although six events is less than once per month, we see dramatic decreases in population risk in going from one or more to six or more events during a season, which is why we report on six or more events rather than a higher number.

In the figures and tables that follow, “base” indicates the base case scenario of recent air quality for the indicated year. “75,” “70,” “65,” and “60” respectively indicate the existing O_3 standard and alternative standard levels of 0.070, 0.065, and 0.060 ppm. “75 6-8” indicates the 0.075 ppm existing 8-hour standard based on rollback for the 2006-2008 period, while “75 8-10” indicates the existing standard scenario based on rollback for the 2008-2010 period. There are two estimates of results for the 2008 existing and alternative standard scenarios (because 2008 overlaps the two rollback periods) and one for each of the other four years. These two estimates for 2008 can be quite different because of the relationship between the design value over the three-year period and the amount of adjustment to the air quality distribution in 2008 that can result.

6.3.1 Lung Function Risk Estimates Based on the McDonnell-Stewart-Smith Model

Results based on the McDonnell-Stewart-Smith (2012) threshold model are summarized in this section; detailed results can be found in Appendix 6-B.

Figure 6-7 shows the results for school-aged children in the same format used in exposure results, explained in Section 5.3.1. Figure 6-7 depicts results for all cities, year, and scenarios for ages 5 to 18 with ≥ 1 occurrences of FEV_1 decrements ≥ 10, 15, 20% and illustrates the variation of results across cities, year, and scenarios.

Figure 6-8 shows the variation across cities (horizontally) and years (vertically) for the percent of school-aged children with ≥ 1 occurrences of FEV_1 decrements ≥ 10% with air quality just meeting the potential alternative standard of 0.07 ppm. The points above each study area on this graph represent the risk for the six years for the study area (2008 has two points, corresponding to the different 2006-2008 and 2008-2010 design values used to adjust the air quality to meet 0.07 ppm). There is substantial variability both across years and across cities. Denver has the highest overall risks, while Cleveland and New York have the lowest. Los Angeles has the smallest variation across years, with a range of 2.3% (from 14.3% to 16.6%). The other cities have a range of around 4% to 7.5% across years.

Table 6-4 and Table 6-5 present summary results (ranges over cities and years) of FEV_1 decrements ≥ 10 and 15% estimated by the MSS model for the different age groups. The results for asthmatic school-aged children are very similar to the results for all school-aged children and are not presented here.

Figure 6-9 illustrates the distribution of responses (FEV_1 decrements > 10%) across ranges of ambient concentrations of O_3 for school-aged children for one city and scenario (Los Angeles, 2006 recent air quality). The concentrations are daily 8-hour average ambient concentrations during the 8-hour period with maximum 8-hour average exposure for that day.

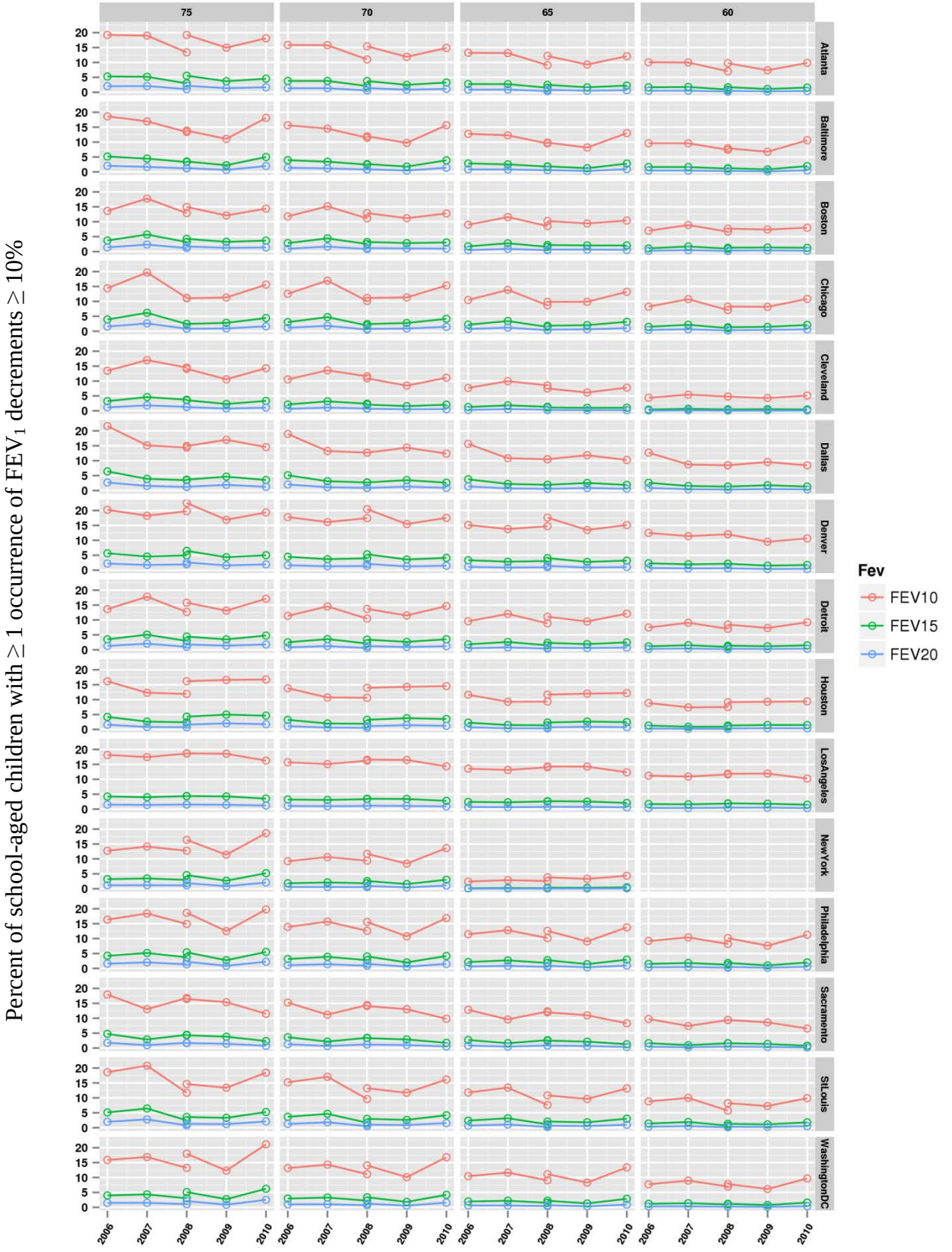


Figure 6-7. Risk results for all school-aged children with ≥ 1 occurrences of FEV_1 decrements ≥ 10, 15, 20% for all cities, year, and scenarios (y-axis is percent of children affected).

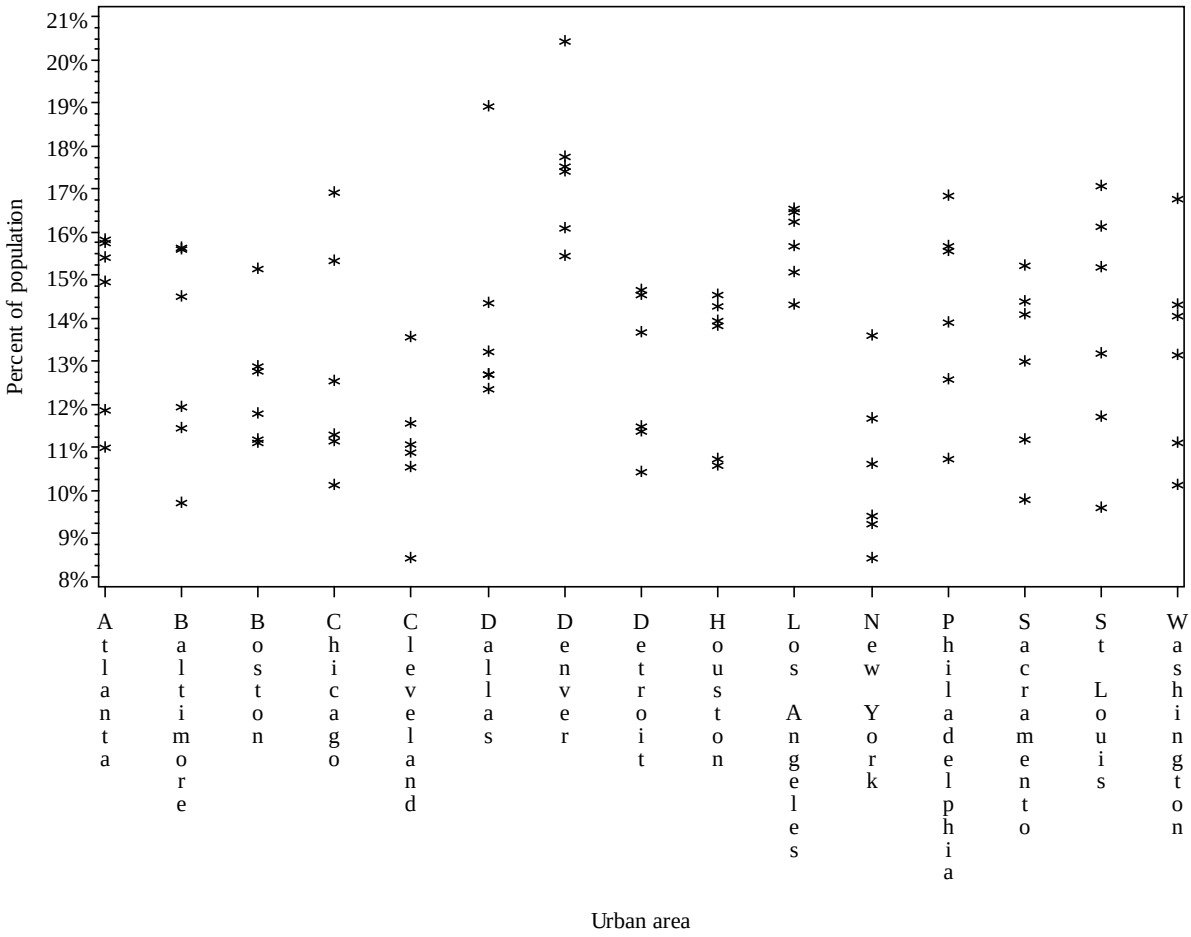


Figure 6-8. Risk results for all school-aged children with ≥ 1 occurrences of FEV_1 decrements ≥ 10% under the 0.07 ppm alternative standard showing variability across cities (horizontally) and years (vertically).

Table 6-4. Ranges of percents of population experiencing one or more days during the O_3 season with lung function decrement (ΔFEV_1) more than 10%. The numbers in this table are the minimum and maximum percents estimated over all cities and years.

		percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 10\%$		percent experiencing ≥ 6 days with $\Delta FEV_1 \geq 10\%$	
Age group	**Scenario**	**minimum**	**maximum**	**minimum**	**maximum**
5 to 18	base	11%	31%	1%	9%
5 to 18	75	11%	22%	1%	6%
5 to 18	70	8%	20%	1%	5%
5 to 18	65	2%	18%	0%	4%
5 to 18	60	4%	13%	0%	3%
19 to 35	base	3%	13%	0%	1%
19 to 35	75	3%	9%	0%	1%
19 to 35	70	2%	8%	0%	1%
19 to 35	65	1%	6%	0%	1%
19 to 35	60	1%	5%	0%	0%
36 to 55	base	1%	4%	0%	0%
36 to 55	75	1%	2%	0%	0%
36 to 55	70	0%	2%	0%	0%
36 to 55	65	0%	2%	0%	0%
36 to 55	60	0%	1%	0%	0%
> 55	All	0%	0%	0%	0%

Table 6-5. Ranges of percents of population experiencing one or more days during the O_3 season with lung function decrement (ΔFEV_1) more than 15%. The numbers in this table are the minimum and maximum percents estimated over all cities and years.

		percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 15\%$		percent experiencing ≥ 6 days with $\Delta FEV_1 \geq 15\%$	
Age group	**Scenario**	**minimum**	**maximum**	**minimum**	**maximum**
5 to 18	base	2%	12%	0%	3%
5 to 18	75	2%	6%	0%	1%
5 to 18	70	2%	5%	0%	1%
5 to 18	65	0%	4%	0%	1%
5 to 18	60	0%	3%	0%	0%
19 to 35	base	0%	3%	0%	0%
19 to 35	75	0%	2%	0%	0%
19 to 35	70	0%	1%	0%	0%
19 to 35	65	0%	1%	0%	0%
19 to 35	60	0%	1%	0%	0%
36 to 55	base	0%	1%	0%	0%
36 to 55	75	0%	0%	0%	0%
36 to 55	70	0%	0%	0%	0%
36 to 55	65	0%	0%	0%	0%
36 to 55	60	0%	0%	0%	0%
> 55	All	0%	0%	0%	0%

These concentrations are less than but close to daily maximum 8-hour average ambient concentrations and are greater than daily maximum 8-hour average exposure concentrations. The percents in this chart reflect the frequencies of person-days with FEV_1 decrements ≥ 10% within a concentration bin as percents of all person-days with FEV_1 decrements ≥ 10%. Figure 6-9 shows that more than 90% of daily instances of FEV_1 decrements ≥ 10% occur when 8-hour average ambient concentrations are above 40 ppb for this modeled scenario. This distribution will be different for different cities, years, and air quality scenarios.

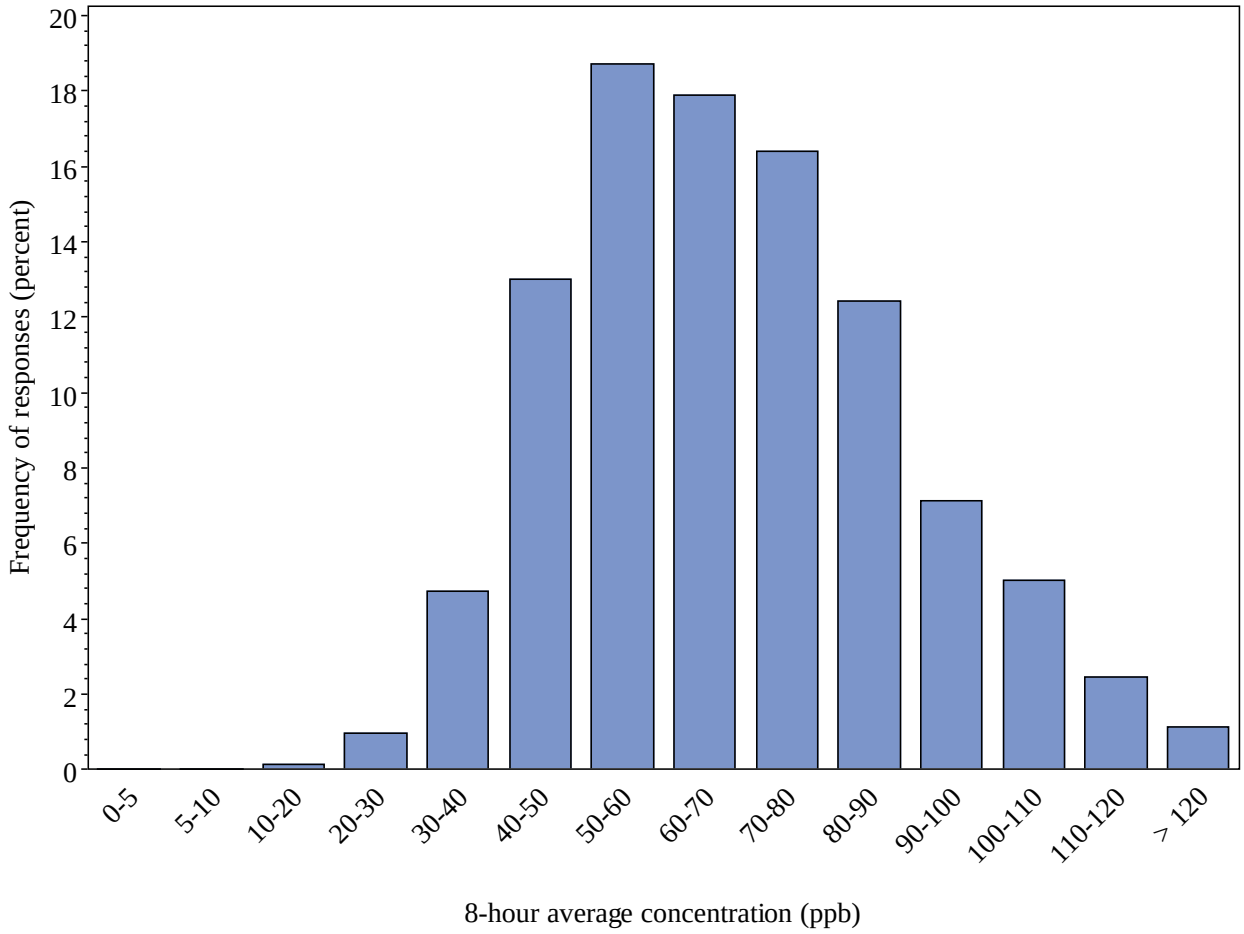


Figure 6-9. Distribution of Daily FEV_1 Decrements ≥ 10% Across Ranges of 8-hour Average Ambient O_3 Concentrations (Los Angeles, 2006 recent air quality).

Outdoor workers spend more time outdoors than the general population and therefore are at higher risk for health effects due to O_3. We conducted simulations of outdoor workers ages 19-35 for Atlanta (2006) for the current and alternative standards to estimate the risk of this group for experiencing FEV_1 decrements ≥ 15%. The methodology for simulating outdoor workers involves modifying activity diaries to represent outdoor workers and is described in Section 5.3.3 in Chapter 5. Table 6-6 shows the results of these simulations and compares them with the results for the general population, ages 19-35. The percents of people experiencing one or more FEV_1 decrements ≥ 15% during the 2006 O_3 season in Atlanta are 3.6 times higher for outdoor workers than for the general population (ages 19-35) under the current standard, and range up to 5.3 times higher for the alternative standards. The percents of people experiencing six or more FEV_1 decrements ≥ 15% during the 2006 O_3 season in Atlanta are 20 times higher for outdoor workers than for the general population under the current standard, and range up to 150 times higher for the alternative standards. As expected, we see that the risk of repeated occurrences of FEV_1 decrements ≥ 15% is much greater for outdoor workers than for the general population.

Part of the reason for this is that APEX tends to underestimate the number of individuals who have very repetitive activity patterns (e.g., 9 to 5 weekdays office workers) when using the CHAD activity database and the method selected for generating longitudinal diary profiles (see Section 5.3.1).

Table 6-6. Percents of the General Population and Outdoor Workers (ages 19-35) Experiencing 1 or More and 6 or More FEV_1 Decrements ≥ 15% (based on Atlanta 2006 APEX simulations)

	General population ages 19-35	Outdoor workers ages 19-35
1 or more occurrences		
Current standard	1.2%	4.3%
70 ppb alt. std.	0.84%	3.2%
65 ppb alt. std.	0.55%	2.5%
60 ppb alt. std.	0.32%	1.7%
6 or more occurrences		
Current standard	0.06%	1.2%
70 ppb alt. std.	0.018%	0.93%
65 ppb alt. std.	0.005%	0.74%
60 ppb alt. std.	0.005%	0.55%

6.3.2 Lung Function Risk Estimates Based on the Exposure-Response Functions Approach Used in Prior Reviews

In this section we present lung function risk estimates for all school-aged children following the methodology used in previous reviews, based on updated exposure-response (E-R) functions. In Appendix 6-C we compare these estimates with those from the previous review.

Table 6-7 provides an overall summary of results for each air quality scenario by tabulating the minimum and maximum estimates over all cities and years of percents of all school-aged children (ages 5 to 18) experiencing one or more days (during the O_3 season) with FEV_1 decrement more than 10 and 15%. This table can be compared with Table 6-4 and Table 6-5, which have analogous results for the MSS model. These results are much lower than the MSS model results. The reasons for this are described in Section 6.3.3 below.

Table 6-7. Ranges of percents of school-aged children experiencing one or more days during the O_3 season with lung function decrement (ΔFEV_1) more than 10 and 15%. The numbers in this table are the minimum and maximum percents estimated over all cities and years.

Scenario	minimum percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 10\%$	maximum percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 10\%$	minimum percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 15\%$	maximum percent experiencing ≥ 1 day with $\Delta FEV_1 \geq 15\%$
base	2%	11%	0%	5%
75	2%	6%	1%	2%
70	2%	6%	0%	2%
65	1%	5%	0%	1%
60	2%	3%	0%	1%

6.3.3 Comparison of the MSS Model with the Exposure-Response Function Approach

There are two key differences between the MSS and E-R models. The E-R model estimates the distribution of FEV_1 decrements across the population or study group, whereas the MSS model estimates FEV_1 decrements at the individual level and then these are aggregated to obtain the population distribution. Thus the MSS model allows for detailed analyses of conditions that influence risk. Second, the E-R model estimates FEV_1 decrements only for 8-hour average exposures when the 8-hour average exertion level is moderate or greater. The MSS model estimates FEV_1 decrements for any averaging time and therefore accounts for a wider range of activities that might result in FEV_1 decrements.

A comparison of the MSS model with the exposure-response function approach for the 2006 existing standard scenarios is summarized in Table 6-8, which lists estimates of the percents of school-aged children estimated to experience lung function responses greater then 10, 15, and 20%. The MSS model estimates are significantly higher than the exposure-response function approach estimates. In most cases, the MSS model gives results about a factor of three higher than the exposure-response function model for school-aged children. This is expected, since, as discussed above, the MSS model includes responses for a wider range of exposure protocols (under different levels of exertion, lengths of exposures, and patterns of exposure concentrations) than the exposure-response model of previous reviews.

Table 6-8. Comparison of responses from the MSS model with responses from the population exposure-response (E-R) method. 2006 existing standard, ages 5 to 18

Urban area	≥ 10% FEV_1 decrement		≥ 15% FEV_1 decrement		≥ 20% FEV_1 decrement	
	MSS model	E-R method	MSS model	E-R method	MSS model	E-R method
Atlanta	19.2%	5.6%	5.3%	1.7%	2.1%	0.7%
Baltimore	18.6%	5.4%	5.2%	1.6%	2.1%	0.7%
Boston	13.6%	4.5%	3.7%	1.2%	1.4%	0.5%
Chicago	14.4%	4.7%	3.9%	1.3%	1.6%	0.5%
Cleveland	13.5%	4.2%	3.3%	1.1%	1.2%	0.4%
Dallas	21.6%	6.0%	6.4%	1.8%	2.7%	0.8%
Denver	20.2%	5.8%	5.6%	1.7%	2.2%	0.7%
Detroit	13.6%	4.4%	3.5%	1.2%	1.3%	0.4%
Houston	16.2%	4.7%	4.2%	1.3%	1.6%	0.5%
Los Angeles	18.2%	4.8%	4.2%	1.2%	1.5%	0.5%
New York	12.7%	4.2%	3.2%	1.1%	1.2%	0.4%
Philadelphia	16.4%	4.8%	4.2%	1.3%	1.6%	0.5%
Sacramento	17.9%	5.1%	4.8%	1.4%	2.8%	0.6%
St. Louis	18.6%	5.4%	5.1%	1.6%	2.0%	0.6%
Washington	15.9%	4.6%	4.0%	1.3%	1.5%	0.5%

Since the E-R method of the previous reviews only looks at 8-hour exposures concomitant with EVR ≥ 13 L/min-m^2 BSA (hereafter, EVR ≥ 13), it is of interest to compare the E-R method results with the corresponding MSS model results (instances of $\Delta FEV_1 \geq 10$, 15, 20% concomitant with EVR ≥ 13).

We performed this comparison for four APEX simulations: the Atlanta March 1-October 30, 2006 base case, ages 18-35; the Los Angeles May 29-July 28, 2006 base case, age 25; the Los Angeles May 29-July 28, 2006 base case, ages 18-35; and the Los Angeles Jan 1-Dec 31, 2006 base case, ages 18-35.

For the Atlanta simulation, the E-R function approach gives 5.0, 1.8, and 0.9% responding for $\Delta FEV_1 \geq 10$, 15, 20%. The MSS model approach gives 11.54, 3.26, and 1.28% responding for $\Delta FEV_1 \geq 10$, 15, 20%. The percents of the population for $\Delta FEV_1 \geq 10$, 15, 20% at the end of the daily max 8-hour average exposure period where the concomitant 8-hour average EVR is ≥ 13 are 6.67%, 2.09%, and 0.84%. 15.17% of the population never have any instances of EVR ≥ 13 and 0.41% have at least one occurrence of $\Delta FEV_1 \geq 10\%$ while never having EVR ≥ 13 for any 8-hour period. 4.46% (not among the 15.17%) have instances of $\Delta FEV_1 \geq 10\%$ but

none of those instances with concomitant EVR ≥ 13. The 11.54% responding is made up of 6.67% of the population with instances of $\Delta FEV_1 \geq 10\%$ concomitant with EVR ≥ 13 and 4.87% with instances of $\Delta FEV_1 \geq 10\%$ not concomitant with EVR ≥ 13.

Table 6-9 and Table 6-10 summarize the pertinent results for the Atlanta and three Los Angeles simulations. Looking at the first rows of these tables shows that these models have similar corresponding results. The broader scope of activity/exposure patterns encompassed by the MSS model, beyond the 8-hour average EVR ≥ 13 restriction of the E-R model, contributes from a third to a half to the total MSS model risk and to a large part explains the differences between the models for ages 18-35. The difference between the MSS and E-R models is larger for school-aged children than for adults ages 18-35 due to the increased EVR and time spent outdoors in children compared to adults.

Table 6-9. Comparison of MSS Model and E-R Model of Previous Reviews for Atlanta, Mar 1-Oct 30, 2006, ages 18-35

Component of results	MSS model $\Delta FEV_1 \geq 10\%$	E-R model $\Delta FEV_1 \geq 10\%$	MSS model $\Delta FEV_1 \geq 15\%$	E-R model $\Delta FEV_1 \geq 15\%$	MSS model $\Delta FEV_1 \geq 20\%$	E-R model $\Delta FEV_1 \geq 20\%$
profiles with instances of $\Delta FEV1 \geq$ cutoff concomitant with 8-hour EVR ≥ 13	6.7%	5.0%	2.1%	1.8%	0.8%	0.9%
profiles with instances of $\Delta FEV_1 \geq$ cutoff never concomitant with 8-hour EVR ≥ 13	4.8%		1.2%		0.5%	
Final result of each model	11.5%	5.0%	3.3%	1.8%	1.3%	0.9%

Table 6-10. Comparison of MSS Model and E-R Model of Previous Reviews for Los Angeles, Jan 1-Dec 31, 2006, ages 18-35

Component of results	MSS model $\Delta FEV_1 \geq 10\%$	E-R model $\Delta FEV_1 \geq 10\%$	MSS model $\Delta FEV_1 \geq 15\%$	E-R model $\Delta FEV_1 \geq 15\%$	MSS model $\Delta FEV_1 \geq 20\%$	E-R model $\Delta FEV_1 \geq 20\%$
profiles with instances of $\Delta FEV1 \geq$ cutoff concomitant with 8-hour EVR ≥ 13	7.9%	6.2%	2.6%	2.6%	1.2%	1.4%
profiles with instances of $\Delta FEV_1 \geq$ cutoff never concomitant with 8-hour EVR ≥ 13	6.5%		1.8%		0.8%	
Final result of each model	14.4%	6.2%	4.4%	2.6%	2.0%	1.4%

Figure 6-10 compares the E-R function to the response curve of the MSS model restricted to 8-hour average EVR ≥ 13 and shows that these curves are very close. The MSS model has a higher response for the low and high ranges of exposure concentrations, while the E-R model is higher in the mid-range of exposures.

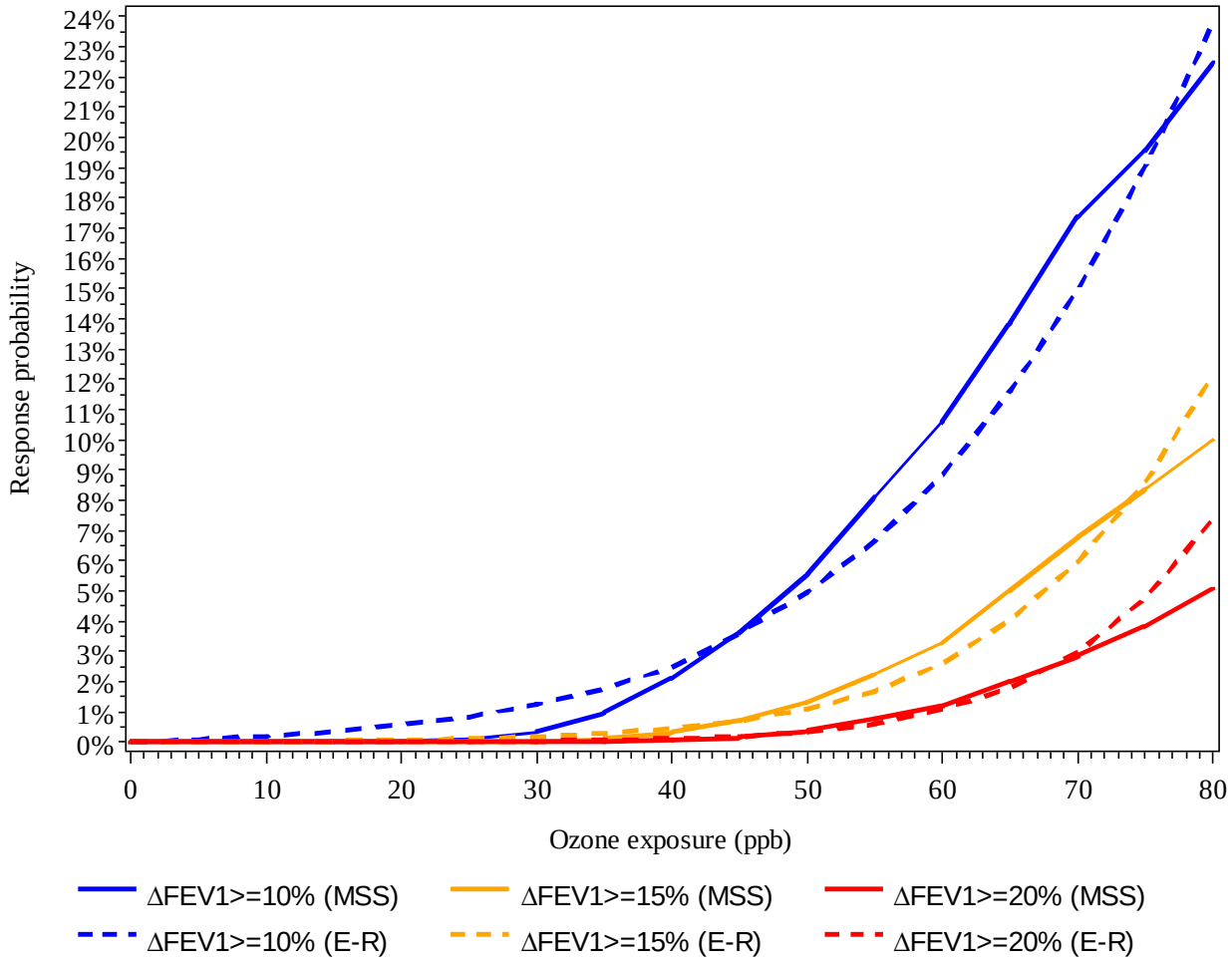


Figure 6-10. Comparison of E-R and MSS Model (restricted to 8-hour average EVR ≥ 13) Response Functions (Atlanta 2006 base case, ages 18-35).

Another element of the difference between the models derives from the distribution of EVR in the clinical studies the E-R approach is based on and how this compares to the distribution of EVR in the APEX simulations. Most of the clinical studies are conducted with a target EVR of 20 L/min-m^2 BSA and the actual EVRs vary somewhat around this value. The rationale for the cutpoint of 13 L/min-m^2 BSA is described in EPA's responses to comments on the 1996 proposed rule on the NAAQS for O_3 (Federal Register, 1996) as "for the 8-hr health risk assessment the range (based on being within 2 standard deviations of the mean) of EVRs observed in the subjects who participated in the study [McDonnell et al., 1991] was 13-27 liters per minute per meter squared [BSA] (L/min-m^2)." Figure 6-11 shows the distribution of EVRs ≥

13 for the Atlanta simulation and is clearly shifted much lower than the distribution of EVR in the clinical studies. This could lead to an overestimation of the percent of responders by the E-R method, since higher EVRs lead to higher lung function decrements and it is applying an E-R function based on EVRs around 20 to a population with median EVRs around 14.5.

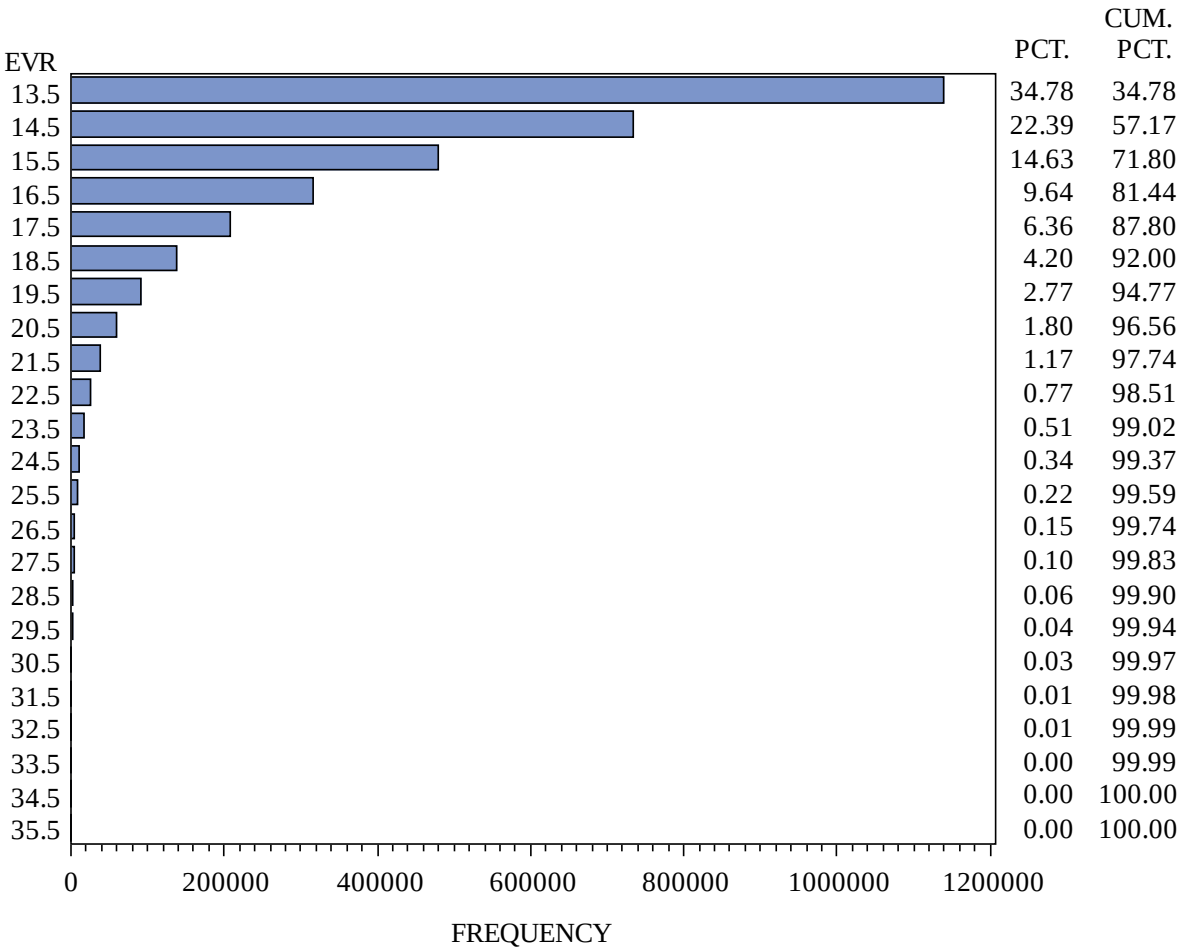


Figure 6-11. Distribution of Daily Maximum 8-hour Average EVR For Values of EVR ≥ 13 (L/min-m^2) (midpoints on vertical axis) (Atlanta 2006 base case, ages 18-35).

6.4 EVALUATION OF THE MSS MODEL

6.4.1 Summary of Published Evaluations

McDonnell et al. (2010) performed a detailed evaluation of their model using two methods: (1) cross-validation and (2) comparison of an independent data set against the predictions of the model.

The cross-validation was based on the data set of 15 EPA studies from which their original model was developed (McDonnell et al., 2007). This data set has 541 subjects, each with multiple measurements during single experiments. Subjects were omitted from the data set, one at a time, the model refit to the reduced data set, and the resulting parameters used to predict the FEV_1 decrements for the omitted subject. The authors then compare the mean predictions and mean observed values for each subject and presented these results in a scatter plot (Figure 1b, McDonnell et al., 2010). The observations exhibit much more variability than the predictions; for observed values of 20%, predicted values range from around 2 to 19%; and all observed values above 20% are underpredicted (the observed values range from -20 to 60%, while the predicted values range from 0 to 20%). These features result from the omission of the inter- and intra-individual variability terms (U_i and ε_{ijk}) in the MSS model (equation 6-3), which are accounted for in the risk estimates in this chapter.

Model predictions were compared against an independent data set of seven clinical studies with a total of 204 subjects (McDonnell et al., 2010). Graphs of predicted and observed study means vs. time show fair to good model fit. The authors do not present overall fit statistics that are directly commensurate with the statistics of interest in this risk assessment: the proportions of people with FEV_1 decrements greater than 10, 15, and 20%.

McDonnell et al. (2012) do compare observed and predicted proportions of people with FEV_1 decrements greater than 10, 15, and 20% and provide the corresponding scatter plots (Figure 4). They find the model to be unbiased, with the slopes of the observed vs. predicted lines for 10, 15, and 20% to be around 1.0 and the R^2 respectively 0.78, 0.73, and 0.67. The higher observed proportions of people with FEV_1 decrements greater than 10, 15, and 20% tended to be substantially underpredicted.

6.4.2 Children

A clinical study with children (ages 8-11; mean, 10 years; n=22), exposed to 120 ppb O_3 over 2.5 hours at heavy exertion levels was done by McDonnell et al. (1985). This study could be used to fit the model for children if all of the measurements of FEV_1 and ventilation rates were available. The paper lists the end-of exposure FEV_1 responses for each individual (but not ventilation rates), which we use to compare with the MSS model with the age term extension described in Section 6.2.4. The numbers of subjects with clean-air adjusted responses greater

than 10%, 15%, and 20% are respectively 4, 2, and 1, corresponding to 18.2%, 9.1%, and 4.5% of the number of subjects. We ran the MSS 2010 model using the mean and standard deviation of the ventilation rates reported in the paper. Resting ventilation rates were assumed to be 10.4 L/min (Avol et al., 1985) and BSA to be 1.08 m^2 (EPA, 2011). Details of this comparison can be found in Appendix 6-D.

Table 6-11 compares the results of this simulation with the results of McDonnell et al. (1985). The agreement is fairly good. Due to the limited sample size of 22 subjects from only one study and the assumptions made in running the MSS model, this does not provide confirmation that the age term extension is correct; on the other hand, this comparison does not indicate that there is a problem with the age term extension. Information is not available that would allow us to provide respectable confidence intervals for these estimates.

Table 6-11. Comparison of Responses from the MSS 2010 Model with Responses from McDonnell et al. (1985)

	≥ 10% FEV1 decrement		≥ 15% FEV1 decrement		≥ 20% FEV1 decrement	
	MSS model	**McDonnell et al. (1985)**	**MSS model**	**McDonnell et al. (1985)**	**MSS model**	**McDonnell et al. (1985)**
Percent responding	18.4%	18.2% (4 subjects)	6.8%	9.1% (2 subjects)	2.3%	4.5% (1 subject)

6.4.3 Threshold vs. Non-Threshold Models

The difference between the results of the MSS threshold and non-threshold models is minor, with the threshold version estimates of lung function decrements almost identical to the no-threshold version for the Atlanta 2006 recent air quality base case, as can be seen by comparing Table 6-12 with Table 6-13. This is consistent with the logistic form of the model, where the impact of exposures to low concentrations on risk is small.

Table 6-12. Percents of the population by age group with one or more days during the O_3 season with lung function (FEV_1) decrements more than 10, 15, and 20% (Atlanta 2006 base case). MSS Threshold model, monitors air quality.[4]

Age Group	$\Delta FEV_1 \geq$ 10%	$\Delta FEV_1 \geq$ 15%	$\Delta FEV_1 \geq$ 20%
5 to 18	31%	13%	6.4%
19 to 35	11%	3.1%	1.3%
36 to 55	3.7%	0.60%	0.14%

Table 6-13. Percents of the population by age group with one or more days during the O_3 season with lung function (FEV_1) decrements more than 10, 15, and 20% (Atlanta 2006 base case). MSS No-Threshold model, monitors air quality.

Age Group	$\Delta FEV_1 \geq$ 10%	$\Delta FEV_1 \geq$ 15%	$\Delta FEV_1 \geq$ 20%
5 to 18	31%	13%	6.6%
19 to 35	11%	3.1%	1.3%
36 to 55	3.8%	0.60%	0.15%

6.5 CHARACTERIZATION OF UNCERTAINTY

In the controlled human exposure study based risk assessment, there are two broad sources of uncertainty to the risk estimates. One of the most important sources of uncertainty is the estimation by APEX of the population distribution of individual time series of O_3 exposures and ventilation rates. The uncertainty regarding these estimated exposures is discussed in Chapter 5; they are not discussed further here.

[4] In the first draft REA, monitor-level air quality was provided as input to the APEX model. As discussed in Chapter 5, tract-level air quality was used in APEX for this second draft REA. Monitor-level air quality is used for the APEX simulations here, since these simulations take less time to run. This does not affect the analyses here, since the two air quality formats yield very similar results (see Appendix 6-F).

In this section, uncertainties associated with the second broad source of uncertainty in the risk calculation are discussed, namely, uncertainties in the lung function risk model. The specific sources of uncertainty covered are:

- Statistical model form
- Convergence of APEX results
- Application of model for all lifestages
- Application of model for asthmatic children
- Interaction between O_3 and other pollutants

6.5.1 Statistical Model Form

The MSS model is a 2-compartment model, the form of which is based on physical considerations. It accomodates these key features of human exposure studies: (1) FEV_1 responses increase with increasing O_3 concentration, ventilation rate, and duration of exposure, (2) the effect of each of these three variables depends on the levels of the other two variables, (3) FEV_1 responses depend on age, (4) certain individuals are consistently more responsive to O_3 exposure, and (5) O_3 –induced FEV_1 decrements improve within a few hours of cessation of exposure (McDonnell et al, 2007). These considerations support the form of the model, as do model evaluation that have been performed (Section 6.4.1). Although the model does not have good predictive ability for individuals (psuedo-R^2 0.28), it does better at predicting the proportion of individuals with FEV_1 decrements $\geq$ 10, 15, and 20% (psuedo-R^2s of 0.78, 0.74, 0.68) (McDonnell et al, 2012).

The clinical studies that these models' estimates are based on were conducted with young adult volunteers rather than randomly selected individuals, so it may be that selection bias has influenced the model parameter estimates.

The parameter estimates are not very precise, as the result of the likelihood surface being somewhat flat in the neighborhood of the maximum likelihood estimates. Table 6-14 gives 95 percent confidence intervals for each of the parameter estimates as percents of the estimates, based on the standard errors reported by McDonnell et al. (2012). Figure 6-10 shows how much the modeled number of children with one or more FEV_1 decrements $\geq$ 10% changes when each parameter is increased by five percent (keeping the other parameters fixed at their estimates). The scenario modeled is the Los Angeles 2006 recent conditions base case. The physiological parameter MET, a measure of the level of exertion for a given activity (see Appendix 6-E), is also included here for comparison. MET is a key variable in calculating ventilation rates and is specified by a distribution for each activity. Here we have shifted all MET distributions by +5% of their means.

Table 6-14. MSS threshold model estimated parameters with confidence intervals

	β1	β2	β3	β4	β5	β6	β9	var(U)	var(E)
parameter estimate	10.916	-0.2104	0.01506	13.497	0.003221	0.8839	59.284	0.9373	17.0816
standard error	0.8446	0.31	0.00333	4.734	0.000207	0.0647	10.192	0.0824	1.1506
95% conf. interval	±15%	±289%	±43%	±69%	±13%	±14%	±34%	±17%	±13%

from McDonnell et al. (2012).

The most influential parameter in Figure 6-12 is β_6, the power to which ventilation rate is raised in the MSS model. An increase of five percent in β_6 leads to 27, 40, and 47 percent increases respectively in the modeled number of children with FEV_1 decrements $\geq$ 10, 15 and 20%. The next most influential parameter is the variance of E, the intra-individual variability term. The least influential parameter is β_2, the slope of the age term. These changes of five percent are much less than the 95 percent confidence intervals of the parameter estimates, so the uncertainty in the risk estimates resulting from parameter uncertainty is likely to be more than is indicated in Figure 6-12.

Age Term Significance

As discussed in Section 6.5.3 below, there are uncertainties in extrapolating the MSS model down to age 5 from the age range of 18 to 35 to which the model was fit. Further considerations indicating that the uncertainty of the extension to children of the MSS model could be substantial are that the age coefficient $\beta_2 = -0.21$ (s.e. 0.31) in the MSS model is not statistically significantly different from zero; and when the MSS model is fit to the U.C. Davis clinical data the age term is positive, $\beta_2 = +0.19$ (0.60), although also not statistically significantly different from zero (McDonnell et al., 2012). Note that, in the previous section, β_2 was found to be the least influential model parameter.

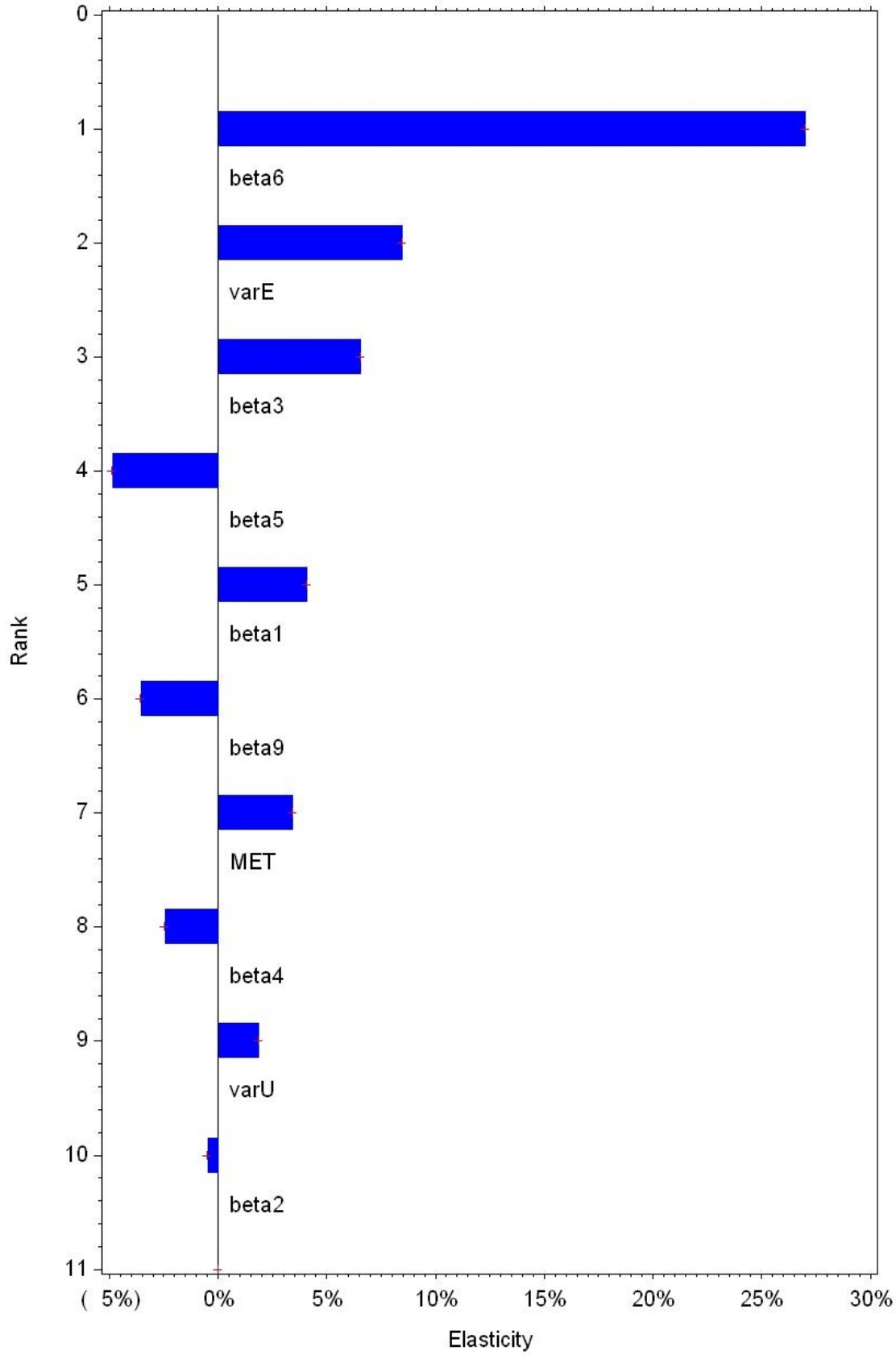


Figure 6-12. Sensitivity (Percent Change) of Population With One or More FEV_1 Decrements ≥ 10% to a 5% Increase in Individual MSS Model Parameter Estimates.

The Variability Term ε

The variability term ε in equation 6-3 is assumed by the MSS model to have a Gaussian distribution with mean zero and estimated standard deviation 4.135 (in the threshold model). Since the actual values are bounded, we truncate the variability term distribution at ±2 standard deviations (±8.27), a convention we use for the distributions of several physiological variables input to APEX in the physiology input file. To look at the effect of truncating the variability distribution, we conducted simulations with the variability term truncated at ±20, the range of the actual values of the variability term. We find that this constraint has a very large effect on estimates of percents of the population with FEV_1 decrements ≥ 10 and 15% and less of an effect for 20%. The percent of children with FEV_1 decrements ≥ 10% increases from 31% to 92% when increasing the truncation point from 8.27 to 20. Details of this comparison and additional results are presented in Appendix 6-F. The assumption that the distribution of the variability term ε is Gaussian is convenient for fitting the model, but is not accurate. The extent to which this mis-specification affects the estimates of the parameters of the MSS model is not clear.

6.5.2 Convergence of APEX Results

APEX accounts for several sources of variability by drawing random variables from specified distributions. Some variables are drawn once for each simulated individual (e.g., age, location of residence), some are drawn every day or every hour for each simulated individual, and others are drawn more frequently, at the event level (e.g., activity). Increasing the number of individuals simulated in an APEX run increases the accuracy of the modeled variability and the results of the APEX runs are more reproducible. In order to assess the number of individuals to simulate to achieve convergence of APEX results, we perform multiple APEX runs with identical inputs except for the random number seed, and look at the variability of the results of these model runs. Table 6-15 summarizes the results of 40 APEX simulations of the Atlanta 2006 base case with 200,000 simulated individuals. For each of these measures, the range of results over the 40 APEX runs is less than one percent. This analysis of the convergence of APEX results shows that modeling 200,000 simulated individuals is adequate for reasonable convergence of the FEV_1 risk measures.

Table 6-15. Convergence results for the Atlanta 2006 base case with 200,000 simulated individuals. Percents of the population by age group with one or more days (and six or more days) during the O_3 season with lung function (FEV_1) decrements more than 10, 15, and 20%. Minimum and maximum values and ranges over 40 APEX runs.

	$\Delta FEV_1 \geq 10\%$			$\Delta FEV_1 \geq 15\%$			$\Delta FEV_1 \geq 20\%$		
Age group	**min**	**max**	**range**	**min**	**max**	**range**	**min**	**max**	**range**
1 or more days in the season									
5 to 18	31.3%	32.1%	0.88%	12.4%	12.9%	0.49%	6.21%	6.71%	0.50%
19 to 35	11.1%	11.5%	0.39%	3.00%	3.26%	0.26%	1.11%	1.32%	0.22%
36 to 55	3.54%	3.79%	0.25%	0.55%	0.68%	0.13%	0.13%	0.20%	0.07%
6 or more days in the season									
5 to 18	9.28%	9.73%	0.45%	2.80%	3.18%	0.38%	1.11%	1.37%	0.27%
19 to 35	1.09%	1.25%	0.16%	0.15%	0.21%	0.06%	0.03%	0.06%	0.03%
36 to 55	0.22%	0.30%	0.08%	0.01%	0.03%	0.02%	0.00%	0.01%	0.01%

6.5.3 Application of Model for All Lifestages

The exposure-response functions derived from controlled human exposure studies involving 18-35 year old subjects were used to estimate responses for school-aged children (ages 5-18). This was in part justified by the findings of McDonnell et al. (1985) who reported that children 8-11 years old experienced FEV_1 responses similar to those observed in adults 18-35 years old when both groups were exposed to 120 ppb O_3 at an EVR of 32-35 L/min/m^2. In addition, a number of summer camp studies of school-aged children exposed in outdoor environments in the Northeast also showed O_3-induced lung function changes similar in magnitude to those observed in controlled human exposure studies using adults, although the studies may not directly comparable. The MSS model predicts increasing responsiveness with younger participants in the age range of 18-35 years, as shown in Figure 6E-4 (Appendix 6-E), which might indicate that responsiveness would continue to increase as age decreases from 18. In extending the MSS model to children, we fixed the age term in the model at its highest value, the value for age 18. If continuing the MSS model trend were to accurately describe continued increased response in children, then the fixed age term for children may have underestimated the effects on children, and particularly younger children. On the other hand, if FEV_1 responses for children are similar to those observed in adults 18-35 years old, as the evidence suggests, then our approach to extending the age term would overestimate the response to children (see Table 6E-3 in Appendix 6-E).

In considering extending the MSS model to ages older than 36, we note that, in general, O_3 responsiveness steadily declines for persons aged 35-55, with persons >55 eliciting minimal responsiveness (ISA, section 6.2.1.1). As described in Section 6.2.4, we extended the age term from the value at 36 linearly to zero at age 55, and set it to zero for ages above 55 (see **Error! Reference source not found.**). The uncertainty of this extrapolation may be substantial, but these age groups are not the primary focus in the clinical risk assessment.

6.5.4 Application of Model for Asthmatic Children

The risk assessment used the same exposure-response relationship, developed from data collected from healthy study subjects, and applied it to all persons, children, and asthmatic children. Based on limited evidence from a few human exposure studies, it is likely that subjects having asthma are at least as sensitive to acute effects of O_3 as other subjects not having this health condition (ISA, page 6-20 to 6-21). An analysis by Romieu et al. (2002) indicated a larger O_3-associated decrement in FEV_1 among children with moderate to severe asthma than among all children with asthma (ISA, page 6-54). This suggests that the lung function decrements presented in this assessment for asthmatic children may be underestimated. The magnitude of influence this element might have on our risk estimates remains unknown at this time. In addition, asthmatic children may have less reserve lung capacity to draw upon when faced with decrements, and therefore a ≥10% decrement in lung function may be a more adverse event in an asthmatic child than a healthy child.

6.5.5 Interaction Between O_3 and Other Pollutants

Because the controlled human exposure studies used in the risk assessment involved only O_3 exposures, it was assumed that estimates of O_3-induced health responses would not be affected by the presence of other pollutants (e.g., SO_2, $PM_{2.5}$, etc). The magnitude of influence that potential interactions might have on our risk estimates remains unknown at this time.

6.5.6 Qualitative Assessment of Uncertainty

EPA staff have identified key sources of uncertainty with respect to the lung function risk estimates. These are: the physiological model in APEX for ventilation rates, the O_3 exposures estimated by APEX, the MSS model applied to ages 18 to 35, and extrapolation of the MSS model to children ages 5 to 18. The first two of these are discussed in Chapter 5. At this time we do not have quantitative estimates of uncertainty for any of these. Table 6-16 provides a qualitative assessment of the uncertainty resulting from each of these key sources. The primary source of uncertainty is the MSS model, applied to ages 18 to 35.

Table 6-16. Summary of Qualitative Uncertainties of Key Modeling Elements in the O_3 Lung Function Risk Assessment

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments
		Direction	Magnitude		
The physiological model in APEX for ventilation rates	The physiological model in APEX takes into account the population distribution of individual physiological characteristics and activities and models minute-by-minute ventilation rates for each simulated individual using a series of physiological relationships known with varying degrees of certainty.	Over	Low-Medium	Low-Medium	Ventilation rates are a key input to the MSS model. Figure 6E-3 in Appendix 6-E gives an overview of the physiological model in APEX for ventilation rates. Comparisons with ventilation rates reported in the literature show fairly good agreement with APEX ventilation rates (Section 5.4.4).
O_3 exposures	The O_3 exposures estimated by APEX and their uncertainties are discussed in Chapter 5.	Both	Low-Medium	Low	O_3 exposures are a key input to the MSS model.
The McDonnell-Stewart-Smith (MSS) FEV_1 model for ages 18 to 35	The MSS model is integrated into APEX and predicts FEV_1 decrements for each simulated individual.	Both	Medium-High	Low	There is a good conceptual foundation for the structure of this model, but the variability in measurements of FEV_1 and estimated parameters of the model introduce uncertainty into the model predictions of large FEV_1 decrements. The estimated parameters have fairly wide confidence intervals (Table 6-1) and the risk results are sensitive to varying the parameters (Figure 6-12). The most influential parameter is β6, the power to which ventilation rate is raised in the MSS model. An increase of five percent in β6 leads to a 27 percent increase in the modeled number of children with FEV_1 decrements ≥ 10%. (The 95 percent confidence interval of this parameter estimate is ±14%.) The variability term ε [in equation 6-3] is assumed by the MSS model to have a Gaussian distribution with mean zero and estimated standard deviation 4.135. Since the actual values are bounded, we truncate the variability term distribution at ±2 standard deviations (±8.27), a convention we use for the distributions of several physiological

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments
		Direction	Magnitude		
					variables input to APEX in the physiology input file. To look at the effect of truncating the variability distribution, we conducted simulations with the variability term truncated at ±20, the range of the actual values of the variability term. We find that this constraint has a large effect on estimates of percents of the population with FEV_1 decrements ≥ 10 and 15% and less of an effect for 20%. The percent of children with FEV_1 decrements ≥ 10% increases from 31% to 92% when increasing the truncation point from 8.27 to 20.
Extrapolation of the MSS model to children	The MSS model is based on studies with subjects ranging in age from 18 to 35 years; therefore prediction for individuals outside this age range involves assumptions for extrapolation of the MSS model for individuals <18 and >35 years of age..	Both	Medium	Low	Summer camp studies and one clinical study of children indicate that FEV_1 responses for children are similar to those observed in adults 18-35 years old. See discussion in Section 6.5.3.

* Refers to the degree of uncertainty associated with our understanding of the phenomenon, in the context of assessing and characterizing its uncertainty

6.6 DISCUSSION

The second draft lung function risk assessment evaluated risks of lung function decrements due to O_3 exposure for all three groups: school-age children ages 5 to 18, young adults ages 19 to 35, and adults ages 36 to 55. Adults older than 55 have minimal O_3-induced lung function risk. Two models were used, one based on application of an individual level exposure-response function, the MSS model introduced in this review, and one based on application of a population level E-R function consistent with the model used in the previous O_3 review which applies probabilistic population-level exposure-response relationships for lung function decrements (measured as percent reductions in FEV_1) associated with 8-hour moderate exertion exposures. The MSS model is preferred, due to its ability to model individual exposures for a wide range of exposure times and levels of exercise (Section 6.2.4; ISA pages 6-15 to 6-16). Both models provide estimates of the percent of the groups experiencing a reduction in lung function for three different levels of impact, 10, 15, and 20% decrements in FEV_1. These levels of impact were selected based on the literature discussing the adversity associated with these types of lung function decrements (US EPA, 2012, Section 6.2.1.1; Henderson, 2006). For the second draft assessment, lung function risks were estimated for 15 cities: Atlanta, Baltimore, Boston, Chicago, Cleveland, Dallas, Denver, Detroit, Houston, Los Angeles, New York, Philadelphia, Sacramento, St. Louis, and Washington, DC.

Based on the MSS model, the percents of population estimated to experience lung function responses greater then 10, 15, and 20%, associated with O_3 exposure while engaged in various levels of exertion, vary considerably for different years and cities under the recent air quality scenarios and also for the existing and alternative standard scenarios (Figure 6-7 and Figure 6-8, Table 6-4 and Table 6-5). The estimates for $\geq$ 10% FEV_1 decrement for school-age children for recent O_3 concentrations range across cities and years from 11 to 31 percent, and range from 11 to 22 percent after simulating just meeting the existing standard. The estimates for $\geq$ 15% FEV_1 decrement for school-age children for recent O_3 concentrations range across cities and years from 2 to 12 percent, and range from 2 to 6 percent after simulating just meeting the existing standards. The estimates for $\geq$ 20% FEV_1 decrement for recent O_3 concentrations range across cities and years from 1 to 6 percent, and range from 1 to 3 percent after simulating just meeting the existing standards.

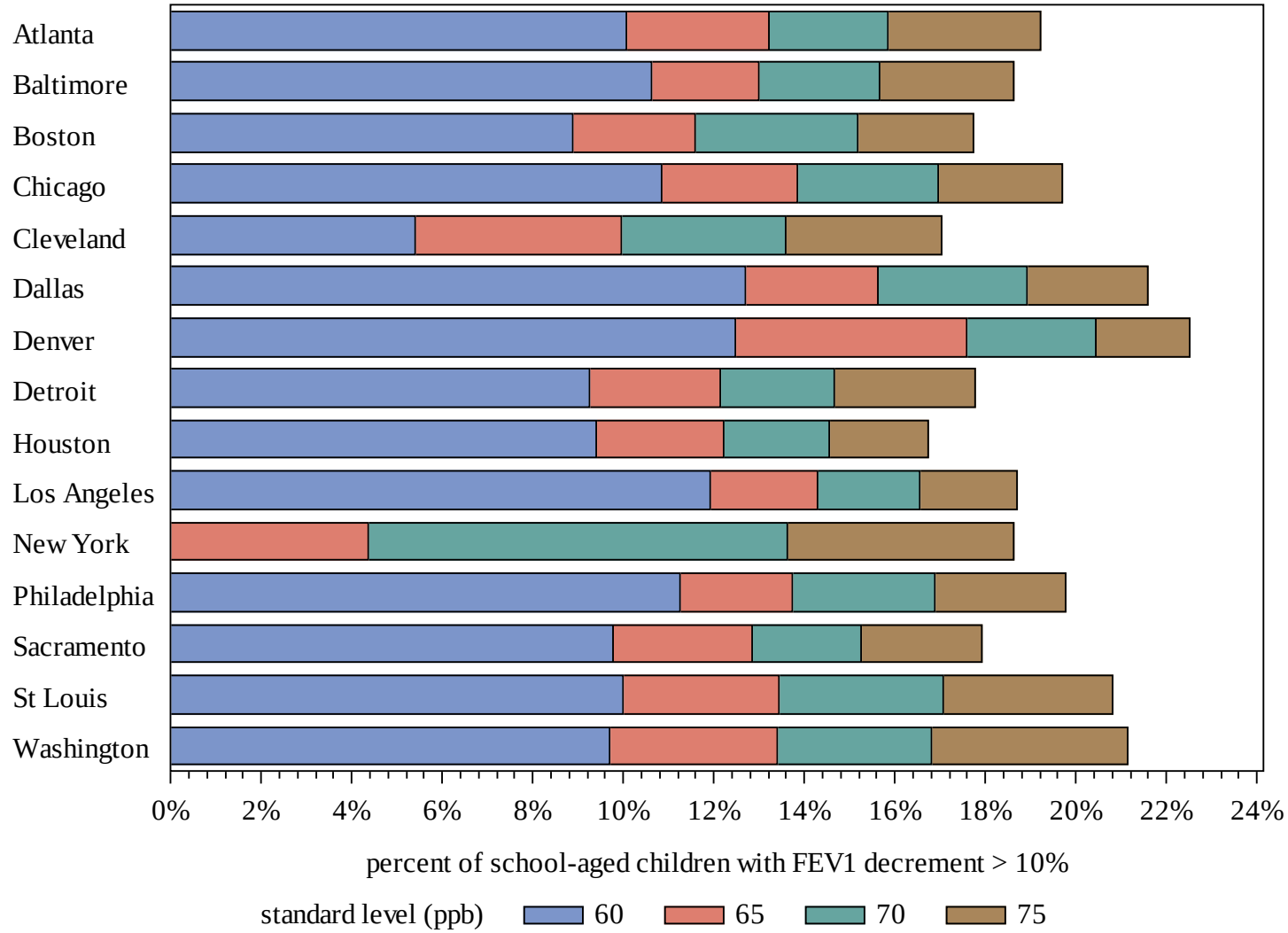
Figure 6-13 displays the risks and the incremental increases in risk for increasing standard levels, where risk is taken to be the highest value for each study area (over years) of the percent of school-aged children with FEV_1 decrement $\geq$ 10%. The risks in this figure for Washington, DC, for example, are about 9.6% for the alternative standard level of 60 ppb and 13.4% for the alternative standard level of 65 ppb. The length of the orange bar is the

incremental risk (3.8%) in going from the 60 ppb to the 65 ppb alternative standards. This figure shows that there are significant increases in incremental risk for all 15 cities in the progression of alternative standard levels from 60 ppb to the level of the existing standard, 75 ppb. The pattern of reductions for lung function decrements larger than 15 and 20% are similar. As discussed in Section 4.3.1, the New York 60 ppb alternative standard was not modeled and the risk for NY for that scenario would not necessarily be zero. Figure 6-14 displays the risks and the incremental increases in risk for increasing standard levels, where risk is taken to be the mean value for each study area (over years) of the percent of school-aged children with FEV_1 decrement $\geq$ 10%.

Similar to the MSS model results, the percents of school-age children estimated to experience lung function responses greater then 10, 15, and 20% based on the population level E-R function exhibit variation across years and cities. However, the MSS model estimates are significantly higher than the E-R approach estimates. For lung function responses greater than 10, 15, and 20% the MSS model gives results typically a factor of three higher than the E-R model for school-aged children. Both models give higher responses for higher concentrations, compared to lower concentrations, as can be seen in Figures 6-6, 6-9, and 6-10.

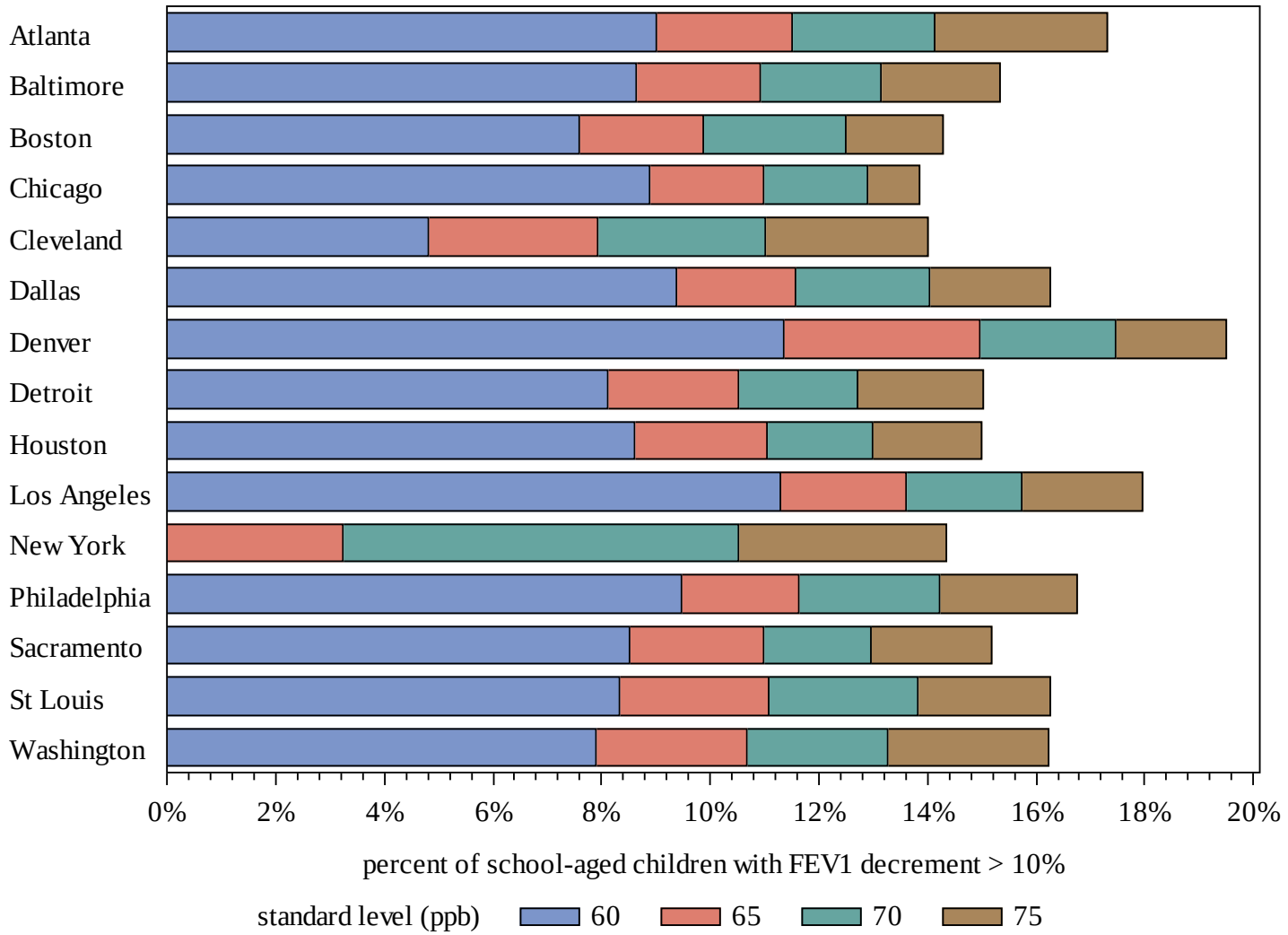
The MSS model was applied to estimate lung function risk for outdoor workers (ages 19-35) in Atlanta for one year (2006). The proportion of outdoor workers with FEV_1 decrements $\geq$ 15% ranges from 3.6 to 5.3 times the proportion of the general population (ages 19-35) with FEV_1 decrements $\geq$ 15% across the different standards simulated. The proportion of outdoor workers with multiple occurrences of FEV_1 decrements $\geq$ 15% is much greater than for the general population.

Figure 6-13. Lung Function Risk Results, Incremental Increases In Risk For Increasing Standard Levels: Percent of All School-aged Children With FEV_1 Decrement ≥ 10%, Highest Value For Each Study area Over Years[5]



[5] New York level 60 was not modeled . We do not know what the percent risk would be for NY under the 60 ppb alternative standard, but it would not necessarily be zero.

Figure 6-14. Lung Function Risk Results, Incremental Increases In Risk For Increasing Standard Levels: Percent of All School-aged Children With FEV_1 Decrement ≥ 10%, Mean Value For Each Study Area Over Years[6]



[6] New York level 60 was not modeled . We do not know what the percent risk would be for NY under the 60 ppb alternative standard, but it would not necessarily be zero.

6.7 REFERENCES

Adams, W.C. 2002. "Comparison of Chamber and Face-mask 6.6-hour Exposures to Ozone on Pulmonary Function and Symptoms Responses." *Inhalation Toxicology,* 14:745-764.

Adams, W.C. 2003. "Comparison of Chamber and Face Mask 6.6-hour Exposure to 0.08 ppm Ozone via Square-wave and Triangular Profiles on Pulmonary Responses." *Inhalation Toxicology*, 15: 265-281.

Adams, W.C. 2006. "Comparison of Chamber 6.6-h Exposures to 0.04-0.08 ppm Ozone via Square-wave and Triangular Profiles on Pulmonary Responses." *Inhalation Toxicology*, 18: 127-136.

American Thoracic Society. 2000. "What Constitutes an Adverse Effect of Air Pollution?" *American Journal of Respiratory and Critical Care Medicine,* 161: pp. 665-673.

Box, G. E. P. and G. C. Tiao. 1973. *Bayesian Inference in Statistical Analysis*. Wesley Publishing Co.: Wiley Classic Library, ISBN 0-471-57428-7.

Burmaster, D.E. 1998. "Lognormal Distributions for Skin Area as a Function of Body Weight." *Risk Analysis,* 18(1):27-32.

Dietert, R.R.; R. A. Etzel; D. Chen, et al. 2000. "Workshop to Identify Critical Windows of Exposure for Children's Health: Immune and Respiratory Systems Work Group Summary." *Environ Health Perspect*, 108(suppl 3): 483-490.

Esmail, S.; Y. Bhambhani and S. Brintnell. 1995. "Gender Differences in Work Performance on the Baltimore Therapeutic Equipment Work Simulator." *American Journal of Occupational Therapy,* 49: 405 - 411.

Federal Register. 1996. Responses to Significant Comments on the 1996 Proposed Rule on the National Ambient Air Quality Standards for Ozone (December 13, 1996; Fed. Reg. 61 FR 65716).

Folinsbee, L.J., et al. 1988. "Pulmonary Function and Symptom Responses after 6.6-hour Exposure to 0.12 ppm Ozone with Moderate Exercise." *Journal of the Air Pollution Control Association*, 38: 28-35.

Graham, S. and T. McCurdy. 2005. *Revised Ventilation Rate (Ve) Equations for Use in Inhalation-Oriented Exposure Models: A NERL Internal Research Report*. Research Triangle Park, NC: EPA National Exposure Research Laboratory.

Hazucha, M.J.; L. J. Folinsbee and P. A. Bromberg. 2003. "Distribution and Reproducibility of Spirometric Response to Ozone by Gender and Age." *Journal of Applied Physiology*, vol. 95, no. 5, pp. 1917-1925.

Hislop, A.A. 2002. "Airway and Blood Vessel Interaction During Lung Development." *Journal of Anatomy*, 201:325–334. Available at: <www.ncbi.nlm.nih.gov/pmc/articles/PMC1570917/>.

Horstman, D.H. et al. 1990. "Ozone Concentration and Pulmonary Response Relationships for 6.6-hour Exposures with Five Hours of Moderate Exercise to 0.08, 0.10, and 0.12 ppm." *American Review of Respiratory Disease,* 142:1158-1163.

Isaacs, K.; G. Glen; T. Mccurdy and L. Smith. 2008. "Modeling Energy Expenditure and Oxygen Consumption in Human Exposure Models: Accounting for Fatigue and EPOC." *Journal of Exposure Science and Environmental Epidemiology,* 18, 289–298.

Isaacs, K. and L. Smith. 2005. *New Values for Physiological Parameters for the Exposure Model Input File 'Physiology.txt.'* Technical memorandum to Tom McCurdy, EPA, December 20.

Kim, C.S.; N. E. Alexis; A. G. Rappold; H. Kehrl; M. J. Hazucha; J. C. Lay; M. T. Schmitt; M. Case; R. B. Devlin; D. B. Peden; D. Diaz-Sanchez. 2011. "Lung Function and Inflammatory Responses in Healthy Young Adults Exposed to 0.06 ppm Ozone for 6.6 hours." *American Journal of Respiratory and Critical Care Medicine,* 183: 1215-1221.

Linn, W.S.; R. D. Buckley; C. E. Spier; R. L. Blessey; M. P. Jones; D. A. Fischer and J. D. Hackney. 1978. "Health Effects of Ozone Exposure in Asthmatics." *The American Review of Respiratory Disease,* vol. 117, no. 5, pp. 835-843.

Lunn, D.; C. Jackson; N. Best; A. Thomas and D. Spiegelhalter. 2012. *The BUGS Book - A Practical Introduction to Bayesian Analysis*. CRC Press: Chapman and Hall.

McCurdy, T.; G. Glen; L. Smith; Y. Lakkadi. 2000. "The National Exposure Research Laboratory's Consolidated Human Activity Database." *Journal of Exposure Analysis and Environmental Epidemiology,*10: 566-578.

McDonnell, W.F.; R. S. Chapman; M. W. Leigh; G. L. Strope; A. M. Collier. 1985. "Respiratory Responses of Vigorously Exercising Children to 0.12 ppm Ozone Exposure." *American Review of Respiratory Disease,* 132: 875-879.

McDonnell, W. F.; H. R. Kehrl; S. Abdul-Salaam; P. J. Ives; L. J. Folinsbee; R. B. Devlin; J. J. O'Neil; D. H. Horstman. 1991. "Respiratory Response of Humans Exposed to Low Levels of Ozone for 6.6 hours." *Archives of Environmental Health,* 46(3):145-150.

McDonnell, W.F. et al. 1993. "Predictors of Individual Differences in Acute Response to Ozone Exposure." *American Review of Respiratory Disease,* 147:818-825.

McDonnell, W.F. et al. 1997. "Prediction of ozone-induced FEV_1 changes." *American Journal of Respiratory and Critical Care Medicine,* 156:715:722.

McDonnell, W.F.; P. W. Stewart and M. V. Smith. 2007. "The Temporal Dynamics of Ozone-induced FEV_1 Changes in Humans: An Exposure-response Model." *Inhalation Toxicology,* 19:483-494.

McDonnell, W.F.; P. W. Stewart and M. V. Smith. 2010. "Prediction of Ozone-induced Lung Function Responses in Humans." *Inhalation Toxicology,* 22(2):160-8.

McDonnell, W.F.; P. W. Stewart; M. V. Smith; C. S. Kim and E. S. Schelegle. 2012. "Prediction of Lung Function Response for Populations Exposed to a Wide Range of Ozone Conditions." *Inhalation Toxicology*, 24:619-633.

Narayanan, M.; J. Owers-Bradley; C. S. Beardsmore; M. Mada; I. Ball; R. Garipov; K. S. Panesar; C. E. Kuehni; B. D. Spycher; S. E. Williams; M. Silverman. 2012. "Alveolarization Continues During Childhood and Adolescence: New Evidence from Helium-3 Magnetic Resonance." *American Journal of Respiratory and Critical Care Medicine*, January 15; 185(2): 86–191. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3410735/>.

Passannante, A. N.; M. J. Hazucha; P. A. Bromberg; E. Seal; L. Folinsbee; G. Koch. 1998. "Nociceptive Mechanisms Modulate Ozone-induced Human Lung Function Decrements." *Journal of Applied Physiology*, 85:1863-1870.

Romieu, I; J. J. Sienra-Monge; M. Ramirez-Aguilar; M. M. Tellez-Rojo; H. Moreno-Macias; N. I. Reyes-Ruiz; B. E. Del Rio-Navarro; M. X. Ruiz-Navarro; G. Hatch; R. Slade; M. Hernandez-Avila. 2002. "Antioxidant Supplementation and Lung Functions Among Children with Asthma Exposed to High Levels of Air Pollutants." *American Journal of Respiratory and Critical Care Medicine*, 166: 703-709.

Samet, J.M. 2011. "Clean Air Scientific Advisory Committee (CASAC) Response to Charge Questions on the Reconsideration of the 2008 Ozone National Ambient Air Quality Standards." (Report number EPA-CASAC-11-004), dated March 30, 2011. Available at: <http://yosemite.epa.gov/sab/sabproduct.nsf/0/F08BEB48C1139E2A8525785E006909AC/$File/EPA-CASAC-11-004-unsigned+.pdf>.

Schelegle, E.S.; C. A. Morales; W. F. Walby; S. Marion; R. P. Allen. 2009. "6.6-Hour Inhalation of Ozone Concentrations from 60 to 87 ppb in Healthy Humans." *American Journal of Respiratory and Critical Care Medicine*, 180:265-272.

Schelegle, E.S.; W. C. Adams; W. F. Walby and S. Marion. (2012). "Modeling of Individual Subject Ozone Exposure Response Kinetics." *Inhalation Toxicology*, 24(7):401-415.

Schofield, W.N. 1985. "Predicting Basal Metabolic Rate, New Standards and Review of Previous Work." *Clinical Nutrition*, 1985; 39C (suppl): 5–41.

Seal, Jr., E.; W. F. McDonnell; D. E. House. 1996. "Effects of Age, Socioeconomic Status, and Menstrual Cycle on Pulmonary Response to Ozone." *Archives of Environmental and Occupational Health*, 1996;51:132-137.

U.S. Environmental Protection Agency. 1996. *Review of National Ambient Air Quality Standards for Ozone: Assessment of Scientific and Technical Information - OAQPS Staff Paper*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA/452/R-96-007). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_pr.html>.

U.S. EPA. 2007a. *Review of National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information - OAQPS Staff Paper*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-07-007). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_cr_sp.html>.

U.S. EPA. 2007b. *Ozone Population Health Risk Assessment for Selected Urban Areas*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/R-07-009). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_cr_td.html>.

U.S. EPA. 2009. *Metabolically Derived Human Ventilation Rates: A Revised Approach Based Upon Oxygen Consumption Rates*. Washington, DC: EPA National Center for Environmental Assessment. (EPA document number EPA/600/R-06/129F). Available at: <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=202543>.

U.S. EPA. 2011. *O_3 National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA-452/P-11-001). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_2008_pd.html>.

U.S. EPA. 2012b. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume I: User's Guide*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards(EPA document number EPA-452/B-12-001a). Available at: <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2012c. *Total Risk Integrated Methodology (TRIM) - Air Pollutants Exposure Model Documentation (TRIM.Expo / APEX, Version 4.4) Volume II: Technical Support Document*. Research Triangle Park, NC. (EPA document number EPA-452/B-12-001b). Available at: <http://www.epa.gov/ttn/fera/human_apex.html>.

U.S. EPA. 2013a. *Integrated Science Assessment of Ozone and Related Photochemical Oxidants*. EPA National Center for Environmental Assessment. (EPA document number EPA/600/R-10/076F, 2013). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/s_o3_2008_isa.html >.

WinBUGS, version 1.4.3. Available at: <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>.

7 CHARACTERIZATION OF HEALTH RISK BASED ON EPIDEMIOLOGICAL STUDIES

This chapter provides an overview of the methods used to estimate health risks in selected urban areas based on application of results of epidemiology studies. Section 7.1.1 discusses the basic structure of the risk assessment, identifying the modeling elements and related sources of input data needed for the analysis and presenting an overview of the approach used in calculating health effect incidence using concentration-response (C-R) functions based on epidemiological studies. Section 7.2 discusses air quality considerations. Section 7.3 discusses the selection of model inputs including: (a) selection of urban study areas, (b) selection of epidemiological studies and specification of C-R functions, (c) specification of baseline health effect incidence and prevalence rates, and (d) estimation of population (demographic) counts. Section 7.4 describes how uncertainty and variability are addressed in the risk assessment, including specification of the sensitivity analyses completed for the risk assessment and how these differ from the core risk estimates. Section 7.5 summarizes the risk estimates that are generated, including both the core estimates and sensitivity analyses. Finally, Section 7.6 provides an assessment of overall confidence in the risk assessment together with a set of key observations regarding the risk estimates generated.

7.1 GENERAL APPROACH

7.1.1 Basic Structure of the Risk Assessment

This risk assessment involves the estimation of the incidence of specific health effect endpoints associated with exposure to ambient O_3 for defined populations located within a set of urban study areas. Because the risk assessment focuses on health effect incidence experienced by defined populations, it represents a form of population-level risk assessment and does not estimate risks to individuals within the population. Furthermore, because it models risk for residents in a set of urban study areas, it is not intended to provide an estimate of national-level risk[1] .

The general approach used in both the prior and current O_3 risk assessments relies on C-R functions based on effect estimates and model specifications obtained from epidemiological studies. Since these studies derive effect estimates and model specifications using averages of ambient air quality data from fixed-site, population-oriented monitors, uncertainty arising from

[1] Chapter 8 provides a limited assessment of national risk focused on the mortality burden associated with recent O_3 levels. This risk and exposure assessment does not provide an analysis of the risk reductions that would be expected for the entire U.S. after meeting either the existing or alternative standards.

the application of these functions in an O_3 risk assessment is decreased if, in modeling risk, we also use ambient air quality data at fixed-site, population-oriented monitors to characterize exposure. Therefore, we developed a composite monitor for each urban study area to represent a surrogate population exposure by averaging O_3 concentrations across the monitors in that study area to produce a single composite hourly time series of values. The O_3 metrics used in evaluating risk are derived from the composite monitor hourly time series distribution (see sections 7.2 and Chapter 4 for additional detail on the characterization of ambient O_3 levels).[2]

The general O_3 health risk model, illustrated in Figure 7-1, combines O_3 air quality data, C-R functions, baseline health incidence and prevalence data, and population data (all specific to a given urban study area) to derive estimates of the annual incidence of specified health effects for that urban study area attributable to O_3 exposure. This risk assessment models risk for 12 urban study areas we selected to provide coverage for the types of urban O_3 scenarios likely to exist across the U.S. (see section 7.3.1). Chapter 8 provides an assessment of the degree to which the 12 selected urban areas are representative of other urban areas in the U.S. that are likely to experience elevated risks from exposure to ambient O_3 under recent conditions.

This risk assessment provides an updated set of estimates for risk under recent O_3 conditions and just meeting the existing standard, and additional estimates of risk if alternative standards are just met, with an emphasis on reductions in risk between just meeting the existing standard and just meeting alternative standards (the full set of risk estimates, including simulation of risk under current conditions is presented in Appendix 7-B). The alternative standard levels evaluated are 70, 65 and 60 ppb (expressed using the current form of the O_3 standard).

We simulated just meeting the existing and alternative O_3 standards by adjusting hourly O_3 concentrations measured over the O_3 season using a model-based adjustment methodology that estimates O_3 sensitivities to precursor emissions changes.[3] These sensitivities, which estimate the response of O_3 concentrations to reductions in anthropogenic NOx and VOC emissions, are developed using the Higher-order Decoupled Direct Method (HDDM) capabilities in the Community Multi-scale Air Quality (CMAQ) model. More details on the HDDM-adjustment approach is presented in Chapter 4 of this REA and in Simon et al. (2013).

As discussed in Chapters 2 and 3, in modeling risk we employ continuous non-threshold C-R functions relating O_3 exposure to health effect incidence. The use of non-threshold

[2] This holds for all air quality metrics used in modeling short-term mortality and morbidity endpoints. However, the air metric used in modeling long-term mortality is based on a seasonal average of maximum hourly values derived for each O_3 monitor within an urban study area with those individual averages then combined to generate a single seasonal average composite monitor value for each study area (see section 7.2 for more detail).

[3] In the first draft of this REA, we used a statistical quadratic rollback approach to simulate just meeting the existing O_3 standards. In that first draft, we proposed using the model based approach used in this draft, and received support for the model based approach from CASAC (Frey, H.D., 2012).

functions reflects the discussion of the relevant studies in the O_3 ISA (see O_3 ISA, section 2.5.4.4, U.S. EPA 2013a). However, also consistent with the conclusions of the O_3 ISA, we recognize that the evidence from the studies indicates less confidence in specifying the shape of the C-R function at O_3 concentrations towards the lower end of the distribution of data used in fitting the curve due to the reduction in the number of data points available. The O_3 ISA noted that the studies indicate reduced certainty in specifying the shape of the C-R function specifically for short-term O_3-attributable respiratory morbidity and mortality, in the range generally below 20 ppb (for both 8hr-maximum and 24hr metrics) (O_3 ISA, section 2.5.4.4). However, care needs to be taken in interpreting this range of reduced confidence indicated in the studies and applying it to the interpretation of risk estimates generated for a specific urban study area. This is because there is considerable heterogeneity in the effect of O_3 on mortality across urban study areas (O_3 ISA section 6.6.2.3). Additionally, it is likely that levels of confidence associated with C-R functions (including ranges of reduced confidence in specifying the function) also vary across urban study areas reflecting underlying differences in factors impacting the exposure-response relationship for O_3, such as demographic differences and exposure measurement error. For these reasons, the ≤20 ppb range discussed in the O_3 ISA should be viewed as a more generalized range to be considered qualitatively or semi-quantitatively, along with many other factors, when interpreting the risk estimates rather than as a fixed, bright-line.[4]

Based on comments we received from CASAC on the 1st draft REA, we are no longer including estimates of risk down to the lowest measured level (LML).[5] Instead, through the use of heat map tables, we focus on providing estimates of total risk, and the distribution of risk over concentrations of O_3.[6] Coupled with information about what the studies indicate about the C-R function at lower O_3 concentrations, this provides for a more complete understanding of confidence in estimated risk than simply truncating risk at the LML.

In modeling risk for all health endpoints included in the analysis, for recent O_3 conditions and just meeting the existing standard, we estimated total risk (down to zero). For meeting the

[4] This range of reduced confidence in the shape of the C-R function is most appropriately applied to area-wide averages (i.e., composite monitor values) of the type often used in epidemiological studies rather than to the range of O_3 associated with a particular monitor. This reflects the fact that the observations presented in the O_3 ISA are themselves based on consideration for epidemiological studies which use composite monitor values.

[5] Based on their November 19, 2012 letter commenting on the 1st draft REA, CASACrecommended against inclusion of risk estimates based on the LML in the core analysis due to the fact that there is little difference between these estimates and risk estimates based on total O_3 exposure and that LML information is not available for many of the epidemiological studies used in the REA (Frey and Samet, 2012). However, they recommend a more limited exploration of the LML and its implications for risk for one or more areas. In response, we have included coverage for LML as part of our discussion of the heat maps results presented in section 7.5.1 .

[6] Heat map tables illustrate the distribution of estimated O_3-related deaths across daily O_3 levels for each urban study and allow a quick visual comparison of trends (in the distribution of total O_3 risk as we as risk reductions) across ambient O_3 ranges both within and across study areas (see section 7.5).

existing and alternative standards, we estimated both total risk as well as the difference in risk, representing the degree of risk reduction associated with just meeting the existing and alternative standard levels. When calculating risk differences, we focus on comparing total risk after just meeting each alternative standard with total risk after just meeting the existing standards. We also evaluate the incremental change in risk from meeting increasingly lower alternative standard levels. Risk results are presented in terms of absolute numbers and changes in the O_3 attributable incidence of mortality and morbidity, and in terms of the percent of baseline mortality and morbidity attributable to O_3. We also provide risks per 100,000 population (to normalize risks across urban areas with different size populations to facilitate comparisons).

As with previous NAAQS-related risk assessments, for this analysis we have generated two categories of risk estimates, including a set of core (or primary) estimates and an additional set of sensitivity analyses. The core risk estimates utilize C-R functions based on epidemiological studies for which we have relatively greater overall confidence and which provide the best coverage for the broader O_3 monitoring period (rather than focusing only on the summer season). Although it is not strictly possible to assign quantitative levels of confidence to these core risk estimates due to data limitations, they are generally based on inputs having higher overall levels of confidence relative to risk estimates that are generated using other C-R functions. Therefore, emphasis is placed on the core risk estimates in making observations regarding total risk and risk reductions associated with recent conditions and after just meeting the existing and alternative standard levels. By contrast, the sensitivity analysis results typically reflect application of C-R functions covering a wider array of design elements which can impact risk (e.g., length of season, copollutants models, lag structures, statistical modeling methods etc). The sensitivity analysis results provide insights into the potential impact of these design elements on the core risk estimates, thereby informing our characterization of overall confidence in the core risk estimates.[7] We have significantly expanded our sensitivity analysis relative to that completed for the 1st draft REA to address a wider range of modeling elements which can impact the core risk estimates. Details of the design of the core and sensitivity analyses (including modeling element composition) for each of the health effect endpoints categories covered in this risk assessment are presented in section 7.4.3 and briefly summarized below.

For short-term exposure related mortality, our core analysis is based on application of C-R functions obtained from the Smith et al., 2009 epidemiological study (see section 7.3.2). In

[7] In presenting both the core and sensitivity analysis, we include both point estimates and 95th percentile confidence intervals (CIs). The 95th percentile CIs reflect the statistical fit of the underlying effect estimates and therefore reflect the statistical power of the epidemiological studies supplying the effect estimates. Often in comparing sensitivity analysis with core risk estimates, we focus not only on the point estimates, but also on the confidence intervals since these inform our understanding of confidence in the respective risk estimates.

addition, we have completed an expanded array of sensitivity analyses which provide coverage for a number of modeling elements including: (a) time period reflected in risk modeling (summer season versus full monitoring period), (b) peak O_3 metric (8hr maximum versus 8hr mean) (c) use of regional versus national-based Bayesian adjustment in deriving effect estimates,[8] (d) use of single (O_3-only) versus copollutant (O_3 and PM_{10}) models, [9] (e) application of alternative C-R functions based on Zanobetti and Schwartz, 2008 (see section 7.3.2) and (f) size of the urban study area (CBSA versus smaller multi-county study area)[10] (see sections 7.4.3 and 7.5.3 for additional detail on the sensitivity analyses completed). In addition to these sensitivity analyses, we have considered alternative methods for adjusting air quality to attain existing and alternative standards (NOx-only versus combination of VOC and NOx reductions). Additional sensitivity analyses exploring lag structure may also provide useful information, but are not possible due to the lack of availability of Bayes adjusted estimates for alternative lag structures.

For short-term exposure morbidity, we have effect estimates covering a wide range of design elements including co-/single-pollutant models and lag structure. However, we were not in a position to differentiate between these alternative model forms in terms of overall confidence and have therefore included all of these estimates in the core analysis. This range of risk estimates can also be viewed as a sensitivity analysis where there is no clear "core" estimate and instead, the full range of risk estimates is considered to provide the best overall picture of risk for a specific endpoint (see section 7.3.2 and 7.4.3).

Our analysis also includes estimates of long-term exposure related respiratory mortality, including a core estimate based on a co-pollutant model (with $PM_{2.5}$) together with sensitivity analyses exploring regional heterogeneity in the effect estimate and application of a national-

[8] Short-term O_3-attributable mortality in this analysis is modeled using Bayesian-adjusted effect estimates. This approach involves adjustment of each city's effect estimate using a prior distribution reflecting the O_3-mortality relationship seen across the broader set of cities considered in the epidemiological study. For the sensitivity analysis, we compare the use of a national prior distribution (the core approach) for the Bayesian adjustment with use of a regional prior.

[9] The copollutants model results are limited by the reduced number of days with copollutants sampling (either 1 in 3 or 1 in 6) which makes it difficult to evaluate the statistical significance of these results in view of the large posterior standard deviations (Smith et al., 2009). This increased uncertainty associated with the estimates prevents these results from being treated as part of the core analysis. Never the less, they provide perspective on the potential magnitude of risk associated with copollutants modeling and as such make an important contribution as a sensitivity analysis for short-term O_3-attributable mortality.

[10] Core based statistical areas (CBSAs) are U.S. geographic areas defined by the Office of Management and Budget (OMB). They include an urban center of at least 10,000 people combined with adjacent urban and surburban areas that are socioeconomically tied to the urban center by commuting. CBSAs tend to be significantly larger than the study areas used in the epidemiological studies providing effect estimates. We have used risk estimates based on CBSAs in the core analysis in order to better represent the changes in risk that could be experienced in the broader urban areas and to avoid the introduction of known bias into the risk assessment. We have included risk estimates based on the smaller study areas from the original epidemiological studies as sensitivity analyses (see discussion later in this section for additional detail).

level estimates focusing only on O_3 (see section 7.5.3).[11] The decision to model this endpoint is based on our evaluation of the evidence as summarized in the O_3 ISA and comments received from CASAC based on the 1st draft risk assessment (Frey H.D., 2012 p.).

As noted earlier, for this draft, we have modeled all core risk estimates using study areas based on the core-based statistical area (CBSA) regardless of whether the epidemiological studies providing the effect estimates used the CBSA spatial definition or a different spatial study area definition. The decision to use CBSA-based study areas in all core simulations for this draft reflects our desire to better represent the changes in risk that could be experienced in the urban areas and avoid introducing substantial known bias into the risk estimates. As discussed in Chapter 4 (section 4.3.1.2), most nonattaining O_3 monitors are not located in the center of the urban area, but instead in the surrounding areas, reflecting the transport and atmospheric chemistry governing O_3 formation. The monitors in the urban core areas are usually most affected by local sources of NOx and experience lower concentrations of O_3 since the NO is titrating the O_3 in these areas. For these monitors, simulating attainment of the existing and alternative standard levels can result in an increase in O_3 concentrations, while areas further out from the core experience the expected reduction in O_3 level . Had we focused risk estimates on the smaller urban core areas used in some of the epidemiological studies, we would not have fully captured the changes in risk estimated to be experienced by the broader urban area since we would have been focusing only on those areas experiencing net increases in O_3 (when simulating attainment of the existing and alternative standard levels). By modeling risk for the core analysis using the more inclusive CBSA study areas, we insure that risk estimates will include consideration both for the relatively smaller core urban areas experiencing increases in O_3 as well as the broader urban and suburban area experiencing risk reductions. We will also insure that, to a greater extent, the analysis includes the county with the design value monitor in the assessment of risk (see section 7.2).

There is a degree of uncertainty introduced through application of effect estimates to study areas (i.e., CBSAs) that do not match those used in the underlying epidemiological studies. This uncertainty should be viewed within the context of the overall larger uncertainty associated with transferring effect estimates from the context of the epidemiological studies to the context of the risk assessment. The epidemiological studies used in modeling short-term exposure-related endpoints generate effect estimates based on day to day variation in O_3 and health effects, using the area wide average O_3 concentrations. Area wide O_3 averaging masks the specific population distribution of O_3 exposures which reflects the times and durations of exposures to O_3 measured at individual monitors in an urban area. We apply those effect estimates to the air quality

[11] The seasonal average metrics used in the long-term mortality estimate are not very sensitive to the reduced number of days with co-pollutant monitoring, and as such it is appropriate to use the co-pollutant model in generating the core risk estimates.

scenarios of just meeting existing and alternative standards, where we are shifting the entire distribution of daily O_3 concentrations, and altering the relationships between O_3 concentrations at different monitors, and thus likely altering the relationship between area wide average O_3 and the population distribution of O_3 exposures. By doing so, we introduce an additional source of exposure measurement error, which goes beyond the impact that measurement error has on the effect estimate, and introduces additional uncertainty into the estimates of risk associated with simulating meeting existing and alternative standards.

Our decision to use the CBSA to define the spatial extent of each urban study area reflects the greater weight we place on minimizing biases relative to minimizing uncertainty, although we strive to minimize both where possible. The sensitivity analysis related to using study-based spatial definitions for urban areas shows clearly that using the smaller urban areas biases downward the risk reductions across an urban area. Thus, to avoid this bias in risk estimates we accept a measure of increased uncertainty associated with the application of effect estimates to study areas that are larger than those used in some of the original epidemiological studies providing those effect estimates.

Using the CBSA definitions of urban areas can partially address the bias caused by focusing only on urban core areas. However, it does not address this bias fully in some areas because of the unevenness in monitoring throughout urban areas. In some urban areas the monitors are more evenly distributed across the CBSA, while in other areas they are not. For example, in some urban areas, there is a high density of monitors in the urban core counties, with less density of monitors in surrounding counties also in the CBSA. Because we use a simple average (to match the averaging used in the epidemiology studies) of monitors across the CBSA, this means that O_3 concentrations in areas where there are more monitors (e.g. in urban core counties) will get a higher weight in the average O_3 concentrations relative to O_3 concentrations in other parts of the CBSA. To the extent that the area with the higher density of monitors experiences increases in O_3 while the remaining area experiences decreases in O_3, the overall average O_3 concentrations applied to populations in the entire CBSA will be weighted more towards O_3 increases, which will attenuate the overall risk reduction that may be associated with meeting alternative O_3 standards. We are not able to determine the magnitude of this remaining bias; however, it is expected to be higher in locations with a high percentage of total CBSA monitors concentrated in urban core counties.

The risk assessment reflects consideration for five years of recent air quality data from 2006 through 2010, with these five years reflecting two three-year attainment simulation periods that share a common overlapping year (i.e., 2006-2008 and 2008-2010 - see section 7.2). We selected these two attainment simulation periods to provide coverage for a more recent time period with relatively elevated O_3 levels (2006-2008) and recent time period with relatively

lower O_3 levels (2008-2010). For the REA, we model risk for the middle year of each three-year attainment simulation period in order to provide estimates of risk for a year with generally higher O_3 levels (2007) and a year with generally lower O_3 levels (2009). In modeling risk, we matched the population data used in the risk assessment to the year of the air quality data. For example, when we used 2007 air quality data, we used 2007 population estimates. For baseline incidence and prevalence, rather than interpolating rates for the two specific years modeled in the risk assessment, we selected the closest year for which we had existing incidence/prevalence data (i.e., for simulation year 2007, we used available data for 2005 and for simulation year 2009, we used data from 2010). The calculation of baseline incidence and prevalence rates is described in section 7.3.4.

The risk assessment procedures described in more detail below are diagramed in Figure 7-1. To estimate the change in incidence of a given health effect resulting from a given change in ambient O_3 concentrations in an assessment location, the following analysis inputs are necessary:

- **Air quality information including:** (1) O_3 air quality data from each of the simulation years included in the analysis (2007 and 2009) from population-oriented monitors in the assessment location (these are aggregated to form composite monitor values used to represent population exposure), and (2) a method for adjusting the air quality data to simulate just meeting the current or alternative suite of O_3 standards. (These air quality inputs are discussed in more detail in Chapter 4).
- **C-R function(s**): which provide an estimate of the relationship between the health endpoint of interest and O_3 concentrations (for this analysis, C-R functions used were applied to urban study areas matching the assessment locations from the epidemiological studies used in deriving the functions, in order to increase overall confidence in the risk estimates generated - see section 7.3.2). For O_3, epidemiological studies providing information necessary to specify C-R functions are readily available for O_3-related health effects associated with short-term exposures (Section 7.1.2 describes the role of C-R functions in estimating health risks associated with O_3). In addition, the Jerrett et al. (2009) study provided a C-R function for modeling mortality risks associated with longer-term exposures to O_3.
- **Population information (*baseline health affects incidence and prevalence rates and population*):** The baseline incidence provides an estimate of the incidence rate (number of cases of the health effect per year or day, depending on endpoint, usually per 10,000 or 100,000 general population) in the assessment location corresponding to recent ambient O_3 levels in that location. The baseline prevalence rate describes the prevalence of a given disease state or conditions (e.g., asthma) within the population

(number of individuals with the disease state/condition, usually per 10,000 or 100,000 general population). To derive the total baseline incidence or prevalence per year, this rate must be multiplied by the corresponding population number (e.g., if the baseline incidence rate is number of cases per year per 100,000 population, it must be multiplied by the number of 100,000s in the population) (Section 7.3.4 summarizes considerations related to the baseline incidence and prevalence rates and population data inputs to the risk assessment).

In addition to the inputs described above, it is also necessary to specify the spatial extent of the study areas that will be modeled. These study areas definitions determine the composition of (a) the composite monitor values (which specific set of monitors are used in constructing the composite monitor, reflecting the area-wide average across monitors for each study area), (b) the specific set of effect estimates that will be used (matching the study areas to the specific set of effect estimates in the epidemiological studies being used to support modeling of endpoints), (c) the baseline incidence data and (d) the population demographic (count) data for each study area. As mentioned earlier, for this REA we have modeled 12 urban study areas and have used the CBSA spatial definition to specify the extent of each of these urban areas (see section 7.3.1 for additional details on study area selection).

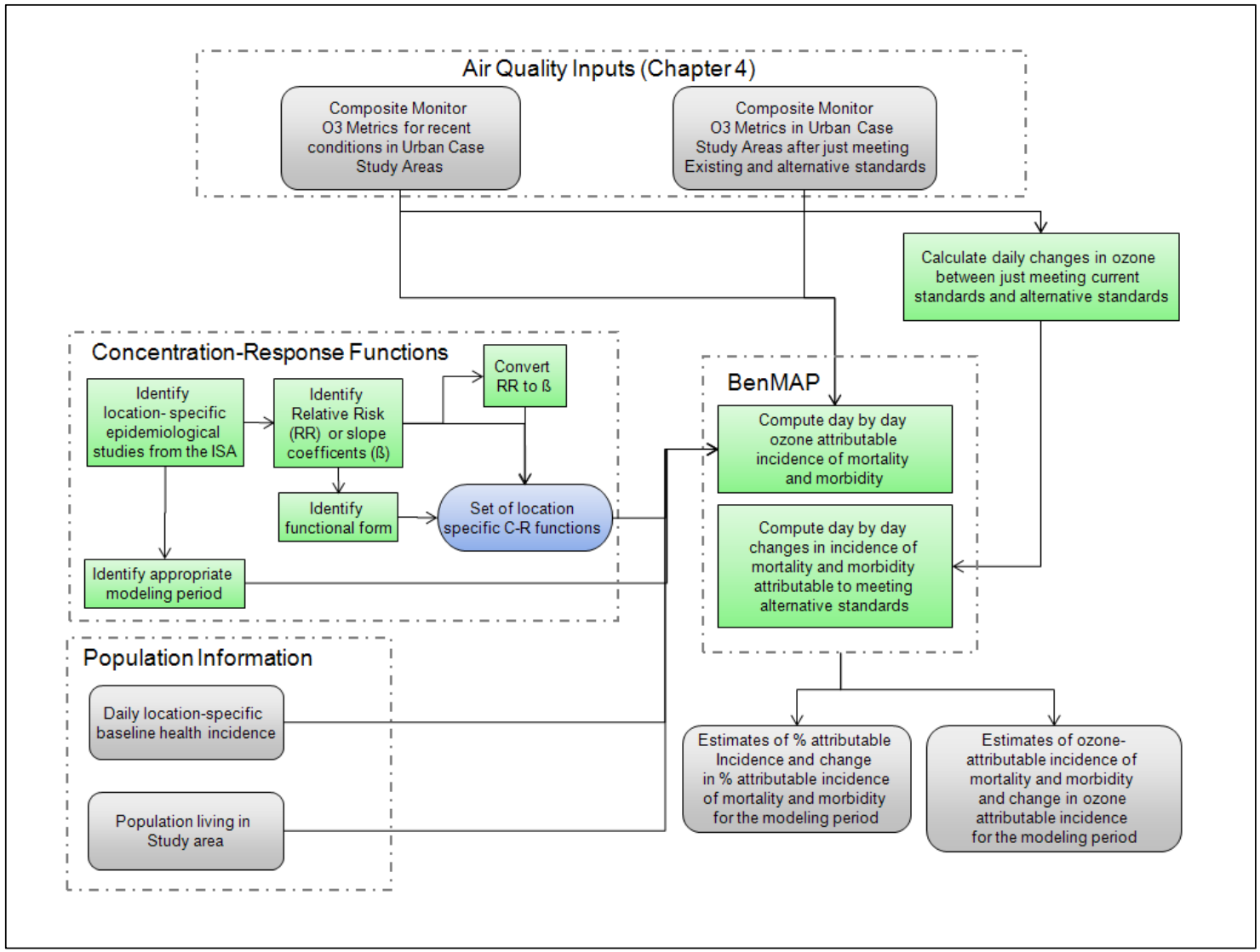


Figure 7-1 Flow Diagram of Risk Assessment for Short-term Exposure Studies

This risk assessment was implemented using the EPA's Environmental Benefits Mapping and Analysis Program—Community Edition, Version 0.63 (BenMAP-CE) (U.S. EPA, 2013b). This GIS-based computer program draws upon a database of population, baseline incidence/prevalence rates and effect coefficients to automate the calculation of health impacts. For this analysis, the standard set of effect coefficients and health effect incidence data available in BenMAP has been augmented to reflect the latest studies and data available for modeling O_3 risk. EPA has traditionally relied upon the BenMAP program to estimate the health impacts avoided and economic benefits associated with adopting new air quality rules. For this analysis, EPA used the model to estimate O_3-related risk for the suite of health effects endpoints described in section 3.2. There are three primary advantages to using BenMAP for this analysis, as compared to the procedure for estimating population risk followed in the last review. First, once we have configured the BenMAP software for this particular O_3 analysis, the program can produce risk estimates for an array of modeling scenarios across a large number of urban areas. Second, the program can more easily accommodate a variety of sensitivity analyses. Third, BenMAP allowed us to complete the national assessment of O_3 mortality described in Chapter 8, which plays in important role in assessing the representativeness of the urban study area analysis.

7.1.2 Calculating O_3-Related Health Effects Incidence

The C-R functions used in the risk assessment are empirically estimated associations between average ambient concentrations of O_3 and the health endpoints of interest (e.g., mortality, hospital admissions, emergency department visits). This section describes the basic method used to estimate changes in the incidence of a health endpoint associated with changes in O_3, using a "generic" C-R function of the most common functional form.

Although some epidemiological studies have estimated linear C-R functions and some have estimated logistic functions, most of the studies used a method referred to as "Poisson regression" to estimate exponential (or log-linear) C-R functions in which the natural logarithm of the health endpoint is a linear function of O_3:

$$y = Be^{\beta x} \quad (1)$$

where x is the ambient O_3 level, y is the incidence of the health endpoint of interest at O_3 level x, β is the coefficient relating ambient O_3 concentration to the health endpoint, and B is the incidence at x=0, i.e., when there is no ambient O_3. The relationship between a specified ambient O_3 level, x_0, for example, and the incidence of a given health endpoint associated with that level (denoted as y_0) is then

$$y_0 = Be^{\beta x_0} \quad (2)$$

Because the log-linear form of a C-R function (equation (1) is by far the most common form, we use this form to illustrate the "health impact function" used in the O_3 risk assessment.

If we let x_0 denote the baseline (upper) O_3 level, and x_1 denote the lower O_3 level, and y_0 and y_1 denote the corresponding incidences of the health effect, we can derive the following relationship between the change in x, $\Delta x = (x_0 - x_1)$, and the corresponding change in y, Δy, from equation (1).[12]

$$\Delta y = (y_0 - y_1) = y_0[1 - e^{-\beta \Delta x}]. \quad (3)$$

Alternatively, the difference in health effects incidence can be calculated indirectly using relative risk. Relative risk (RR) is a measure commonly used by epidemiologists to characterize the comparative health effects associated with a particular air quality comparison. The risk of mortality at ambient O_3 level x_0 relative to the risk of mortality at ambient O_3 level x_1, for example, may be characterized by the ratio of the two mortality rates: the mortality rate among individuals when the ambient O_3 level is x_0 and the mortality rate among (otherwise identical) individuals when the ambient O_3 level is x_1. This is the RR for mortality associated with the difference between the two ambient O_3 levels, x_0 and x_1. Given a C-R function of the form shown in equation (1) and a particular difference in ambient O_3 levels, Δx, the RR associated with that difference in ambient O_3, denoted as $RR_{\Delta x}$, is equal to $e^{\beta \Delta x}$. The difference in health effects incidence, Δy, corresponding to a given difference in ambient O_3 levels, Δx, can then be calculated based on this $RR_{\Delta x}$ as:

$$\Delta y = (y_0 - y_1) = y_0[1 - (1/RR_{\Delta x})]. \quad (4)$$

Equations (3) and (4) are simply alternative ways of expressing the relationship between a given difference in ambient O_3 levels, $\Delta x > 0$, and the corresponding difference in health effects incidence, Δy.[13] These health impact equations are the key equations that combine air quality information, C-R function information, and baseline health effects incidence information to estimate ambient O_3 health risk.

[12] If $\Delta x < 0$ – i.e., if $\Delta x = (x_1 - x_0)$ – then the relationship between Δx and Δy can be shown to be $\Delta y = (y_1 - y_0) = y_0[e^{\beta \Delta x} - 1]$. If $\Delta x < 0$, Δy will similarly be negative. However, the *magnitude* of Δy will be the same whether $\Delta x > 0$ or $\Delta x < 0$ – i.e., the absolute value of Δy does not depend on which equation is used.

[13] When calculating total risk associated with a specific air quality scenario, Δx is the total O_3 concentration associated with a given study area (as noted earlier in section 7.1.1, we are not incorporating thresholds, such as LMLs into this analysis).

7.2 AIR QUALITY CONSIDERATIONS

Air quality data are discussed in detail in Chapter 4 of this report. Here we describe those air quality considerations that are directly relevant to the estimation of health risks in the epidemiology based portion of the risk assessment. As described in section 7.1.1, the risk assessment uses composite (area-wide average) monitor values derived for each urban study area as the basis for characterizing population exposure in modeling risk. The use of composite monitors reflects consideration for the way ambient O_3 data are used in the epidemiological studies providing the C-R functions (see section 7.1.1). For the short-term exposure related health endpoints, the composite monitor values derived for this analysis include hourly time series for each study area (where the O_3 value for each hour is the average of measurements across the monitors in that study area reporting values for that hour). Once these composite monitor hourly time series are constructed, we can then extract short-term peak O_3 metrics needed to model specific health effects endpoints. For short-term O_3-attributable endpoints, reflecting consideration for available evidence in the published literature (see section 7.3.2), we have focused the analysis on short-term peak O_3 metrics including 1hr maximum, 8hr mean and 8hr maximum. The 24 hour average has been deemphasized for this analysis, although it is still used in risk modeling when use of C-R functions based on this metric allow us to cover a specific health effect endpoint/location of particular interest[14] (see section 7.3.2).

For modeling mortality risk associated with long-term O_3-attributable we construct seasonally-averaged maximum hourly O_3 values (see section 7.3.2). The derivation of composite monitor distributions used in modeling this health effect endpoint is different than that used for short-term O_3-attributable endpoints. Specifically, for the long-term O_3-attributable endpoint we first construct the seasonally-averaged peak O_3 metric for each monitor within a given study area and then average those monitor-specific metric values together to generate a single composite value to use in generating risk estimates for that study area.

In applying effect estimates obtained from epidemiological studies we attempted to match the modeling period (e.g. O_3 monitoring season) associated with each epidemiology study. This increases overall confidence in the risk compared with using a single more generalized specification of the modeling period. As discussed earlier, we modeled all health effect endpoints for the core analysis using a CBSA-based study area. The use of the CBSA-based study areas addresses potential bias that would have occurred had we focused the risk assessment on the smaller core urban study areas. (see section 7.1.1). Table 7-1 identifies (a) the counties associated with the CBSA definition for each of the 12 urban study areas, (b) the number of O_3 monitors associated with each CBSA (and a flag for whether the design value monitor is

[14] In order to provide estimates of respiratory-related HA for LA, we did include a C-R function based on Linn et al., 2000, which utilizes a 24 hour average exposure metric.

contained within the CBSA), (c) the number of monitors associated with the smaller Smith et al., 2009-based study areas, and (d) the specific O_3 modeling period for each study area. A map showing the counties and monitors for these 12 urban areas can be found in Chapter 4 (figure 4-5, Section 4.3.2.1).

Table 7-1 Information on the 12 Urban Case Study Areas in the Risk Assessment

Study Area	Counties associated with the CBSA definition	# of O_3 Monitors within the CBSA[a]	Required O_3 Monitoring Season
Atlanta	Barrow, Bartow, Butts, Carroll, Cherokee Clayton, Cobb, Coweta, Dawson, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Haralson, Heard, Henry, Jasper, Lamar, Meriwether, Newton, Paulding, Pickens, Pike, Rockdale, Spalding, Walton	13 (3)	March - October
Baltimore	Anne Arundel, Baltimore, Carroll, Harford, Howard, Queen Anne's, Baltimore	7 (1)	April - October
Boston	Essex, Middlesex, Norfolk, Plymouth, Suffolk, Rockingham, Strafford	11* (2)	April - September
Cleveland	Cuyahoga, Geauga, Lake, Lorain, Medina	10* (4)	April - October
Denver	Adams, Arapahoe, Broomfield, Clear Creek, Denver, Douglas, Elbert, Gilpin, Jefferson, Park	16 (6)	March - September
Detroit	Lapeer, Livingston, Macomb, Oakland, St. Clair, Wayne	8 (4)	April - September
Houston	Austin, Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, San Jacinto, Waller	22 (17)	January - December
Los Angeles	Los Angeles, Orange	21* (17)	January - December
New York	Bergen, Essex, Hudson, Hunterdon, Middlesex, Monmouth, Morris, Ocean, Passaic, Somerset, Sussex, Union, Bronx, Kings, Nassau, New York, Putnam, Queens, Richmond, Rockland, Suffolk, Westchester, Pike	22 (7)	April - October
Philadelphia	New Castle, Cecil, Burlington, Camden, Gloucester, Salem, Bucks, Chester, Delaware, Montgomery, Philadelphia	15 (4)	April - October
Sacramento	El Dorado, Placer, Sacramento, Yolo	17 (8)	January - December
St. Louis	Bond, Calhoun, Clinton, Jersey, Macoupin, Madison, Monroe, St. Clair, Franklin, Jefferson, Lincoln, St. Charles, St. Louis, Warren, Washington, St. Louis	17 (2)	April - October

a – This column presents the number of monitors within each CBSA, whether the design value falls outside of the CBSA (denoted with an "*") and the number of monitors within the smaller Smith et al., 2009-based study area (in parenthesis).

We estimate risk associated with recent O_3 conditions as well as risk associated with simulating just meeting the existing and alternative standards. While the derivation of composite

monitor hourly O_3 distributions (and associated peak exposure metrics) for recent conditions is relatively straightforward, the generation of these estimates for the scenarios of just meeting the existing and alternative standards is more complex. The procedures for simulating attainment of both existing and alternative O_3 standards are presented in Chapter 4 and Chapter 4 appendices.

Summary statistics for the air metrics used in modeling risk for each of the 12 urban study areas under recent conditions and simulated attainment of the existing and alternative standard levels are presented in Chapter 4 (see section 4.3.3.2, Figures 4-10 (2007) and 4-11 (2009)).

7.3 SELECTION OF MODEL INPUTS AND ASSUMPTIONS

7.3.1 Selection of Urban Study Areas

This analysis focuses on modeling risk for a set of urban study areas, reflecting the goal of providing risk estimates that have greater overall confidence due to the use of location-specific data when available for these urban locations. In addition, given the greater availability of location-specific data, a more rigorous evaluation of the impact of uncertainty and variability can be conducted for a set of selected urban study areas than would be possible for a broader regional or national-scale analysis. We considered the following factors in selecting the 12 urban study areas included in this analysis:

- **Air Quality Data**: An urban area has reasonably comprehensive monitoring data for the period of interest (2006-2010) to support the risk assessment. This criterion was evaluated qualitatively by considering the number of monitors within the CBSA of the prospective urban areas. Locations with one or two monitors would be excluded since they had relatively limited spatial coverage in characterizing O_3 levels.

- **Elevated Ambient O_3 Levels**: Because we are interested in evaluating the potential magnitude of risk reductions associated with just meeting the existing and alternative O_3 standard levels, we focus on study areas with elevated ambient O_3 levels at or above the existing standard, such that just meeting alternative O_3 standard levels would result in some degree of risk reduction.

- **Location-specific C-R Functions**: Given the health endpoints selected for inclusion in the analysis (see section 7.3.2), there are epidemiological studies of sufficient quality available for these urban study areas to provide the C-R functions necessary for modeling risk. This criterion primarily applies to short-term epidemiological studies since the associated health effect endpoints are the primary focus of the REA. Short-term O_3-

attributable epidemiological studies often include city-specific effect estimates, and in some cases are multi-city studies that provide estimates for multiple cities.

- **Baseline Incidence Rates and Demographic Data**: The required urban area-specific baseline incidence rates and population data are available for a recent year for at least one of the health endpoints.

- **Geographic Heterogeneity**: Because O_3 distributions and population characteristics vary geographically across the U.S., we selected urban study areas to provide coverage for regional variability in factors related to O_3 risk including variability in the spatial pattern of O_3 in the urban area, population exposure (differences in residential housing density, air conditioning use and commuting patterns), demographic characteristics (baseline incidence rates, SES) and variability in effect estimates. The degree to which the set of urban study areas provided coverage for regional differences across the U.S. in many of these O_3 risk-related factors was evaluated as part of the representativeness analysis presented in Chapter 8.

Application of the above criteria resulted in the selection of 12 urban study areas for inclusion in the risk assessment including:

- Atlanta, GA
- Baltimore, MD
- Boston, MA
- Cleveland, OH
- Denver, CO
- Detroit, MI
- Houston, TX
- Los Angeles, CA
- New York, NY
- Philadelphia, PA
- Sacramento, CA
- St. Louis, MO

The specific set of counties used in defining each of the 12 urban study areas based on the CBSA is presented in Table 7-1.

7.3.2 Selection of Epidemiological Studies and Specification of Concentration-Response Functions

Once the set of health effect endpoints to be included in the risk assessment has been specified, the next step was to select the set of epidemiological studies that will provide the effect estimates and model specifications used in the C-R functions. This section describes the approach used in completing these tasks and presents a summary of the epidemiological studies and associated C-R functions specified for use in the risk assessment.

In Chapter 2, section 2.5 we identified the set of health effect categories and associated endpoints to be included in this assessment, based on review of the evidence provided in the O_3 ISA (U.S. EPA, 2013a). The selection of specific health effect endpoints to model within a given health effect endpoint category is an iterative process involving review of both the strength of evidence (for a given endpoint) as summarized in the O_3 ISA together with consideration for the available epidemiological studies supporting a given endpoint and the ability to specific key inputs needed for risk modeling, including effect estimates and model forms. Ultimately, endpoints are only selected if (a) they are associated with an overarching effect endpoint category selected for inclusion in the risk assessment and (b) they have sufficient epidemiological study support to allow their modeling in the risk assessment. Health effect endpoints selected for inclusion in the second draft REA include:

Short-term O_3-attributable:

- Mortality (likely to be a casual relationship)
  - All-cause (non-accidental)
  - Cardiovascular
  - Respiratory
- Respiratory effects (causal relationship)
  - ED (asthma, wheeze, all respiratory symptoms)
  - HA (COPD, asthma, all respiratory) [15]
  - Respiratory symptoms

Long-term O_3-attributable:

[15] Regarding COPD-related HA, the O_3 ISA states that "Although limited in number, both single- and multi-city studies consistently found positive associations between short-term O_3 exposures and asthma and COPD hospital admissions." (U.S. EPA 2013a, p. 6-128). It is also important to point out that when modeling of COPD-related HA is limited to the summer months (as was done for the REA), available effect estimates have tighter confidence intervals and are generally positive, which increases overall confidence in the resulting risk estimate (see U.S. EPA 2013a, Figure 6-19).

- Respiratory effects, focusing on respiratory-related mortality (likely causal relationship).[16]

We selected epidemiological studies to support modeling of the health effect endpoints listed above by applying a number of criteria including[17]:

- The study was peer-reviewed, evaluated in the O_3 ISA, and judged adequate by EPA staff for purposes of inclusion in the risk assessment. We considered the following criteria: whether the study provides C-R relationships for locations in the U.S., whether the study has sufficient sample size to provide effect estimates with a sufficient degree of precision and power, and whether adequate information is provided to characterize statistical uncertainty.
- Preference for multicity studies given that they typically have greater power and reflect patterns of O_3 related health effects over a range of urban areas (and regions) which can display variability in key risk-related factors such as exposure measurement error. In the case of short-term O_3-attributable mortality, we also favored those multi-city studies for which we could obtain Bayesian-adjusted city-specific estimates from the study authors, since these incorporate both city-specific effect information with information from the broader array of cities included in the study. In those instances where we did not have multi-city studies (e.g., with many of the short-term respiratory-related morbidity endpoints) we use single-city studies.
- The study design is considered robust and scientifically defensible, particularly in relation to methods for covariate adjustment, including treatment of confounders, as well as treatment of effect modifiers. For example, if a given study used ecological-defined variables (e.g., smoking rates) as the basis for controlling for confounding, concerns may be raised as to the effectiveness of that control.
- The study is not superseded by another study (e.g., if a later study is an extension or replication of a former study, the later study would effectively replace the former study), unless the earlier study has characteristics that are clearly preferable (e.g., inclusion of copollutants models, or use of a peak exposure metric of interest).

[16] The O_3 ISA classifies long-term O_3-attributable respiratory health effects, including respiratory-related mortality, as having a likely causal classification. By contrast, it classifies long-term O_3-attributable total mortality as having a suggestive of a causal relationship classification (O_3 ISA, 2012, Chapter 1). We have focused on modeling long-term O_3-attributable respiratory-related mortality given the greater support for this health endpoint relative to total mortality.

[17] In addition to the criteria listed here, we also attempted to include studies that provide coverage for populations considered particularly at-risk for a particular health (e.g., children, individuals with preexisting disease). However, a study would have to meet the criteria listed here (in addition to providing coverage for an at-risk population) in order for that study to be used to derive C-R functions.

We applied the above criteria and selected the set of epidemiological studies presented in Table 7-2 for use in specifying C-R functions (Table 7-2 also describes elements of the C-R functions specified using each epidemiological study, as discussed below).

As part of methods refinement for this risk assessment, we considered studies that utilized more sophisticated and potentially representative exposure surrogates in characterizing population exposure (e.g., using population-weighted O_3 monitor values instead of equally-weighted monitors, linking exposures in individual counties or U.S. Census tracts to the nearest monitor, rather than using a composite monitor value to represent the entire study area). However, analysis conducted by EPA demonstrated that use of the simpler composite monitor approach (as used for other short-term O_3-attributable morbidity endpoints) generated risk estimates that were very close to those generated using the population-weighted O_3 metric (see REFERENCE- Karen Wesson???). Therefore, in order to conserve time and resources, we modeled this endpoint using the more generalized composite monitor-based metric. And finally, a number of the long-term O_3-attributable morbidity studies originally considered for modeling this endpoint category did involve more complex O_3 metrics (e.g., Atkinbami et al., 2010, Meng et al., 2010, and Moore et al., 2008). However, limitations in the study-level data required to support risk assessment prevents us at this point from completing a quantitative risk assessment for this category of health endpoints with a reasonable degree of confidence.[18]

Based on additional evaluation of the literature, we have substituted Smith et al., 2009 for Bell et al., 2004 as a source of Bayes-adjusted city-specific effect estimates to support modeling short-term O_3-attributable mortality. This decision reflects a number of factors. The Smith et al., 2009 study includes a wider range of simulations exploring sensitivity of the mortality effect to different model specifications including (a) regional versus national Bayes-based adjustment, (b) copollutants models considering PM_{10}, and (c) all - year versus O_3-season based estimates. This is contrasted with the Bell et al., 2004 study which does not provide this degree of model exploration. In obtaining the city-specific Bayes-adjusted effect estimates for the Smith et al., 2009 study from the study authors, we were provided with estimates reflecting this range of alternative model specifications which allowed us to incorporate them into both the core and sensitivity analysis portions of the REA (see section 7.4.3). In addition, the Smith et al., 2009 study does not use the trimmed mean approach employed in the Bell et al., 2004 study in preparing O_3 monitor data. We have a number of concerns regarding the trimmed mean approach including (1) the potential loss of temporal variation in the data when the approach is used (this could impact the size of the effect estimate) and (2) a lack of complete documentation for the

[18] However, as noted in section 7.7.3 of the first draft REA, these limitations do not prevent the use of this evidence from informing consideration of the levels of exposure at which specific types of health effects may occur (i.e., the evidence analysis, which is an important aspect of the O_3 NAAQS review). Rather, these limitations only prevent the quantitative estimation of risk with a reasonable degree of confidence.

approach which prevents us from fully reviewing the technique and using it in preparing O_3 metrics for the REA. Given these concerns, we view it as advantageous that the Smith et al., 2009 study does not use the trimmed mean approach.

With the exception of the trimmed mean approach, the Smith et al., 2009 study was intended to reproduce the results of the Bell et al., 2004 analysis. Thus, the core risk results based on Smith et al 2009 are comparable to the 1st draft REA estimates based on Bell et al 2004, while the alternative models provided in Smith et al 2009 allow for an expanded set of sensitivity analyses. The comparability of the Smith et al 2009 and Bell et al 2004 estimates is confirmed by the graphical comparison in Smith et al 2009 of mortality effect estimates (for the 24hr O_3 metric) with matching effect estimates from Bell et al., 2004. This comparison demonstrates the close match of the two studies (for this particular scenario).

Reflecting the points made above, in modeling short-term O_3-attributable mortality, we have included a core analysis based on the national-Bayesian adjusted city-specific effect estimates (reflecting the full O_3 monitoring period in each city) obtained from Smith et al., 2009. As sensitivity analyses, we have included effect estimates obtained from Smith et al., 2009 which reflect application of copollutants models (including PM_{10}), Bayes adjustment using a regional prior,[19] and a shorter fixed O_3 measurement period (April-October). In the 1st draft REA, we had also included national Bayes-adjusted effect estimates (reflecting a fixed June-August period) obtained from Zanobetti and Schwartz, 2008 as part of the core analysis. However, we have decided to instead include these as part of the sensitivity analysis in this 2nd draft of the REA since these effect estimates cover a more limited warm-weather period and consequently will generate only partial characterizations of mortality risk (since they exclude risk occurring during the non-summer months).

We have also included estimates of respiratory-related mortality associated with long-term O_3 exposures based on effect estimates obtained from Jerrett et al., 2009. The decision to model long-term O_3-attributable mortality reflects consideration for evidence supporting a likely to be a causal relationship for long-term O_3-attributable respiratory effects, including mortality (O_3 ISA, section 2.5.2, U.S. EPA, 2013a). After considering its strengths and weaknesses, we consider the Jerrett et al. (2009) study to be an appropriate basis for estimating long-term O_3-related respiratory mortality risk. Key strengths of this study are that it (a) included 1.2 million participants in the American Cancer Society cohort from all 50 states, DC, and Puerto Rico; included O_3 data from 1977 (5 years before enrollment in the cohort began) to 2000; (b) considered co-pollutant models that controlled for $PM_{2.5}$; and (c) explored the potential for a threshold concentration associated with the long-term mortality endpoint. Importantly, this study

[19] With application of a regional prior within Bayesian adjustment, city-specific effect estimates are adjusted towards the regional value rather than a national value as is the case with the application of a national prior.

was also the first to explore the relationship between long-term O_3 exposure and respiratory mortality (rather than focusing on cardiopulmonary mortality). Key limitations are possible exposure misclassification and uncontrolled confounding by temperature, which are endemic to most long-term epidemiological studies. While Jerrett et al. (2009) found negative associations between O_3 exposure and cardiovascular mortality when controlling for $PM_{2.5}$, null or negative associations for O_3 are consistent with the evidence that $PM_{2.5}$ is the pollutant most strongly associated with cardiovascular disease (EPA 2009 PM ISA).

Our analysis includes a core estimate based on a co-pollutant model (with $PM_{2.5}$). The seasonal average metrics used in the long-term exposure mortality estimate are not very sensitive to the reduced number of days with co-pollutant monitoring, and as such it is appropriate to include the copollutant model as the core estimate. We also include two sensitivity analyses for long-term O_3-attributable respiratory mortality including: (a) application of regionally-differentiated effect estimates (although these do not include a copollutants model specification) and (b) application of a single pollutant (O_3-only) national-based effect estimate.

The effect estimates used in modeling long-term O_3-attributable mortality (see Table 7-2) utilize a seasonal average of peak (1hr maximum) measurements. These long-term exposure metrics can be viewed as long-term exposures to daily peak O_3 over the warmer months, as compared with annual average levels such as are used in long-term PM exposure calculations. This increases the need for care in interpreting these long-term O_3-attributable mortality estimates together with the short-term O_3-attributable mortality estimates, in order to avoid double counting. It is also important to keep in mind that our estimates of short-term O_3-attributable mortality are for all-causes, while estimates of long-term O_3-attributable mortality are focused on respiratory-related mortality. This further limits the ability to compare estimates of long-term and short-term exposure related mortality.

Once the set of epidemiology studies described above was selected, the next step was to specify C-R functions for use in the risk assessment. Several factors were considered in identifying the effect estimates and model forms used in specifying C-R functions for each endpoint. These factors are described below:

- O_3 **Exposure Metric:** In the risk assessment supporting the previous O_3 NAAQS review, for short-term exposure, we included C-R functions based on 24hr averages as well as a number of peak O_3 measurements. However, given that the the current O_3 NAAQS standard uses an 8hr form and given that many of the clinical studies involving O_3 also utilize shorter exposures (on the order of 2 to 8 hrs – see O_3 ISA, section 6.2.1.1), we wanted to see if the latest epidemiological studies for O_3 also supported use of an 8hr averaging time in modeling risk. Several epidemiological studies completed since the last

review provide limited support for stronger associations between health endpoints and peak O_3 metrics (i.e., 1hr maximum, 8hr maximum and 8hr means) relative to 24hr averages. Specifically, a study of respiratory ED visits in Atlanta (Darrow et al., 2011) found stronger associations with peak metrics (including 1hr and 8hr maximum measurements) compared with 24hr averages (see O_3 ISA section 6.2.7.3 and Figure 6-17, U.S. EPA, 2013a). Similarly, for short-term exposure-related mortality, there are also a limited number of epidemiologic studies that have compared mortality associations with peak O_3 metrics and the 24hr average metric. Although the O_3 ISA recognizes that 24hr exposure metrics when used in time series studies may result in smaller risk estimates, ultimately it concludes that "Overall, the evidence from time-series and panel epidemiologic studies does not indicate that one exposure metric is more consistently or strongly associated with mortality or respiratory-related health effects" (U.S. EPA, 2013a, section 2.5.4.2). Based on consideration for the evidence summarized in the O_3 ISA, we have decided to focus on peak exposure metrics because of the limited evidence that these metrics may be associated with higher risk estimates relative to the 24 hr exposure metric. However, we recognize that, as summarized in the O_3 ISA, there is only weak support for differentiating between these two categories of short-term exposure metric.

Table 7-2 Overview of Epidemiological Studies Used in Specifying C-R Functions

Epidemiological study (*stratified by* O_3-*attributable health endpoints*)	Health endpoints	Location (urban study area(s) covered)	Exposure metric (and modeling period)	Additional study design details	Notes regarding application in the analysis
Short-term O_3***-attributable mortality***					
Smith et al., 2009	Non-accidental, respiratory, cardiovascular	95 large urban communities (provides coverage for all 12 urban study areas)	24hr avg, 8hr max, 1hr max. April through October and all year	Adjusting for time-varying confounders (PM, weather, seasonality). Lag structure included 0, 1, 2 and day 3 lag as well as 0-6 day distributed lag. Age range: all ages.	Focused on the 8hr max-based metric C-R functions for the REA (see text discussion later in this section). Obtained Bayes-adjusted city-specific effect estimates for non-accidental all-cause mortality from Dr. Smith (personal communication, Dr. Richard L. Smith, January 15, 2013) reflecting consideration for the following modeling elements: (a) regional- versus national-prior Bayes model adjustment, (b) single pollutant versus copollutants (PM_{10}) models, and (c) full monitoring period versus summer only (April-October). For the core analysis, we focused on the single pollutant (O_3-only) model covering the full monitoring period. The copollutants model (with PM_{10}) was included as a sensitivity analysis (see section 7.4.3).
Zanobetti and Schwartz (2008)	Non-accidental, respiratory, cardiovascular	48 U.S. cities (provides coverage for the 12 urban study areas)	8hr max. June-August	Effect controlled for season, day of week, and temperature. Lag structure included 0-3d, 0-20 and 4-20 day). Age range: all ages	Obtained Bayes-adjusted city-specific effect estimates for non-accidental, respiratory and cardiovascular from Dr. Zanobetti (personal communication, Dr. Antonella Zanobetti, January 5, 2012). These effect estimates reflect a 0-3 day distributed lag and are based on 8hr mean O_3 levels measured between June and August. Estimates were generated for each study area using this constrained warm-season period.
Short-term O_3***-attributable morbidity - HA for respiratory effect)***					
Medina-Ramon et al., 2006.	HA: COPD, pneumonia	36 cities (provides coverage for all 12 urban study areas)	8hr mean. warm (May-September), cool (October-April), all year	Distributed lag (0-1 day). Age range: ≥ 65yrs. Controlled for day of the week and weather (including temperature).	Generated risk estimates based on warm season for COPD only (May-September).
Linn et al., 2000	HA: unscheduled	LA only	24hr mean, LA O_3 season (all year),	Lag 0. Age range: all ages. Used subgroup analysis to explore the	Included effect estimate based on 24hr avg metric (for summer) since this provided additional

Epidemiological study (*stratified by* O_3-*attributable health endpoints*)	Health endpoints	Location (urban study area(s) covered)	Exposure metric (and modeling period)	Additional study design details	Notes regarding application in the analysis
	for pulmonary illness		winter, spring, summer and autumn	effect of temporal variation, weather and autocorrelation on O_3 effect.	coverage for HA in L.A. Modeled using air quality for June-August.
Lin et al., 2008	HA: respiratory disease	NY State (used to cover NYC)	1hr max (for 10am-6pm interval), warm season (April-October)	Lag 0, 1, 2, 3. Age range: <18yrs. Models adjusted for the confounding effects of demographic characteristics, particulate matter(PM_{10}), meteorological conditions, day of the week, seasonality, long-term trends, and different lag periods of exposure.	Used 1hr max metric applied to the city-specific O_3 season for NYC (April-October).
Katsouyanni et al 2009	HA: cardiovascular disease, chronic obstructive pulmonary disease, pneumonia, all respiratory	14 cities (provides coverage for Detroit only)	1hr max. Summer only and all year	Lag 0-1day. Age range: ≥ 65yrs. Models accounted for seasonal patterns, but also, for weekend and vacation effects, and for epidemics of respiratory disease. The data were also analyzed to detect potential thresholds in the concentration–response relationships.	C-R function applied only for all respiratory endpoint. Used June-August-based composite monitor.
Silverman et al., 2010	HA: asthma (ICU and non-ICU)	NYC	8hr max. Warm season (April-August)	Includes control for $PM_{2.5}$. Lag 0-1 day. Age range: children 6-18yrs. The model adjusted for temporal trends, weather, and day of the week.	Applied C-R function (for O_3 and O_3 with control for $PM_{2.5}$) to the city-specific O_3 season for NYC (slightly longer than the modeling period used in the study).
Short-term O_3-attributable morbidity– ED and ER visits (respiratory)					
Ito et al., 2007	ED: asthma	NYC	8hr max. Warm season (April-September)	Includes models controlling for SO_2, NO_2, CO and $PM_{2.5}$. Lag: 0, 1, and distributed lag (0-1 day). Age range: all ages. Model adjusts for temporal trends, weather terms, day-of-week and other pollutants.	Applied C-R functions (for O_3 alone and O_3 with control for listed pollutants) to the city-specific O_3 season for NYC (slightly longer than the modeling period used in the study).
Tolbert et al., 2007	ED: all respiratory	Atlanta	8hr max. Summer (March-October)	Includes models controlling for NO_2, CO, PM_{10} and NO_2/ NO_2. Age range: all ages. Model controls for temporal trends, temperature,	Applied C-R functions (for O_3 alone and O_3 with control for listed pollutants) to the city-specific O_3 season for Atlanta.

Epidemiological study (*stratified by* O_3-*attributable health endpoints*)	Health endpoints	Location (urban study area(s) covered)	Exposure metric (and modeling period)	Additional study design details	Notes regarding application in the analysis
				other pollutants.	
Strickland et al., 2010	ER: respiratory	Atlanta	8hr max (based on population weighted average across monitors). Warm season (May to October) and cool (November to April)	Lag: average of 0-2 day, distributed lag 0-7 day. Age range: 5-17yrs. Model controls for seasonal trends and meteorology.	Included effect estimates based on both lag structures and used composite monitor values for city-specific O_3 season.
Darrow etl al., 2011	ED: all respiratory	Atlanta	8hr max, 1hr max, 24hr avg for summer (March-October).	Lag: 1day. Age range: all ages. The study used a time series analysis similar to case-crossover with crossover matching based on daily temperature (rather than day of the week) to provide control for this key risk-related factor.	Used city-specific O_3 season-based composite monitor values.
Short-term O_3-attributable morbidity – respiratory symptoms					
Gent et al., 2003	Respiratory symptoms: wheeze, persistent cough, chest tightness, shortness of breath	Springfield MA (study used to cover Boston)	1hr max, 8hr max	Lag: 0 and 1 day. Age range: asthmatic children <12 yrs. Model adjusted for temperature.	Included effect estimates for different symptoms based on both 8hr max and 1hr max metrics (for city-specific O_3 season composite monitor values for Boston). The study area (which focuses on Springfield and the northern portion of Connecticut) does not encompass Boston. However, we are willing to accept uncertainty associated with using effect estimates from this study to provide coverage for Boston given the goal of providing coverage for this morbidity endpoint. However, there is increased uncertainty associated with modeling for this endpoint.
Long-term O_3-attributable respiratory mortality					
Jerrett et al., 2009	Respiratory, cardiovascular, cardiopulmonary	96 metropolitan statistical areas (provides coverage for all 12 study areas)	Seasonal average (i.e. Apr-Sep) of the peak daily 1hr max values.	>30 yrs of age, includes national-level and regional effect estimates (only national-level estimate has copollutants modeling considering PM2.5 along with O_3). Modeling included consideration for a range of potential confounders evaluated	Included national copollutants model-based effect estimates in core analysis and single-pollutant model regional effect estimates and national effect estimates as sensitivity analyses.

Epidemiological study (*stratified by* O_3-*attributable health endpoints*)	Health endpoints	Location (urban study area(s) covered)	Exposure metric (and modeling period)	Additional study design details	Notes regarding application in the analysis
				at both the ecological level and personal level.	

- **Single-and Multi-pollutant Models (*pertains to both short-term and long-term exposure studies*)**: Epidemiological studies often consider health effects associated with ambient O_3 using both single-pollutant and co-pollutant models. To the extent that any of the co-pollutants present in the ambient air may have contributed to health effects attributed to O_3 in single pollutant models, risks attributed to O_3 may be overestimated or underestimated if C-R functions are based on single pollutant models. This would argue for inclusion of models reflecting consideration of co-pollutants. Conversely, in those instances where co-pollutants are highly correlated with O_3, inclusion of those pollutants in the health impact model can produce unstable and statistically insignificant effect estimates for both O_3 and the co-pollutants. Furthermore, there are often significant differences in sampling frequencies for each pollutant included in copollutants models, which can lead to a loss of statistical power in copollutants models (relative to single pollutant models). These last points could argue for inclusion of a model based exclusively on O_3. Given that single and multi-pollutant models each have potential advantages and disadvantages, to the extent possible, given available information we have included both types of C-R functions in the risk assessment.
- **Multiple Effect Estimates within a Given CBSA-based Study Area**: As noted earlier in section 7.1.1, for this analysis, all health endpoints, including short-term O_3-attributable mortality are modeled using CBSA-based study areas. In the case of both Smith et al., 2009 and Zanobetti and Schwartz 2008, these CBSA-based study areas are larger than the study areas used in these epidemiological studies to derive effect estimates. Furthermore, for some of the CBSA-based urban study areas, several of the smaller study areas evaluated in the epidemiological study fall within a single larger CBSA-based study area. For example, with the Smith et al., 2009 study, multiple effect estimates are available for the CBSA-defined study areas of Los Angeles and New York City. Specifically, the Smith et al., 2009 study provides separate effect estimates for (a) Santa Anna/Anaheim and Los Angeles study areas, both of which fall within the larger CBSA-based Los Angeles study area and (b) New York, Jersey City and Newark study areas, all of which fall within the larger CBSA-defined New York study area (see Table 7-3). This raises the question of how to specify the effect estimate for these larger CBSA-based study areas when there are multiple effect estimates available from the epidemiological study. For this analysis, in those instances where there are multiple effect estimates, we have decided to use the effect estimate that represents the largest number of residents within each CBSA-based study area. There is uncertainty associated with this decision which is discussed both in section 7.4.2 and section 7.5.3 (as part of the air quality-related sensitivity analysis discussion).

Table 7-3 CBSA-based Study Areas with Multiple Effect Estimates from the Smith et al., 2009 Study*

CBSA Study Area	Smith et al., 2009 (smaller) study areas with CBSA-based study area	Population totals	Mortality effect estimate	Comments
New York City	New York, NY	9,100,000	0.0009	New York study area dominates from a population standpoint, so that effect estimate was chosen to represent the entire CBSA. An additional 8.3 million people live in portions of the New York CBSA not covered by the Smith et al., 2009 study areas.
	Jersey City, NJ	630,000	0.0001	
	Newark, NJ	780,000	0.0005	
Los Angeles	Santa Ana/Anaheim, CA	3,000,000	0.0002	Los Angeles dominates from a population standpoint, so that effect estimate was chosen to represent the entire CBSA. In this case, the full CBSA-based study area is covered by the Smith et al., 2009-based subareas.
	Los Angeles, CA	9,800,000	0.0001	

* Source: obtained from Dr. Smith (personal communication, Dr. Richard L. Smith, January 15, 2013)

- **Single-city Versus Multi-city Studies**: All else being equal, we judge C-R functions estimated in the assessment location as preferable to a function estimated in some other location, to avoid uncertainties that may exist due to differences associated with geographic location. There are several advantages, however, to using estimates from multi-city studies versus studies carried out in single cities. Multi-city studies are applicable to a variety of settings, since they estimate a central tendency across multiple locations. Multi-city studies also tend to have more statistical power and provide effect estimates with relatively greater precision than single-city studies due to larger sample sizes, reducing the uncertainty around the estimated health coefficient. By contrast, single-city studies, while often having lower statistical power and varying study designs which can make comparison across cities challenging, reflect location-specific factors such as differences in underlying health status, and differences in O_3 exposure-related factors such as air conditioner use and patterns of urban residential density. There is a third type of study design that generates Bayes-adjusted city-specific effect estimates, thereby combining the advantages of both city-specific and multi-city studies. Bayes-adjusted city-specific estimates begin with a city-specific effect estimate and shrink that towards a multi-city mean effect estimate based on consideration for the degree of variance in both estimates. We have elected to place greater confidence on these types of Bayesian-adjusted effect estimates when they are available. Otherwise, given the advantages for both city-specific and multi-city effect estimates, we have used both types when available.

- **Multiple Lag Models**: Based on our review of evidence provided in the O_3 ISA, we believe there is increased confidence in modeling both short-term O_3-attributable mortality and respiratory morbidity risk based on exposures occurring up to a few days prior to the health effect, with less support for associations over longer exposure periods or effects lagged more than a few days from the exposure (see O_3 ISA section 2.5.4.3, U.S. EPA, 2013a). Consequently, we have favored C-R functions reflecting shorter lag periods (e.g., 0, 1 or 1-2 days). With regard to the specific lag structure (e.g, single day versus distributed lags), the O_3 ISA notes that epidemiological studies involving respiratory morbidity have suggested that both single day and multi-day average exposures are associated with adverse health effects (see O_3 ISA section 2.5.4.3). Therefore, when available both types of lag structures where considered in specifying C-R functions for short-term O_3-attributable mortality and morbidity.
- **Seasonally-differentiated Effects Estimates**: The previous O_3 Air Quality Criteria Document (AQCD) (published in 2006) concluded that aggregate population time-series studies demonstrate a positive and robust association between ambient O_3 concentrations and respiratory-related hospitalizations and asthma ED visits during the warm season (see O_3 ISA section 2.5.3.1 U.S. EPA, 2013a). The current O_3 ISA notes that recent studies of short-term O_3-attributable respiratory mortality in the U.S. suggest that the effect is strengthened in the summer season (O_3 ISA section 2.5.3.1, U.S. EPA, 2013a). In addition, many of the key epidemiological studies discussed in the current O_3 ISA exploring both short-term exposure related mortality and morbidity have larger (and more statistically significant) effect estimates when evaluated for the summer (O_3) season, relative to the full year (see O_3 ISA Figures 6-20 and 6-27, U.S. EPA, 2013a). However, if we focus the assessment of risk on the warm season, we bias our estimate by excluding potential effects associated with cooler (non-summer) months. Given our desire to provide a more complete picture of overall risk in each of the study areas, we have favored (for the core analysis) effect estimates that cover the full O_3 monitoring period specific to each study area, rather than the more limited warm (summer) period.
- **Shape of the Functional Form of the Risk Model (including threshold):** The current O_3 ISA concludes that there is little support in the literature for a population threshold for short-term O_3-attributable effects. However, specifically in relation to mortality, the O_3 ISA concludes that a national or combined analysis may not be appropriate to identify whether a threshold exists (see O_3 ISA, section 2.5.4.4, U.S. EPA, 2013a).[20] Given the

[20] Specifically, given the multi-city nature of these mortality studies combined with the variability in O_3 and other factors related to exposure and risk, the O_3 ISA concludes that these studies are not well positioned to evaluate the potential for a threshold in the mortality effect.

above general observation from the O_3 ISA regarding the low potential for thresholds, we did not include C-R functions for any of the short-term O_3-attributable health endpoints modeled that incorporated a threshold.[21]

Application of the above criteria resulted in an array of C-R functions specified for the risk assessment (see Table 7-2), including functions covering short-term O_3-attributable mortality and morbidity and long-term O_3-attributable mortality. In presenting the C-R functions in Table 7-2, we have focused on describing key attributes of each C-R function (and associated source epidemiological study) relevant to a review of their use in the risk assessment. More detailed technical information including effect estimates and model specification is provided in Appendix 7A. Specific summary information provided in Table 7-2 includes:

- *Health endpoints:* identifies the specific endpoints evaluated in the study. Generally we included all of these in our risk modeling, however, when a subset was modeled, we reference that in the "Notes" column (last column in the table).
- *Location*: identifies the specific urban areas included in the study and maps those to the set of 12 urban study areas included in the risk assessment.
- *Exposure metric*: describes the exposure metric used in the study, including the specific modeling period (e.g., O_3 season, warm season, full year). We developed two categories of composite monitor values to match the modeling periods used in the two short-term O_3-attributable mortality studies providing C-R functions for the analysis. For the remaining morbidity endpoints, we mapped specific C-R functions to whichever of these two composite monitor categories most closely matched the modeling period used in the underlying epidemiological study. This mapping (for morbidity endpoint C-R functions) is described in the "Notes" column (the seasons reflecting in modeling for each C-R function are also presented in Appendix 7A).
- *Additional study design details*: this column provides additional information primarily covering the lag structure and age ranges used in the study.
- *Notes regarding application in second draft analysis*: as the name implies, this column provides notes particular to the application of a particular epidemiological study and associated C-R functions in the risk assessment.

[21] While clinical studies have suggested the presence of a threshold for respiratory effects, these should not be used to support specification of population-level thresholds for use in the epidemiological-based risk assessment. The clinical studies focus on relatively small and clearly defined populations of healthy adults which are not representative of the broader residential populations typically associated with epidemiological studies, including older individuals and individuals with existing health conditions which place them at greater risk for O_3-related effects. Therefore, the clinical studies are unlikely to have the power to capture population thresholds in a broader and more diverse urban residential population, should those thresholds exist.

7.3.3 Baseline Health Effect Incidence and Prevalence Data

As discussed earlier (section 7.1.2), the most common epidemiological-based health risk model expresses the change in health risk (Δy) associated with a given change in O_3 concentrations (Δx) as a percentage of the baseline incidence (y). To accurately assess the impact of O_3 air quality on health risk in the selected urban areas, information on the baseline incidence of health effects (i.e., the incidence under recent air quality conditions) in each location is needed. In some instances, health endpoints are modeled for a population with an existing health condition, necessitating the use of a prevalence rate. Where at all possible, we use county-specific incidences or incidence rates (in combination with county-specific populations). In some instances, when county-level incidence rates were not available, BenMAP can employ more generalized regional rates (see BenMAP Guidance Manual for additional detail, Abt Associates, Inc. 2010). For prevalence rates (which were only necessary for modeling respiratory symptoms among asthmatic children using Gent et al., (2008) - see Table 7-2), we utilized a national-level prevalence rate appropriate for the age group being modeled. A summary of available baseline incidence data for specific categories of effects (and prevalence rates for asthma) is presented below:

- *Baseline incidence data on mortality*: County-specific (and, if desired, age- and race-specific) baseline incidence data are available for all-cause and cause-specific mortality from CDC Wonder.[22] The most recent year for which data are available online is 2005 and this was the source of incidence data for the risk assessment.[23]
- *Baseline incidence data for hospital admissions and emergency room (ER) visits:* Cause-specific hospital admissions baseline incidence data are available for each of 40 states from the State Inpatient Databases (SID). Cause-specific ER visit baseline incidence data are available for 26 states from the State Emergency Department Databases (SEDD). SID and SEDD are both developed through the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ). In addition to being able to estimate State-level rates, SID and SEDD can also be used to obtain county-level hospital admission and ER visit counts by aggregating the discharge records by county.
- *Asthma prevalence rates:* state-level prevalence rates that are age group stratified are available from the Centers for Disease Control and Prevention (CDC) Behavioral Risk Factor Surveillance System (BRFSS) (U.S. CDC, 2010).

[22] http://wonder.cdc.gov/mortsql.html

[23] Note: For years 1999 – 2005, CDC Wonder uses ICD-10 codes; for years prior to 1999, it uses ICD-9 codes. Since most of the studies use ICD-9 codes, this means that EPA will have to create or find a mapping from ICD-9 codes to ICD-10 codes if the most recent data available are to be used.

Incidence and prevalence rates are presented as part of the full set of model inputs documented in Appendix 7A. The incidence rates and prevalence rates provided in Table 7A-1 are weighted average values for the age group associated with each of the C-R functions. These weighted averages are calculated within BenMAP using more refined age-differentiated incidence and prevalence rates originally obtained from the data sources listed in the bullets above.

7.3.4 Population (demographic) Data

To calculate baseline incidence rates, in addition to the health baseline incidence data we also need the corresponding population. We obtained population data from the 2010 U.S. Census (http://www.census.gov/popest/counties/asrh/). These data are then used as the basis for back-casting estimates for simulation years (in this case, 2007 and 2009) (see Appendix J of the BenMAP User's Manual for additional detail, U.S. EPA, 2012b). Total population counts used in modeling each of the health endpoints evaluated in the analysis (differentiated by urban study area and simulation year) are provided as part model inputs presented in Appendix 7A.

7.4 ADDRESSING VARIABILITY AND UNCERTAINTY

An important component of a population risk assessment is the characterization of both uncertainty and variability. *Variability* refers to the heterogeneity of a variable of interest within a population or across different populations. For example, populations in different regions of the country may have different behavior and activity patterns (e.g., air conditioning use, time spent indoors) that affect their exposure to ambient O_3 and thus the population health response. The composition of populations in different regions of the country may vary in ways that can affect the population response to exposure to O_3 – e.g., two populations exposed to the same levels of O_3 might respond differently if one population is older than the other. Variability is inherent and cannot be reduced through further research. Refinements in the design of a population risk assessment are often focused on more completely characterizing variability in key factors affecting population risk – e.g., factors affecting population exposure or response – in order to produce risk estimates whose distribution adequately characterizes the distribution in the underlying population(s).

Uncertainty refers to the lack of knowledge regarding the actual values of inputs to an analysis. Models are typically used in analyses, and there is uncertainty about the true values of the parameters of the model (parameter uncertainty) – e.g., the value of the coefficient for O_3 in a C-R function. There is also uncertainty about the extent to which the model is an accurate representation of the underlying physical systems or relationships being modeled (model uncertainty) – e.g., the shapes of C-R functions. In addition, there may be some uncertainty

surrounding other inputs to an analysis due to possible measurement error—e.g., the values of daily O_3 concentrations in a risk assessment location, or the value of the baseline incidence rate for a health effect in a population.[24] In any risk assessment, uncertainty is, ideally, reduced to the maximum extent possible through improved measurement of key variables and ongoing model refinement. However, significant uncertainty often remains, and emphasis is then placed on characterizing the nature of that uncertainty and its impact on risk estimates. The characterization of uncertainty can be both qualitative and, if a sufficient knowledgebase is available, quantitative.

The selection of urban study areas for the O_3 risk assessment was designed to cover the range of O_3-related risk experienced by the U.S. population and, in general, to adequately reflect the inherent variability in those factors affecting the public health impact of O_3 exposure. Sources of variability reflected in the risk assessment design are discussed in section 7.4.1, along with a discussion of those sources of variability which are not fully reflected in the risk assessment and consequently introduce uncertainty into the analysis.

The characterization of uncertainty associated with risk assessment is often addressed in the regulatory context using a tiered approach in which progressively more sophisticated methods are used to evaluate and characterize sources of uncertainty depending on the overall complexity of the risk assessment (WHO, 2008). Guidance documents developed by EPA for assessing air toxics-related risk and Superfund Site risks (U.S.EPA, 2004 and 2001, respectively) as well as recent guidance from the World Health Organization (WHO, 2008) specify multi-tiered approaches for addressing uncertainty.

The WHO guidance, in particular, presents a four-tiered approach for characterizing uncertainty (see Chapter 3, section 3.2.6 for additional detail on the four tiers included in the WHO's guidance document). With this four-tiered approach, the WHO framework provides a means for systematically linking the characterization of uncertainty to the sophistication of the underlying risk assessment. Ultimately, the decision as to which tier of uncertainty characterization to include in a risk assessment will depend both on the overall sophistication of the risk assessment and the availability of information for characterizing the various sources of uncertainty. We used the WHO guidance as a framework for developing the approach used for characterizing uncertainty in this risk assessment.

The overall analysis in the O_3 NAAQS risk assessment is relatively complex, thereby warranting consideration of a full probabilistic (WHO Tier 3) uncertainty analysis. However, limitations in available information prevent this level of analysis from being completed at this

[24] It is also important to point out that failure to characterize variability in an input used in modeling can also introduce uncertainty into the analysis. This reflects the important link between uncertainty and variability with the effort to accurately characterize variability in key model inputs actually reflecting an effort to reduce uncertainty about population means and population variability.

time. In particular, the incorporation of uncertainty related to key elements of C-R functions (e.g., competing lag structures, alternative functional forms, etc.) into a full probabilistic WHO Tier 3 analysis would require that probabilities be assigned to each competing specification of a given model element (with each probability reflecting a subjective assessment of the probability that the given specification is the "correct" description of reality). However, for many model elements there is insufficient information on which to base these probabilities. One approach that has been taken in such cases is expert elicitation; however, this approach is resource- and time-intensive and consequently, it was not feasible to use this technique in the current O_3 NAAQS review to support a WHO Tier 3 analysis.[25]

For most elements of this risk assessment, rather than conducting a full probabilistic uncertainty analysis, we have included qualitative discussions of the potential impact of uncertainty on risk results (WHO Tier1). As discussed in section 7.1.1, for this draft of the risk assessment, we have also expanded the sensitivity analysis considerably to cover a range of model elements (this represents a WHO Tier 2 analysis). The specific modeling elements covered in the sensitivity analysis for each health effects endpoint together with the specification of the core analysis is presented in section 7.4.3. As part of the sensitivity analysis, we have also completed an influence analysis using estimated elasticities of response[26] designed to determine which of the input factors used in calculating risk are primarily responsible for inter-city variability in risk. This influence analysis focuses on the response of core short-term exposure-related mortality risk to inputs since this is one of the key risk metrics completed for the REA (see section 7.4.3).

In addition to the qualitative and quantitative treatment of uncertainty and variability which are described here, we have also completed an analysis to evaluate the representativeness of the selected urban study areas against national distributions for key O_3 risk-related attributes to determine whether they are nationally representative or more focused on a particular portion of the distribution for a given attribute (see Chapter 8, section 8.2.1). In addition, we have completed a second analysis addressing the representativeness issue, which identified where the 12 urban study areas included in this risk assessment fall along a distribution of national-level short-term and long-term exposure-related mortality risk (see Chapter 8, section 8.2.2). This analysis allowed us to assess the degree of which the 12 urban study areas capture locations within the U.S. likely to experience elevated levels of risk related to O_3 exposure (for both short-term and long-term O_3-attributable mortality).

[25] While a full probabilistic uncertainty analysis was not completed for this risk assessment, we were able to use confidence intervals associated with effects estimates (obtained from epidemiological studies) to incorporate statistical uncertainty associated with sample size considerations in the presentation of risk estimates.

[26] Elasticities are a measure of sensitivity calculated as the percent change in the response variable for a one percent change in the input variable.

The remainder of this section is organized as follows. Key sources of variability which are reflected in the design of the risk assessment, along with sources excluded from the design, are discussed in section 7.4.1. A qualitative discussion of key sources of uncertainty associated with the risk assessment (including the potential direction, magnitude and degree of confidence associated with our understanding of the source of uncertainty – the knowledge base) is presented in section 7.4.2. The design of the core analysis and sensitivity analysis completed for each of the health effect endpoint categories modeled in the risk assessment is discussed in section 7.4.3.

7.4.1 Treatment of Key Sources of Variability

The risk assessment was designed to cover the key sources of variability related to population exposure and exposure response, to the extent supported by available data. Here, the term *key sources of variability* refers to those sources that we believe have the potential to play an important role in impacting population incidence estimates generated for this risk assessment. Specifically, hawse have concluded that these sources of variability, if fully addressed and integrated into the analysis, could result in adjustments to the core risk estimates which might be relevant from the standpoint of interpreting the risk estimates in the context of the O_3 NAAQS review. The identification of sources of variability as "key" reflects consideration for sensitivity analyses conducted for previous O_3 NAAQS risk assessments, which have provided insights into which sources of variability can influence risk estimates, as well as information presented in the O_3 ISA.

As with all risk assessments, there are sources of variability which have not been fully reflected in the design of the risk assessment and consequently introduce a degree of uncertainty into the risk estimates. While different sources of variability were captured in the risk assessment, it was generally not possible to separate out the impact of each factor on population risk estimates, since many of the sources of variability are reflected collectively in a specific aspect of the risk model. For example, inclusion of urban study areas from different regions of the country likely provides some degree of coverage for a variety of factors associated with O_3 risk (e.g., air conditioner use, differences in population commuting and exercise patterns, weather). However, the model is not sufficiently precise or disaggregated to allow the individual impacts of any one of these sources of variability on the risk estimates to be characterized.

Key sources of potential variability that are likely to affect population risks are discussed below, including the degree to which they are captured in the design of the risk assessment:

- **Heterogeneity in the Effect of O_3 on Health Across Different Urban Areas:** A number of studies cited in the O_3 ISA have found evidence for regional heterogeneity in the short-term O_3-attributable mortality effect (Smith et al., 2009 and Bell and

Dominici, 2008, Bell et al., 2004, Zanobetti an Schwartz 2008 – see O_3 ISA section 6.6.2.2, U.S. EPA, 2013a). These studies have demonstrated that differences in effect estimates between cities can be quite substantial (see O_3 ISA Figures 6-32 and 6-33). Therefore, for the short-term O_3-attributable mortality endpoint modeled using Smith et al., 2009-based effect estimates, we have included Bayes-adjusted city-specific effect estimates reflecting application of both a regional- and national-prior, both of which are intended to capture cross-city differences in effect estimates for the mortality endpoint, while still reflecting input from the more stable regional, or national-level signal. The national-prior based estimates are included in the core analysis since they have greater overall power, while the regional-prior based estimates are included as sensitivity analyses to explore the impact of using regional prior in developing the Bayes-adjusted estimates (see section 7.4.3).[27] For short-term morbidity endpoints, typically we have used city-specific effect estimates; however, for most endpoints, we only have estimates for a subset of the urban study areas (typically NYC, Atlanta and/or LA). Therefore, although our risk estimates do reflect the application of city-specific effect estimates, because we do not have estimates for all 12 urban study areas, we do not provide comprehensive coverage for heterogeneity in modeling the respiratory morbidity endpoint category. Long-term O_3-attributable mortality has been shown to demonstrate regional heterogeneity. Specifically, Jerrett et al., 2009 presented regional effect estimates that demonstrated considerable heterogeneity ranging from essentially a no-effect (for the Northeast and Industrial Midwest) to effects substantially larger than the national effect (Southeast, Southwest and Upper Midwest) (see Table 4 in Jerrett et al., 2009). There are many potential explanations for regional heterogeneity including differences in O_3-attributable factors and potential confounding, potential for the presence of (and regional differences in) averting behavior, and variation in sample sizes which can impact stability of effect estimates. For the core analysis, we use a national effect estimate in modeling long-term exposure related mortality. Consideration of regional effect estimates are included as a sensitivity analysis (see section 7.4.3 and 7.5.3).

- **Exposure Measurement Error Associated with O_3 Effect Estimates:** Exposure measurement error refers to uncertainty associated with using ambient monitor based exposure surrogate metrics to represent the actual exposure of an individual or population. As such, this factor can be an important contributor to variability in epidemiological study results across locations, and uncertainty in results for any

[27] Note, that in some instances, there may be insufficient variance between cities to generate city-specific estimates using a regional prior, which compromises their use in the core analysis.

specific city (O_3 ISA, p. 1xii). Exposure measurement error can result from a number of factors (e.g., central site monitors not representing actual patterns of personal exposure including activity patterns, presence of non-ambient sources of exposure for the pollutant of interest) (O_3 ISA, 1xii). These factors can vary across urban study areas (and even within urban study areas), thereby contributing to differences in the nature and magnitude of exposure measurement error across locations and ultimately to differences in effect estimates and associated confidence levels. Exposure measurement error is related to heterogeneity in effect estimates, since regional differences in effect estimates can result in part, from differences in exposure measurement error as noted here.

- **Intra-urban Variability in Ambient O_3 Levels**: The picture with regard to within city variability in ambient O_3 levels and the potential impact on epidemiologic-based effect estimates is somewhat more complicated. The O_3 ISA notes that spatial variability in O_3 levels is dependent on spatial scale with O_3 levels being more homogeneous over a few kilometers due to the secondary formation nature of O_3, while levels can vary substantially over tens of kilometers. Community exposure may not be well represented when monitors cover large areas with several subcommunities having different sources and topographies as exemplified by Los Angeles which displays significantly greater variation in inter-monitor correlations than does, for example, Atlanta or Boston (see O_3 ISA section 4.6.2.1 U.S. EPA 2013a). Despite the potential for substantial variability across monitors the O_3 ISA notes that studies have tended to demonstrate that monitor selection has only a limited effect on the association of short-term O_3 exposure with health effects. The likely explanation for this is that, while absolute values for a fixed point in time can vary across monitors in an urban area, the temporal patterns of O_3 variability across those same monitors tends to be well correlated. Given that most of the short-term O_3-attributable O_3 epidemiological studies are time series in nature, the O_3 ISA notes that the stability of temporal profiles across monitors within most urban areas means that monitor selection will have little effect on the outcomes of an epidemiological study examining short-term O_3-attributable mortality or morbidity (see O_3 ISA section 4.6.2.1 U.S. EPA 2013a). For this reason, we conclude that generally intra-city heterogeneity in O_3 levels is not a significant factor likely to impact estimates of short-term O_3-attributable risk. One exception is LA which, due to its size and variation in O_3 sources and other factors impacting O_3 patterns such as topography, displays significant variation in ambient O_3 levels with a subsequent impact on risk. However, in the case of LA (as with the other urban study areas), we model risk using

composite monitors which do not provide spatially-differentiated representations of exposure and consequently, we do not address this source of variability in the risk assessment. As discussed in the uncertainty section, short-term exposure mortality effect estimates for the New York CBSA (Smith et al., 2009) display significant variability. However, it is not clear which factors are primarily responsible for this heterogeneity (e.g., differences in the urban structure, residential behavior, or ambient O_3 levels within the CBSA). The potential for intra-city heterogeneity in O_3 levels to affect risk is more pronounced with long-term O_3-attributable mortality where the relationship between annual trends in ambient O_3 (as represented using composite monitor values) and annual mortality is compared between urban study areas in order to derive effect estimates. Here, pronounced heterogeneity in O_3 levels within a given city can result in exposure misclassification, if that heterogeneity is not well represented by the composite monitor for that city. Different degrees of exposure misclassification across urban study areas can introduce uncertainty into the overall national-level effect estimate for long-term exposure-related mortality. Furthermore, if that exposure measurement error has a regional trend, then measurement error can potentially result in apparent regional heterogeneity in the effect estimates. The degree to which there is true regional heterogeneity is made uncertain by the presence of differential measurement error across regions.

- **Variability in the Patterns of Ambient O_3 Reduction Across Urban Areas:** The simulated patterns of ambient O_3 concentrations across an urban area can vary based on the methodology used to adjust ambient O_3 concentrations to represent just meeting the current or alternative suites of standards. For the 1st draft REA, we used a statistical approach called the "quadratic rollback" method for simulating just meeting the current O_3 standard. Although the quadratic rollback method replicates historical patterns of air quality changes better than some alternative methods, its implementation relies on a statistical relationship instead of on a mechanistic characterization of physical and chemical processes in the atmosphere. Because of its construct as a statistical fit to measured O_3 values, the quadratic rollback technique cannot capture spatial and temporal heterogeneity in O_3 response and also cannot account for nonlinear atmospheric chemistry that causes increases in O_3 as a result of emissions reductions of certain O_3 precursors under some circumstances. As noted in section 7.1.1, for this draft of the REA, we have employed a model-based O_3 adjustment methodology in the risk assessment for simulating O_3 concentrations under current and alternate standard levels. Use of this model-based approach allows the risk assessment results to more fully account for non-linearities in O_3 formation

and to reflect spatial and temporal heterogeneity in O_3 response, including NOx titration conditions under which a reduction in NOx causes an increase in O_3 concentrations, in some core urban locations.

- **Demographics and Socioeconomic-status (SES)-related Factors:** Variability in population density, particularly in relation to elevated levels of O_3 has the potential to influence population risk, although the significance of this factor also depends on the degree of intra-urban variation in O_3 levels (as discussed above). In addition, community characteristics such as pre-existing health status, ethnic composition, SES and the age of housing stock (which can influence rates of air conditioner use thereby impacting rates of infiltration of O_3 indoors) can contribute to observed differences in O_3-related risk (discussed in O_3 ISA – section 2.5.4.5, U.S. EPA, 2013a). Some of the heterogeneity observed in effect estimates between cities in the multicity studies may be due to these community characteristics, and while we cannot determine how much of that heterogeneity is attributable to these factors, the degree of variability in effect estimates between cities in our analysis should help to capture some of the latent variability in these community characteristics.
- **Baseline Incidence of Disease**: We collected baseline health effects incidence data (for mortality and morbidity endpoints) from a number of different sources (see section 7.3.4). Often the data were available at the county-level, providing a relatively high degree of spatial refinement in characterizing baseline incidence given the overall level of spatial refinement reflected in the risk assessment as a whole. Otherwise, for urban study areas without county-level data, either (a) a surrogate urban study area (with its baseline incidence rates) was used, or (b) less refined state-level or national incidence rate data were used.

7.4.2 Qualitative Assessment of Uncertainty

As noted in section 7.4, we have based the design of the uncertainty analysis carried out for this risk assessment on the framework outlined in the WHO guidance document (WHO, 2008). That guidance calls for the completion of a Tier 1 qualitative uncertainty analysis, provided the initial Tier 0 screening analysis suggests there is concern that uncertainty associated with the analysis is sufficient to significantly impact risk results (i.e., to potentially affect decision making based on those risk results). Given previous sensitivity analyses completed for prior O_3 NAAQS reviews, which have shown various sources of uncertainty to have a potentially significant impact on risk results, we believe that there is justification for conducting a Tier 1 analysis.

For the qualitative uncertainty analysis, we have described each key source of uncertainty and qualitatively assessed its potential impact (including both the magnitude and direction of the impact) on risk results, as specified in the WHO guidance. Similar to our discussion of variability in the last section, the term *key sources of uncertainty* refers to those sources that the we believe have the potential to play an important role in impacting population incidence estimates generated for this risk assessment (i.e., these sources of uncertainty, if fully addressed could result in adjustments to the core risk estimates which might impact the interpretation of those risk estimates in the context of the O_3 NAAQS review). These key sources of uncertainty have been identified through consideration for sensitivity analyses conducted for previous O_3 NAAQS risk assessments, together with information provided in the final O_3 ISA and comments provided by CASAC on the analytical plan for the risk assessment.

Table 7-4 includes the key sources of uncertainty identified for the O_3 REA. For each source of uncertainty, we have (a) provided a description, (b) estimated the direction of influence (over, under, both, or unknown) and magnitude (low, medium, high) of the potential impact of each source of uncertainty on the risk estimates, (c) assessed the degree of uncertainty (low, medium, or high) associated with the knowledge-base (i.e., assessed how well we understand each source of uncertainty), and (d) provided comments further clarifying the qualitative assessment presented.

The categories used in describing the potential magnitude of impact for specific sources of uncertainty on risk estimates (i.e., low, medium, or high) reflect our consensus on the degree to which a particular source could produce a sufficient impact on risk estimates to influence the interpretation of those estimates in the context of the O_3 NAAQS review.[28] Sources classified as having a "low" impact would not be expected to impact the interpretation of risk estimates in the context of the O_3 NAAQS review; sources classified as having a "medium" impact have the potential to change the interpretation; and sources classified as "high" are likely to influence the interpretation of risk in the context of the O_3 NAAQS review. Because this classification of the potential magnitude of impact of sources of uncertainty is not based on our direct quantitative assessments, we use qualitative judgments, in some cases informed by other relevant quantitative analyses. Therefore, the results of the qualitative analysis of uncertainty are not useful for making quantitative estimates of confidence, e.g. probabilistic statements about risk. However, they can be used to support the interpretation of the risk estimates, including the assessment of overall confidence in the risk estimates. In addition, they can also be used in guiding future research to reduce uncertainty related to O_3 risk assessment. As with the qualitative discussion of

[28] For example, if a particular source of uncertainty were more fully characterized (or if that source was resolved, potentially reducing bias in a core risk estimate), could the estimate of incremental risk reduction in going from the current to an alternative standard level change sufficiently to produce a different conclusion regarding the magnitude of that risk reduction in the context of the O_3 NAAQS review?

sources of variability included in the last section, the characterization and relative ranking of sources of uncertainty addressed here is based on our consideration of information provided in previous O_3 NAAQS risk assessments (particularly past sensitivity analyses), the results of risk modeling completed for the current O_3 NAAQS risk assessment and information provided in the third draft O_3 ISA as well as earlier O_3 Criteria Documents. Where appropriate, in Table 7-4, we have included references to specific sources of information considered in arriving at a ranking and classification for a particular source of uncertainty.

Table 7-4 Summary of Qualitative Uncertainty Analysis of Key Modeling Elements in the O_3 NAAQS Risk Assessment

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
A. Adjustment of recent air quality measurements of O_3 to simulate attainment of both existing and alternative standard levels	See Chapter 4 for details	Both	Low-Medium	Low-medium	See Chapter 4 for more details (specific call-outs to be added)
B. Use of CBSA-based study areas in modeling risk (i.e., potential mismatch between study areas used in the REA and study areas used in the epidemiological studies providing the effect estimates used in modeling health effect endpoints)	If the set of monitors used in a particular urban study area to characterize population exposure as part of an ongoing risk assessment do not match the ambient monitoring data used in the original epidemiological study, then uncertainty can be introduced into the risk estimates. This uncertainty is balanced in part by the reduction in bias that results from using the expanded CBSA definition. (See section 7.1.1 for more details.) However, it should be noted that because these epidemiological studies occurred in the past, sometimes it can be impossible to exactly match the monitors used in the study using recent air quality data given that monitors may have moved to a different location or there may not be measurements available at specific monitors in the more recent time period.	Both	Low-medium	Low-medium	KB and INF: In modeling risk for the current draft of the REA, we used CBSA-based study areas for all health effect endpoints. As discussed in section 7.1.1, the use of the larger CBSA study areas allows us to better reflect how the change in air quality affects risk across the entire urban area and to avoid introducing known bias into the REA by focusing risk estimates on that subpopulation living in areas likely to experience potential increases in O_3 (and excluding the larger population of urban and suburban areas likely to experience reductions in ambient O_3 levels). While the use of the larger CBSA-based study areas addresses this source of known bias, it also introduces uncertainty into the REA since we are no longer matching the REA study areas to the study areas in the epidemiological studies providing the effect estimates used in modeling health effects endpoints. Given available data, it is not possible at this point to reliably characterize the degree of uncertainty introduced into the REA by having this mismatch in study areas. However, the potential bias avoided through the use of the larger CBSA study areas (with its acknowledged uncertainty) is substantial, as illustrated in the sensitivity analyses exploring spatial study area (see section 7.5.3).
C. Application of C-R functions based on a specific temporal and spatial pattern of correlations between O_3 monitors in an urban area (as reflected in the	The effect estimates used in this risk assessment reflect a specific spatial and temporal pattern of ambient O_3 (as represented by the particular monitoring network providing data for the underlying epidemiological study). However, if the spatial and	Both	Low-medium	Low-medium	KB and INF: With application of the HDDM adjustment approach, we simulate potential changes in the spatial and temporal pattern of O_3 for a study areas when just meeting the existing and alternative standards relative to patterns under recent conditions. This introduces uncertainty into the application of the original effect estimates, since the exposure surrogate represented by the composite monitor values may no longer match that of the underlying epidemiological study. However, it is not

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
epidemiological study providing the effect estimates) to a simulated change in the patterns of those correlations when we estimate risk in the REA.	temporal pattern of O_3 levels in the study areas being modeled differ significantly from the patterns in the original epidemiological study (for those same study areas), then uncertainty can be introduced into the risk estimates.				possible, given available data, to characterize quantitatively the magnitude of this uncertainty. This is probably most true in the urban areas of New York and Los Angeles where simulation meeting the existing and alternative standards using the HDDM-adjustment approach relied on large NOx reductions and there is very little day-to-day variability in the resulting O_3 concentrations.
D. Characterizing intra-urban population exposure in the context of epidemiology studies linking O_3 to specific health effects	Exposure misclassification within communities that is associated with the use of generalized population monitors (which may miss important patterns of exposure within urban study areas) introduces uncertainty into the effect estimates obtained from epidemiology studies.	Under (generally)	Low-medium	Medium	KB and INF: Despite the potential for substantial variability in O_3 levels across monitors (particularly in larger urban areas with greater variation in sources and topography such as L.A.), the O_3 ISA notes that studies have tended to demonstrate that monitor selection has only a limited effect on the association of short-term O_3 exposure with health effects (see O_3 ISA section 4.6.2.1, US EPA, 2013a). However, this issue could be more of a concern in larger urban areas which may exhibit greater variation in O_3 levels due to diverse sources, topography and patterns of commuting.
E. Statistical fit of the C-R functions	Exposure measurement error combined with other factors (e.g., size of the effect itself, sample size, control for confounders) can effect the overall level of confidence associated with the fitting of statistical effect-response models in epidemiological studies.	Both	Medium (short-term health endpoints)	Medium	INF: For short-term mortality and morbidity health endpoints, there is greater uncertainty associated with the fit of models given the smaller sample sizes often involved, difficulty in identifying the etiologically relevant time period for short-term O_3 exposure, and the fact that models tend to be fitted to individual counties or urban areas (which introduces the potential for varying degrees of confounding and effects modification across the locations). These studies can also have effects estimates that are not statistically significant. For this risk assessment, in modeling short-term mortality, we are not relying on location-specific models. Instead, we are using city-specific effects estimates derived using Bayesian techniques (these combine national-scale models with local-scale models). Exposure measurement error (uncertainty associated with the exposure metrics used to represent exposure of an individual or population) can also be an important contributor to uncertainty in effect estimates associated both with short-term and long-term O_3-attributable studies (O_3 ISA, p. 1xii). Together with other factors (e.g., low data density), exposure measurement error can result in the smoothing of epidemiologically-derived response functions and the obscuring of thresholds should they exist (O_3 ISA, p. Ixix). In addition, exposure measurement error can vary across different populations even within the same urban study area. For

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
					example a particular group could have an activity pattern that results in central site monitors (in that urban study area) being particularly poor at representing that group's exposure to ambient O_3. In this example, an effect estimate derived for that specific population based on O_3 exposure characterized using central site monitors would have increased uncertainty relative to effect estimates generated for other population with different activity patterns and lower levels of exposure measurement error.
F. Shape of the C-R functions	Uncertainty in predicting the shape of the C-R function, particularly in the lower exposure regions which are often the focus in O_3 NAAQS regulatory reviews.	Both	Medium	Low-medium	KB and INF: Studies reviewed in the O_3 ISA that attempt to characterize the shape of the O_3 C-R curve along with possible "thresholds" (i.e., O_3 concentrations which must be exceeded in order to elicit an observable health response) have indicated a generally linear C-R function with no indication of a threshold (for analyses that have examined 8-h max and 24-h avg O_3 concentrations). However, the ISA notes that the studies from which the C-R functions are derived indicate there is less certainty in the shape of the C-R curve at the lower end of the distribution of O_3 concentrations (in the range below 20 ppb) due to the low density of data in the studies in this range.
G. Addressing co-pollutants	The inclusion or exclusion of co-pollutants which may confound, or in other ways, affect the O_3 effect, introduces uncertainty into the analysis.	Both	Low-medium	Medium	KB and INF: The O_3 ISA notes that across studies, the potential impact of PM indices on O_3-mortality risk estimates tended to be much smaller than the variation in O_3-mortality risk estimates across cities. This suggests that O_3 effects are independent of the relationship between O_3 and mortality. However, interpretation of the potential confounding effects of PM on O_3-mortality risk estimates requires caution. This is because the PM-O_3 correlation varies across regions, due to the difference in PM components, complicating the interpretation of the combined effect of PM on the relationship between O_3 and mortality. Additionally, the limited PM or PM component datasets used as a result of the every-3rd- and 6th-day PM sampling schedule instituted in most cities limits (in most cases) the overall sample size employed to examine whether PM or one of its components confounds the O_3-mortality relationship (O_3ISA section 2.5.4.5, US EPA, 2013a).
H. Specifying lag structure (short-term exposure studies)	There is uncertainty associated with specifying the exact lag structure to use in modeling short-term O_3-	Both	Low-Medium	Low	KB and INF: The majority of studies examining different lag models suggest that O_3 effects on mortality occur within a few days of

Source	Description	Potential influence of uncertainty on risk estimates		Knowledge-Base uncertainty*	Comments (KB: knowledge base, INF: influence of uncertainty on risk estimates)
		Direction	Magnitude		
	attributable mortality and respiratory-related morbidity.				exposure. Similar, studies examining the impact of O_3 exposure on respiratory-related morbidity endpoints suggests a rather immediate response, within the first few days of O_3 exposure (see O_3ISA section 2.5.4.3, US EPA, 2013a). Consequently, while the exact nature of the ideal lag models remains uncertain, generally, we are fairly confident that they would be on the order of a day to a few days following exposure.
I. Using studies from one geographic area to cover urban areas outside of the study area	In the case of Gent et al., 2003 (used in modeling asthma exacerbations in Boston), we are using C-R functions based on an epidemiological study of a region (northern Connecticut and Springfield) that does not encompass the actual urban study area assessed for risk (Boston).	Both	Medium	Low	INF: Factors related to O_3 exposure including commuting patterns, exercise levels etc may differ between the region reflected in the epidemiological study and Boston. If these differences are great, then applying the effect estimate from the epidemiological study to Boston could be subject to considerable uncertainty and potential bias.
J. Characterizing baseline incidence rates	Uncertainty can be introduced into the characterization of baseline incidence in a number of different ways (e.g., error in reporting incidence for specific endpoints, mismatch between the spatial scale in which the baseline data were captured and the level of the risk assessment).	Both	Low-medium	Low	INF: The degree of influence of this source of uncertainty on the risk estimates likely varies with the health endpoint category under consideration. There is no reason to believe that there are any systematic biases in estimates of the baseline incidence data. The influence on risk estimates that are expressed as incremental risk reductions between alternative standards should be relatively unaffected by this source of uncertainty. KB: The county level baseline incidence and population estimates at the county level were obtained from data bases where the relative degree of uncertainty is low.

* Refers to the degree of uncertainty associated with our understanding of the phenomenon, in the context of assessing and characterizing its uncertainty (specifically in the context of modeling PM risk)

7.4.3 Description of Core and Sensitivity Analyses

As discussed in section 7.1.1, this risk assessment includes a set of core (higher confidence) risk estimates which are supplemented by sensitivity analyses. The sensitivity analyses explore the potential impact that variation in specific model design elements can have on the core risk estimates. This section specifies which design elements are included in both the core and sensitivity analyses completed for each of the health effect endpoint categories included in the risk assessment. We divided the sensitivity analyses into two categories: (a) those involving air quality characterization and (b) those associated directly with the specification of the C-R functions used in estimating risk. We recognize that there can be overlap between these categories with some modeling elements (e.g., modeling period) affecting both the composite monitor distribution as well as representing an element of C-R function specification. However, we have retained these two categories to aid in the presentation and discussion of sensitivity analysis results.[29] The sensitivity analyses also included an initial influence analysis designed to evaluate which of the model inputs are primarily responsible for inter-city variability (heterogeneity) in risk. The influence analysis uses estimatedelasticities of risk with respect to the risk function input variables, focusing on the short-term exposure-related mortality endpoint and associated input parameters since this is one of the key risk estimates generated for the REA (additional detail on how the influence analysis was conducted is presented in section 7.5.3).

Table 7-5 presents the alternative approaches for adjusting the O_3 distributions used in the sensitivity analysis and also identifies the approaches used in the core analysis for each of the study areas. The alternative air quality adjustment approaches examine the differences in changes in air quality and risk when applying NOx-only versus NOx and VOC reductions in the HDDM-adjustment approach. It should be noted that when NOx and VOC reductions were used in the HDDM-adjustment approach in this sensitivity analysis, the same percent reduction for both pollutants was used in the air quality adjustment for meeting the existing and alternative standard in each urban area. More details on these alternative air quality adjustment approaches are discussed in Chapter 4 and appendices.

Besides the approach used to adjust the distributions of O_3, another fact which has a direct impact on composite monitor composition is the specification of the study area (since this determines the mix of monitors that will be included in constructing the composite monitor). As discussed in section 7.1.1, for the core analysis, we modeled all endpoints (for all study areas)

[29] As noted in 7.1.1, in presenting both the core and sensitivity analyses, we include both point estimates and 95th CIs, with the latter reflecting the statistical fit of the effect estimates (and hence the power of the underlying epidemiological study). In comparing core and sensitivity analyses, we not only focus on point estimates, but also on the CIs since they provide insights into differences in the degree of statistical support for the effect estimates underlying the risk estimates and therefore, overall confidence in those estimates.

using CBSA-based study areas. For the sensitivity analysis (for the short-term O_3-attributable mortality endpoint), we included a smaller study area based on the original study area definition used in the Smith et al., 2009 study.[30]

Table 7-6 presents the model elements included in sensitivity analyses exploring alternative C-R function specifications. These sensitivity analyses were applied both to short-term O_3-attributable mortality and long-term O_3-attributable mortality. As discussed in section 7.1.1, we were not able to differentiate between alternative C-R function specifications for short-term O_3-attributable morbidity endpoints and therefore included the full set of alternative C-R function specifications in the core analysis. This results in a distribution of core risk estimates for each endpoint which can be used to gain insights into the impact of different C-R function specifications on risk. Because separate sensitivity analyses were not completed for short-term O_3-attributable morbidity endpoints, this category is not included in Table 7-6.

Table 7-5 Specification of the Core and Sensitivity Analyses (air quality simulation)

Study Area	Core simulation (type of precursor reduced to adjust O_3 distribution)	Sensitivity analysis
Atlanta, GA	NOx-only	Alternative modeling approach not evaluated
Baltimore, MD	NOx-only	
Boston, MA	NOx-only	
Cleveland, OH	NOx-only	
Denver, CO	**NOx & VOC**	NOx-only
Detroit, MI	NOx-only	NOx & VOC
Houston, TX	NOx-only	NOx & VOC
Los Angeles, CA	**NOx-lower bound***	NOx & VOC-lower bound*
New York, NY	**NOx-lower bound* (exclude 60)**	NOx& VOC-lower bound*
Philadelphia, PA	NOx-only	NOx & VOC
Sacramento, CA	NOx-only	NOx & VOC

[30] We did not include an alternative study area simulation as a sensitivity analysis for long-term exposure related mortality, since we are using a single (national) effect estimate in modeling this endpoint, and consequently, the use of an effect estimate from a smaller study area to represent a somewhat larger area (as is the case with short-term O_3-attributable mortality) is likely to introduce less uncertainty.

Study Area	Core simulation (type of precursor reduced to adjust O_3 distribution)	Sensitivity analysis
St. Louis, MO	NOx-only	Alternative modeling approach not evaluated

A lower-bound fit of the HDDM-based O_3 sensitivities (reflecting a greater increment of O_3 reduction per unit of VOC and/or NOx reduction) was required in simulation of the alternative standard levels.

Table 7-6 Specification of the Core and Sensitivity Analyses (alternative C-R function specification)

Health effect endpoint category	Modeling elements included	
	Core analysis	Sensitivity analysis
Short-term O_3-attributable mortality	- Full monitoring period (specific to each study area), 8hr max metric, national-Bayes adjusted, single pollutant model. - effect estimates obtained from: Smith et al., 2009 study	- summer (warm month), 8hr mean, regional-bayes adjusted, multi-pollutant (with PM_{10}). - effect estimates obtained from Zanobetti and Schwartz, 2008 and Smith et al., 2009
Long-term O_3-attributable mortality	- Single national estimate, two-pollutant model (PM2.5), long-term peak trend metric (based on daily 1hr max values), CBSA-based study area. - effect obtained from Jerrett et al., 2009 study	- Regional-differentiated effect estimates, single pollutant model. - National-level effect estimate, single pollutant model. - effect estimates also obtained from Jerrett et al., 2009 study

7.5 URBAN STUDY AREA RESULTS

This section discusses risk estimates generated for the set of 12 urban study areas, including both the core risk estimates and accompanying sensitivity analyses. In summarizing risk estimates, this discussion focuses on results most relevant to two policy-related questions: (a) to what extent is the existing O_3 standard protective of public health , and, (b) what is the nature and magnitude of additional public health protection provided by the suite of alternative standards under consideration? Consequently, we focus on two types of risk estimates including the magnitude of O_3-attributable risk after simulation of just meeting the existing standard and

the degree of risk reduction potentially provided by each of the alternative standards relative to just meeting the existing standard.[31]

This section is organized as follows. We begin by presenting the core risk estimates in both tabular and graphical format at the end of this section. We then present key observations about the risk estimates for just meeting the existing standard (for core risk) in section 7.5.1. Key observations related to risk estimates for just meeting alternative standard levels, and for estimates of risk changes comparing alternative standards to just meeting the existing standard (again, for core risk) are presented in section 7.5.2. After presenting key observations related to the core risk estimates, we then present key observations resulting from the sensitivity analyses (section 7.5.3).

A number of details regarding the design of the core risk assessment should be kept in mind when reviewing the core risk estimates presented in this section (see section 7.1.1 for additional detail on these design elements):

- All risk estimates reflect application of a CBSA-based study area.
- Estimates are presented for two simulation years (2007 and 2009):
- Short-term O_3-attributable mortality estimates are generated for all 12 urban study areas, while most short-term O_3-attributable morbidity estimates (depending on the specific health endpoint) are generated for only a subset of urban study areas. Long-term O_3-attributable mortality is modeled for all 12 urban study areas.
- For all health effect endpoints, we model risk down to zero O_3 and do not include either consideration for LML or alternative threshold levels.

There are several categories of risk metrics generated for the core mortality and morbidity endpoints modeled in this analysis. Below we describe both the types of risk metrics generated for the core analysis and the specific types of tables and figures used in presenting those metrics.

I. Core short-term O_3-attributable mortality estimates

- **Table presenting estimates of O_3-attributable mortality incidence after just meeting the existing standard and the estimated change in incidence associated with meeting each of the alternative standard levels relative to the existing standard (Table 7-7):** These estimates include point estimates and 95th percentile confidence intervals representing uncertainty associated with the statistical fit of the effect estimates.

[31] As part of this draft of the risk assessment, we have also generated estimates of risk under recent conditions as well as estimates of the degree of risk reduction (relative to risk under recent conditions) associated with the simulated attainment of the existing standard. See Appendix 7B.

- **Table presenting estimates of the percent of total mortality attributable to O_3 after just meeting the existing standard and the percent reduction in O_3-attributable risk associated with each alternative standard (Table 7-8).**

- **Heat maps for mortality illustrating distribution across daily O_3 levels of total O_3-attributable risk after just meeting the existing standard and risk reductions after meeting alternative standards (Figures 7-2 and 7-3):** Heat maps are provided for each of the 12 urban areas. The color gradient in each figure reflects the distribution of mortality (or the change in mortality) across the range of daily 8-hour O_3 levels and provides a visual tool to explore trends in mortality counts across daily O_3 levels and between cities. Visual patterns in the figures presenting total risk and risk reduction are interpreted differently:

  - For figures depicting total O_3-attributable risk, colors range from blue (lower mortality count) to red (higher mortality count). As an example, with Figure 7-2, top heat map (which presents total O_3-attributable risk for the existing standard in 2007, based on Smith et al., 2009 C-R functions), if we focus on the first row (Atlanta, GA), we see a value of 38 under the column 55-60 ppb. This value reflects the fact that 38 of the 270 O_3-attributable deaths estimated for Atlanta after just meeting the existing 75 ppb standard in 2007 occurred on days with composite monitor O_3 levels between 55 and 60 ppb. Similarly, in the same row, we see that only 3 O_3 attributable deaths occurred on days when the composite monitor value was between 20 and 25 ppb. We also include the total O_3-attributable incidence (for each study area) in the final column marked "Total".

  - For figures depicting changes in risk associated with simulation of existing and alternative standard levels, we see that the pattern is more complex since we can have a combination of increases and decreases in risk in the heat maps, with increases in risk identified as red to yellow and decreases in risk identified as yellow to blue. Increases in risk are negative numbers, decreases are positive. In addition, in the final three columns of each map, we provide estimates of the total O_3-attributable incidence, as well as the total broken down into the subtotals across days with increases (negative) and days with decreases (positive) in that incidence. The increase and decrease for a given study area should sum (accounting for rounding in these subtotals) to the overall total for O_3-attributable deaths for that study area. Several factors can contribute to the patterns of changes in O_3-attributable risk reflected in these maps. For example, non-linearities in O_3

formation can result in increases in O_3 on some days, even when simulating attainment of a lower alternative standard (see section 7.1.1). In addition, simulation of alternative standard levels can result in a change in the overall distribution of the composite monitor ambient O_3 distribution. Often, this change will take the form of a shift in the upper tail of the distribution towards the mean, given that simulated attainment of alternative standard levels targets higher O_3 days. If we look at figure 7-2 at the second map (Decrease 75 to 70) and specifically at the row for Houston, we see that there is a -4 increase in deaths distributed across 20-35 ppb days and a decrease in deaths of 9, primarily distributed across 40-60 ppb days.

- **Graphic plots of O_3-attributable deaths per 100,000 population for just meeting the existing and alternative standards (Figures 7-4):** O_3This plot provides estimates that are adjusted for the size of the underlying urban population, thereby allowing the mortality estimates and associated trends to be more readily compared across urban study areas (consideration of absolute O_3 mortality is complicated by the role that underlying urban population plays in driving total O_3-attributable mortality – larger study areas like Los Angeles and New York having substantially larger mortality estimates primarily due to their higher underlying populations). These figures allow us to evaluate the overall magnitude of risk reductions across standard levels and determine the degree to which those trends differ for different study areas.

Tables summarizing incidence, percent of baseline incidence and percent reduction in O_3-attributable risk for short-term O_3-attributable morbidity (Tables 7-9 through 7-11): Three categories of short-term O_3-attributable mortality effects were modeled for the analysis (respiratory related HA, respiratory-related ER visits and asthma exacerbations). As discussed in section 7.1.1, these morbidity effects were modeled for a combination of all 12 urban study areas and a subset of those study areas depending on the endpoint (see below). The C-R functions available for modeling many of these morbidity endpoints included consideration for a number of design elements (e.g., copollutants and lag structure). However, as noted earlier in section 7.1.1, for short-term exposure morbidity endpoints with multiple C-R functions, we were not able to differentiate between C-R functions in terms of overall confidence and consequently we could not identify a single core model. Therefore, when we have multiple C-R functions reflecting different treatments of key design elements such as lag structure, we consider the risk estimates that result from the full set of C-R functions to represent a core range of risk. Each of the tables summarizing short-term O_3-attributable morbidity risk present several risk metrics

including: (a) total O_3-attributable incidence (after just meeting the existing standard), (b) reductions in O_3-attributable incidence (for each of the alternative standard levels relative to just meeting the existing standard), (c) percent of baseline incidence attributable to O_3 (after just meeting the existing standard) and (d) percent reductions in O_3-attributable risk (for each of the alternative standard levels). In presenting these morbidity risk estimates, we do not include 95th percentile confidence intervals in order to conserve space. Specific tables summarizing these morbidity incidence estimates include:

- HA visits (for respiratory symptoms including asthma): Table 7-9 presents estimates of the incidence of HA (for respiratory symptoms, chronic lung disease and asthma). Risk estimates are generated for a subset of the urban study areas for some of the health endpoints (e.g., New York City for HA [chronic lung disease and asthma]), while HA (respiratory-related) estimates cover all 12 urban study areas.

- ER visits (for respiratory symptoms including asthma): Table 7-10 presents estimates of the incidence of ER visits for respiratory symptoms and asthma) specifically for New York City and Atlanta based on C-R functions obtained from several epidemiological studies.

- Asthma exacerbations: Table 7-11 presents estimates of the incidence of asthma exacerbations (including estimates for a range of symptoms) for Boston, the only urban study area with C-R functions supporting modeling for this endpoint.

- **Graphic plots of O_3-attributable respiratory-related HA per 100,000 residents for the existing and alternative standard levels (Figures 7-5):** This figure is intended to complement Figure 7-4 which presents the same type of risk information for short-term O_3-attributable mortality. By plotting respiratory HA per 100,000, we adjust for the underlying population which makes trends in risk more comparable across urban study areas. We have only created this graphic for respiratory HA (based on application of Medina-Ramon et al., 2006) since that is the only morbidity endpoint modeled for all 12 urban study areas. As with the mortality figure, this figure allows us to evaluate the overall magnitude of risk reductions across standard levels and determine the degree to which those trends differ for different study areas.

III. Core long-term O_3-attributable mortality estimates

- **Table presenting estimates of long-term O_3-attributable mortality incidence including total risk after just meeting the existing standard and risk reductions based on comparing risks after meeting alternative standards to risks after meeting the existing standard (Table 7-12):** Estimates presented in Table 7-12 reflect respiratory mortality and include 95th percentile confidence intervals representing uncertainty associated with the statistical fit of the effect estimates used. Estimates presented in these tables allow for consideration for the magnitude of risk associated with just meeting the existing standard and the pattern of risk reduction (in incidence) in meeting alternative standards relative to the existing standard.

- **Table presenting estimates of the percent of respiratory mortality attributable to O_3 and percent reductions in O_3-attributable risk for long-term O_3-attributable mortality (Table 7-13).**

- **Graphic plots of O_3-attributable deaths per 100,000 population for just meeting the existing and alternative standards (Figures 7-6):** This plot provides estimates that are adjusted for the size of the underlying urban population, thereby allowing the mortality estimates and associated trends to be more readily compared across urban study areas (consideration of absolute O_3 mortality is complicated by the role that underlying urban population plays in driving total O_3-attributable mortality – larger study areas like Los Angeles and New York having substantially larger mortality estimates primarily due to their higher underlying populations). These figures allow us to evaluate the overall magnitude of risk reductions across standard levels and determine the degree to which those trends differ for different study areas.

Table 7-7 Short-Term O_3-attributable All Cause Mortality Incidence (2007 and 2009) (Smith et al., 2009 C-R Functions) (O_3 season, CBSA-based study area, no threshold)

Study Area	Air Qualtiy Scenario			
	Absolute Incidence	Change in Incidence		
	75ppb	75-70	75-65	75-60
2007 Simulation Year				
Atlanta, GA	270	12	21	34
	(-370 - 890)	(-16 - 39)	(-30 - 72)	(-47 - 110)
Baltimore, MD	440	13	27	45
	(-250 - 1100)	(-7 - 33)	(-15 - 68)	(-25 - 110)
Boston, MA	350	7	20	32
	(-500 - 1200)	(-10 - 24)	(-28 - 67)	(-45 - 110)
Cleveland, OH	430	14	32	64
	(-41 - 890)	(-1 - 28)	(-3 - 67)	(-6 - 130)
Denver, CO	86	2	4	8
	(-280 - 440)	(-6 - 10)	(-14 - 23)	(-25 - 40)
Detroit, MI	660	23	42	69
	(32 - 1300)	(1 - 44)	(2 - 81)	(3 - 130)
Houston, TX	680	5	11	24
	(130 - 1200)	(1 - 9)	(2 - 20)	(4 - 43)
Los Angeles, CA	1300	43	87	160
	(-530 - 3000)	(-18 - 100)	(-36 - 210)	(-66 - 380)
New York, NY	2800	130	640	NA
	(1700 - 3900)	(80 - 190)	(380 - 890)	NA
Philadelphia, PA	1200	35	76	120
	(270 - 2200)	(8 - 62)	(17 - 140)	(25 - 210)
Sacramento, CA	370	7	13	23
	(-390 - 1100)	(-7 - 20)	(-13 - 39)	(-24 - 70)
St. Louis, MO	430	18	39	60
	(-110 - 950)	(-5 - 41)	(-10 - 86)	(-15 - 130)
2009 Simulation Year				
Atlanta, GA	240	9	16	23
	(-340 - 800)	(-12 - 28)	(-22 - 54)	(-32 - 77)
Baltimore, MD	400	7	17	28
	(-220 - 1000)	(-4 - 19)	(-10 - 44)	(-15 - 71)
Boston, MA	320	-1	5	14
	(-450 - 1100)	(2 - -4)	(-7 - 17)	(-19 - 47)
Cleveland, OH	400	12	29	49
	(-38 - 830)	(-1 - 24)	(-3 - 60)	(-5 - 100)
Denver, CO	83	0	2	7
	(-270 - 420)	(-1 - 2)	(-6 - 10)	(-23 - 37)
Detroit, MI	580	-21	-6	15
	(28 - 1100)	(-1 - -41)	(0 - -12)	(1 - 30)
Houston, TX	700	-1	4	14
	(130 - 1200)	(0 - -1)	(1 - 7)	(3 - 26)
Los Angeles, CA	1300	41	89	160
	(-540 - 3100)	(-17 - 100)	(-37 - 210)	(-68 - 390)
New York, NY	2600	84	440	NA
	(1600 - 3700)	(50 - 120)	(260 - 610)	NA
Philadelphia, PA	1100	19	44	69
	(240 - 2000)	(4 - 34)	(10 - 78)	(15 - 120)
Sacramento, CA	370	6	12	21
	(-390 - 1100)	(-7 - 19)	(-13 - 38)	(-22 - 64)
St. Louis, MO	380	8	21	37
	(-96 - 840)	(-2 - 18)	(-5 - 46)	(-9 - 83)

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

.

Table 7-8 Percent of Total All-Cause Mortality Attributable to O_3 and Percent Change in O_3-Attributable Risk (2007 and 2009) (Smith et al., 2009 C-R functions) (O_3 season, CBSA-based study area, no threshold)

Study Area	Air Quality Scenario			
	% of Baseline Incidence	% Change in O_3-Attributable Risk		
	75ppb	75-70	75-65	75-60
2007 Simulation Year				
Atlanta, GA	1.1	4	8	12
Baltimore, MD	1.9	3	6	10
Boston, MA	1.2	2	5	9
Cleveland, OH	2.4	3	7	14
Denver, CO	0.8	2	5	9
Detroit, MI	3.0	3	6	10
Houston, TX	1.9	1	2	3
Los Angeles, CA	1.0	3	7	13
New York, NY	4.1	5	22	NA
Philadelphia, PA	3.2	3	6	9
Sacramento, CA	1.2	2	3	6
St. Louis, MO	2.5	4	9	14
2009 Simulation Year				
Atlanta, GA	1.0	3	7	9
Baltimore, MD	1.8	2	4	7
Boston, MA	1.1	-0.3	2	4
Cleveland, OH	2.3	3	7	12
Denver, CO	0.8	0.3	2	8
Detroit, MI	2.7	-4	-1	3
Houston, TX	1.9	-0.1	0.5	2
Los Angeles, CA	1.1	3	7	13
New York, NY	3.9	3	16	NA
Philadelphia, PA	3.0	2	4	6
Sacramento, CA	1.2	2	3	6
St. Louis, MO	2.3	2	5	9

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Figure 7-2 Heat Maps for Short Term O_3-attributable Mortality (Just meeting existing standard and risk reductions from just meeting alternative standards) (2007) (Smith et al., 2009 C-R functions) (see Key at bottom of figure)

Current Standard (75)

Study area	Daily 8hr Max Ozone Level (ppb)																Total
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75	
Atlanta, GA	0	0	0	0	3	5	18	24	41	52	63	38	15	6	3	0	267
Baltimore, MD	0	0	0	1	1	11	22	43	84	71	69	73	44	12	9	3	443
Boston, MA	0	0	0	0	4	20	45	50	58	57	35	20	30	9	13	11	353
Cleveland, OH	0	0	0	1	5	14	40	65	89	81	43	40	31	12	10	0	431
Denver, CO	0	0	0	0	0	0	1	5	6	13	17	23	15	4	2	0	86
Detroit, MI	0	0	0	0	2	7	42	72	123	147	75	52	56	20	43	17	655
Houston, TX	0	0	0	0	17	49	126	146	148	95	50	49	3	0	0	0	683
Los Angeles, CA	0	0	0	0	0	0	0	17	340	445	388	44	13	5	0	0	1,253
New York, NY	0	0	0	0	21	98	297	544	741	475	364	233	39	0	0	0	2,812
Philadelphia, PA	0	0	0	2	0	34	62	156	213	236	209	165	101	42	9	10	1,238
Sacramento, CA	0	0	0	0	1	18	53	98	67	65	40	20	5	2	0	0	367
St. Louis, MO	0	0	0	1	3	7	18	65	66	76	74	47	29	29	12	3	430

Decrease 75 to 70

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	0	0	0	0	0	1	1	2	3	2	1	0	0	0	12	0	12
Baltimore, MD	0	0	0	0	0	0	0	0	1	2	3	3	2	1	1	0	13	-1	14
Boston, MA	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	7	-1	8
Cleveland, OH	0	0	0	0	0	0	0	0	2	3	2	2	2	1	1	0	14	-2	15
Denver, CO	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	2	0	2
Detroit, MI	0	0	0	0	0	0	-1	0	2	5	3	3	4	2	3	1	23	-2	24
Houston, TX	0	0	0	0	-1	-1	-2	0	2	2	2	2	0	0	0	0	5	-4	9
Los Angeles, CA	0	0	0	0	0	0	0	0	6	16	17	2	1	0	0	0	43	0	43
New York, NY	0	0	0	0	-1	-2	0	12	27	32	35	25	5	0	0	0	134	-11	146
Philadelphia, PA	0	0	0	0	0	-1	0	0	3	7	8	9	6	3	1	1	35	-3	38
Sacramento, CA	0	0	0	0	0	-1	-1	1	2	2	2	1	0	0	0	0	7	-1	8
St. Louis, MO	0	0	0	0	0	0	0	1	2	3	4	3	2	2	1	0	18	0	19

Decrease 75 to 65

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	0	0	0	0	1	1	3	4	6	4	2	1	1	0	21	0	22
Baltimore, MD	0	0	0	0	0	0	0	0	3	4	5	7	5	1	1	0	27	-2	28
Boston, MA	0	0	0	0	0	-1	0	1	2	3	3	2	3	1	2	2	20	-2	22
Cleveland, OH	0	0	0	0	0	0	0	2	6	7	5	5	4	2	2	0	32	-3	34
Denver, CO	0	0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	4	-1	5
Detroit, MI	0	0	0	0	0	0	-1	1	4	8	6	6	7	3	6	3	42	-3	45
Houston, TX	0	0	0	0	-2	-3	-3	0	4	5	4	5	0	0	0	0	11	-9	20
Los Angeles, CA	0	0	0	0	0	0	0	0	13	33	35	4	1	1	0	0	87	0	87
New York, NY	0	0	0	0	-1	2	24	85	149	136	136	90	19	0	0	0	640	-6	646
Philadelphia, PA	0	0	0	0	0	-1	-1	0	7	15	17	18	12	6	1	2	76	-5	81
Sacramento, CA	0	0	0	0	0	-1	-1	2	3	4	3	2	0	0	0	0	13	-3	15
St. Louis, MO	0	0	0	0	0	0	0	3	4	7	8	6	4	4	2	1	39	-1	39

Decrease 75 to 60

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	0	0	0	0	1	2	4	7	9	6	3	1	1	0	34	0	34
Baltimore, MD	0	0	0	0	0	0	0	1	5	7	9	11	8	2	2	1	45	-2	47
Boston, MA	0	0	0	0	0	-1	1	2	4	5	5	3	5	2	3	2	32	-2	34
Cleveland, OH	0	0	0	0	-1	0	0	5	12	14	10	10	8	4	3	0	64	-3	67
Denver, CO	0	0	0	0	0	0	0	0	0	1	2	3	2	1	0	0	8	-1	8
Detroit, MI	0	0	0	0	0	0	-1	2	8	14	10	9	11	5	9	4	69	-4	73
Houston, TX	0	0	0	0	-3	-5	-4	2	8	10	7	8	1	0	0	0	24	-13	37
Los Angeles, CA	0	0	0	0	0	0	0	2	41	59	48	6	2	1	0	0	159	0	159
New York, NY	NA																		
Philadelphia, PA	0	0	0	0	0	-2	-1	1	11	23	26	27	17	8	2	2	116	-7	123
Sacramento, CA	0	0	0	0	0	-2	-2	4	6	7	5	3	1	0	0	0	23	-4	27
St. Louis, MO	0	0	0	0	0	0	0	5	7	11	12	9	6	7	3	1	60	-1	61

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Key: For *current standard (75)* which is an absolute risk metric expressed as incidence of mortality, color gradient ranges from blue (smallest O_3-related mortality count) to red (highest O_3-related mortality count). For *estimates of decreases in risk*, color gradient ranges from red (increase in risk – negative cell values) to blue (reduction in risk – positive cell values).

Figure 7-3 Heat Maps for Short Term O_3-attributable Mortality (Just meeting existing standard and risk reductions from just meeting alternative standards) (2009) (Smith et al., 2009 C-R functions) (see Key at bottom of figure)

Current Standard (75)

Study area	Daily 8hr Max Ozone Level (ppb)																Total
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75	
Atlanta, GA	0	0	1	2	8	16	18	33	49	44	29	30	9	2	0	0	241
Baltimore, MD	0	0	0	1	3	13	41	71	64	92	64	45	11	0	0	0	404
Boston, MA	0	0	1	0	11	25	45	57	50	53	48	7	3	6	9	3	319
Cleveland, OH	0	0	0	0	5	25	46	68	75	81	57	28	11	6	0	0	401
Denver, CO	0	0	0	0	0	1	2	4	9	17	22	20	6	2	1	0	83
Detroit, MI	0	0	1	9	7	26	46	67	114	148	38	51	46	0	21	6	579
Houston, TX	0	0	0	6	28	51	122	123	114	90	84	36	27	7	4	4	695
Los Angeles, CA	0	0	0	0	0	0	2	17	281	328	496	152	9	0	0	0	1,285
New York, NY	0	0	0	6	36	215	427	356	632	469	274	175	56	0	0	0	2,645
Philadelphia, PA	0	0	0	2	16	51	159	126	219	175	198	97	68	0	0	0	1,112
Sacramento, CA	0	0	0	0	2	23	64	69	73	56	42	33	6	0	0	0	367
St. Louis, MO	0	0	1	6	7	17	29	54	52	77	66	53	13	8	0	0	383

Decrease 75 to 70

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	0	0	0	0	0	1	2	3	2	2	1	0	0	0	8	-2	10
Baltimore, MD	0	0	0	0	0	0	-1	0	1	3	2	2	1	0	0	0	7	-2	9
Boston, MA	0	0	0	0	-1	-1	-1	0	0	1	1	0	0	0	0	0	-1	-6	5
Cleveland, OH	0	0	0	0	0	-1	0	1	2	3	3	2	1	0	0	0	12	-2	14
Denver, CO	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	-1	1
Detroit, MI	0	0	-1	-3	-2	-6	-5	-5	-5	-2	0	2	2	0	2	1	-21	-29	8
Houston, TX	0	0	0	-1	-2	-3	-3	-1	1	2	3	2	2	0	0	0	-1	-10	10
Los Angeles, CA	0	0	0	0	0	0	0	0	4	11	20	6	0	0	0	0	41	0	41
New York, NY	0	0	0	-1	-3	-14	-8	8	22	31	22	19	6	0	0	0	84	-38	121
Philadelphia, PA	0	0	0	0	-1	-2	-2	-1	5	5	8	5	4	0	0	0	19	-9	28
Sacramento, CA	0	0	0	0	0	-1	-1	1	2	2	2	2	0	0	0	0	6	-2	8
St. Louis, MO	0	0	0	-1	-1	-1	0	1	1	3	3	3	1	1	0	0	8	-4	12

Decrease 75 to 65

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	0	0	-1	-1	0	1	3	5	3	4	2	0	0	0	16	-3	19
Baltimore, MD	0	0	0	0	0	-1	-1	0	3	6	5	4	1	0	0	0	17	-3	21
Boston, MA	0	0	0	0	-1	-1	-1	0	1	2	3	1	0	0	1	0	5	-6	11
Cleveland, OH	0	0	0	0	-1	-1	0	4	5	8	7	4	2	1	0	0	29	-3	32
Denver, CO	0	0	0	0	0	0	0	0	0	0	1	2	1	0	0	0	2	-1	3
Detroit, MI	0	0	-1	-4	-2	-7	-5	-4	-2	3	2	4	5	0	3	1	-6	-27	21
Houston, TX	0	0	0	-2	-4	-5	-5	-1	2	4	6	3	3	1	1	1	4	-18	21
Los Angeles, CA	0	0	0	0	0	0	0	0	10	23	42	14	1	0	0	0	89	0	89
New York, NY	0	0	0	-1	-5	-17	16	52	107	120	81	63	21	0	0	0	437	-43	479
Philadelphia, PA	0	0	0	-1	-2	-4	-4	-1	10	11	17	10	8	0	0	0	44	-15	59
Sacramento, CA	0	0	0	0	0	-2	-1	2	3	4	3	3	1	0	0	0	12	-4	16
St. Louis, MO	0	0	-1	-2	-1	-1	0	2	3	6	6	6	2	1	0	0	21	-5	26

Decrease 75 to 60

Study area	Daily 8hr Max Ozone Level (ppb)																Total	Change in risk	
	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75		Inc.	Dec.
Atlanta, GA	0	0	-1	0	-1	-1	0	2	5	6	5	5	2	0	0	0	23	-3	26
Baltimore, MD	0	0	0	0	-1	-1	-1	1	5	10	8	6	2	0	0	0	28	-4	32
Boston, MA	0	0	0	0	-2	-1	-1	2	2	5	5	1	1	1	2	1	14	-6	20
Cleveland, OH	0	0	0	0	-1	-1	1	7	9	13	11	6	2	1	0	0	49	-3	53
Denver, CO	0	0	0	0	0	0	0	0	0	1	2	3	1	0	0	0	7	-1	8
Detroit, MI	0	0	-1	-4	-2	-7	-5	-2	2	10	4	8	8	0	4	2	15	-26	41
Houston, TX	0	0	0	-2	-7	-7	-7	0	5	8	11	6	5	1	1	1	14	-25	40
Los Angeles, CA	0	0	0	0	0	0	0	2	32	44	63	22	1	0	0	0	164	0	164
New York, NY	NA																		
Philadelphia, PA	0	0	0	-1	-3	-6	-4	0	16	17	25	14	11	0	0	0	69	-20	88
Sacramento, CA	0	0	0	0	-1	-3	-1	3	6	6	5	5	1	0	0	0	21	-6	27
St. Louis, MO	0	0	-1	-2	-1	-2	0	3	5	10	9	9	3	2	0	0	37	-6	43

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Key: For current standard (75) which is an absolute risk metric, color gradient ranges from blue (smallest O_3-related mortality count) to red (highest O_3-related mortality count). For *estimates of decreases in risk*, color gradient ranges from red (increase in risk – negative cell values) to blue (reduction in risk – positive cell values).

Figure 7-4 Plots of Short-Term O_3-attributable All-Cause Mortality for Meeting Existing standard and Alternative Standards (Smith et al., 2009) (Simulation year 2007 and 2009)

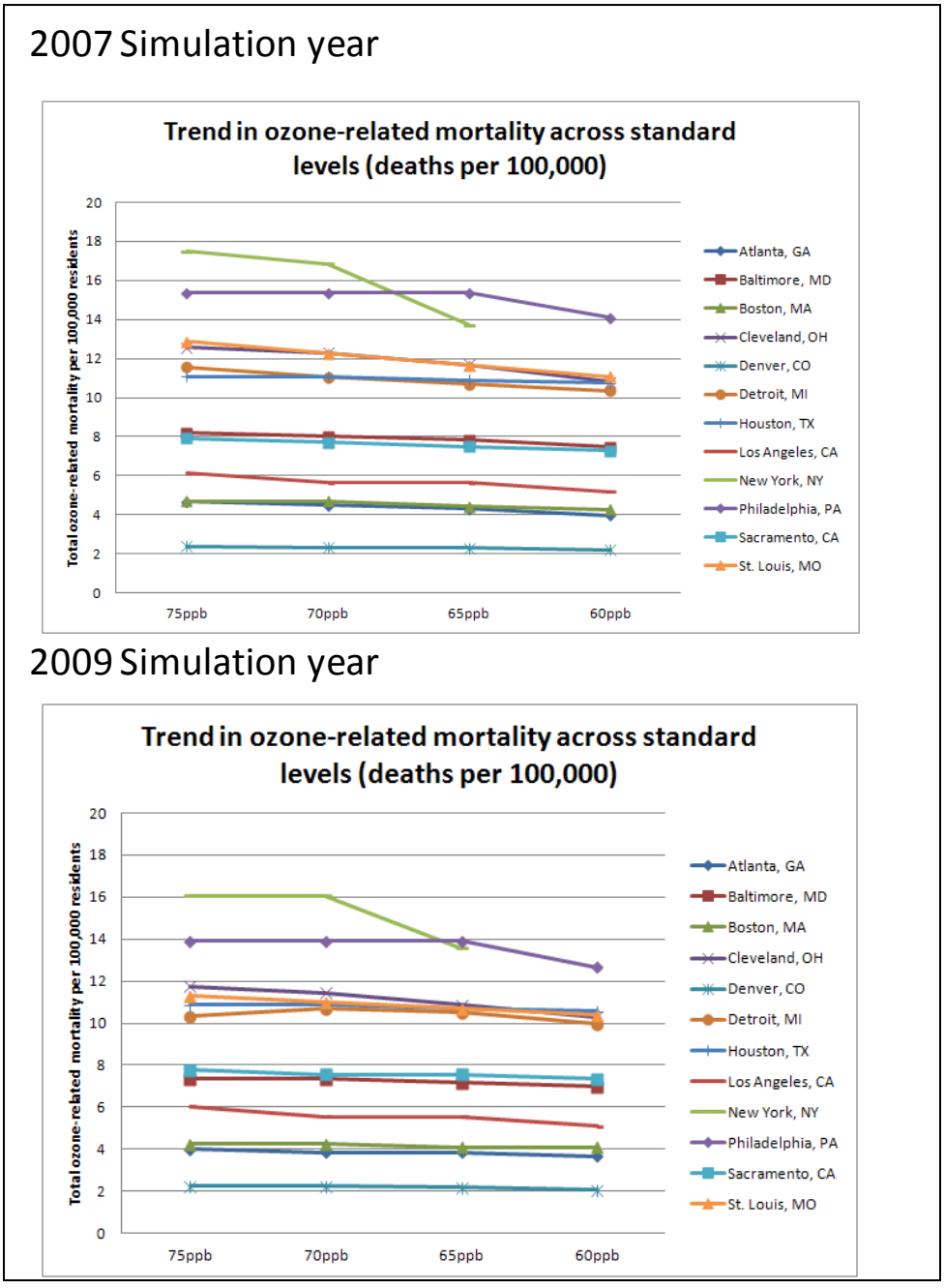


Table 7-9 Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in O_3-attributable Risk – Respiratory-Related Hospital Admissions (2007 and 2009)

Endpoint/Study Area/Descriptor	Air Quality Scenario							
	Absolute Incidence	Change in Incidence			Percent of Baseline	% Change in Ozone-Related Risk		
	75ppb	75-70	75-65	75-60	75ppb	75-70	75-65	75-60
2007 Simulation Year								
HA (respiratory); Detroit (Katsouyanni et al., 2009)								
1hr max, penalized splines	230	13	23	37	2.8	5	10	15
1hr max, natural splines	230	12	22	36	2.7	5	10	15
HA (respiratory); NYC (Silverman and Ito, 2010; Lin et al., 2008)								
HA Chronic Lung Disease (Lin)	120	6.7	29		3.3	5	23	
HA Asthma (Silverman)	420	28	120	NA	27.6	5	21	NA
HA Asthma, PM2.5 (Silverman)	310	20	84		20.1	5	22	
HA (respiratory); LA (Linn et al., 2000)								
1hr max penalized splines	790	19	38	60	2.4	2	5	7
HA (COPD less asthma); all 12 study areas (Medina-Ramon, et al., 2006)								
Atlanta, GA	67	4	6	10	2.5	5	9	15
Baltimore, MD	77	3	6	9	2.6	4	7	12
Boston, MA	100	2	6	10	2.2	2	6	9
Cleveland, OH	61	2	5	10	2.4	4	8	17
Denver, CO	27	1	2	3	2.9	3	6	11
Detroit, MI	90	3	6	9	2.5	3	6	10
Houston, TX	68	1	2	4	2.1	1	3	6
Los Angeles, CA	180	8	16	25	2.7	4	9	13
New York, NY	180	11	50	NA	2.2	6	28	NA
Philadelphia, PA	130	4	10	15	2.5	3	7	11
Sacramento, CA	34	1	2	4	2.5	3	7	11
St. Louis, MO	53	3	5	8	2.6	5	10	15
2009 Simulation Year								
HA (respiratory); Detroit (Katsouyanni et al., 2009)								
1hr max, penalized splines	220	3.6	13	25	2.5	2	6	11
1hr max, natural splines	210	3.4	12	24	2.4	2	6	11
HA (respiratory); NYC (Silverman and Ito, 2010; Lin et al., 2008)								
HA Chronic Lung Disease (Lin)	120	5.1	21		3.2	4	17	
HA Asthma (Silverman)	410	24	96	NA	27.2	4	17	NA
HA Asthma, PM2.5 (Silverman)	310	17	68		19.8	4	18	
HA (respiratory); LA (Linn et al., 2000)								
1hr max penalized splines	640	18	39	62	2.4	2	4	7
HA (COPD less asthma); all 12 study areas (Medina-Ramon, et al., 2006)								
Atlanta, GA	65	3	5	8	2.2	5	8	12
Baltimore, MD	74	2	4	6	2.3	2	5	8
Boston, MA	92	0	1	4	2.0	0	1	4
Cleveland, OH	58	2	5	8	2.2	3	8	14
Denver, CO	27	0	1	3	2.7	1	4	11
Detroit, MI	81	-3	-2	1	2.2	-4	-2	1
Houston, TX	71	1	2	4	2.2	1	2	5
Los Angeles, CA	200	8	16	26	2.7	4	8	13
New York, NY	170	7	35	NA	2.1	4	20	NA
Philadelphia, PA	120	2	6	9	2.3	2	4	7
Sacramento, CA	41	1	3	4	2.4	3	7	11
St. Louis, MO	51	2	4	6	2.4	3	8	12

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Table 7-10 Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in O_3-attributable Risk – Emergency Room Visits (2007 and 2009)

Endpoint/Study Area/Descriptor		Air Quality Scenario							
		Absolute Incidence	Change in Incidence			Percent of Baseline	% Change in Ozone-Related Risk		
		75ppb	75-70	75-65	75-60	75ppb	75-70	75-65	75-60
2007 Simulation Year									
ER Visits (repiratory); Atlanta (Strickland et al., 2007)									
	Distributed lag 0-7 days	7,500	410	740	1,200	19.6	4	8	13
	Average day lag 0-2	4,500	230	420	670	11.6	5	8	13
ER-visits (respiratory); Atlanta (Tolbert et al., 2007, Darrow et al., 2011)									
	Tolbert	8,100	360	670	1,100	5.8	4	8	12
	Tolbert-CO	7,200	320	590	940	5.1	4	8	12
	Tolbert-NO2	6,500	290	530	840	4.6	4	8	12
	Tolbert-PM10	5,100	230	420	660	3.6	4	8	12
	Tolbert-PM10, NO2	4,900	220	400	640	3.5	4	8	12
	Darrow	4,400	190	360	560	3.1	4	8	12
ER-visits (asthma); NYC (Ito et al, 2007)									
	single pollutant model	9,000	530	2,300		19.9	5	22	
	PM2.5	7,100	410	1,800		15.5	5	22	
	NO2	5,800	330	1,500	NA	12.8	5	23	NA
	CO	9,500	570	2,500		21.0	5	22	
	SO2	7,300	420	1,900		16.0	5	22	
2009 Simulation Year									
ER Visits (repiratory); Atlanta (Strickland et al., 2007)									
	Distributed lag 0-7 days	6,800	310	570	800	17.2	4	7	10
	Average day lag 0-2	4,000	170	320	460	10.1	4	7	10
ER-visits (respiratory); Atlanta (Tolbert et al., 2007, Darrow et al., 2011)									
	Tolbert (single pollutant	7,400	270	500	720	5.1	3	6	9
	Tolbert-CO	6,600	240	450	640	4.5	3	6	9
	Tolbert-NO2	6,000	210	400	580	4.1	3	6	9
	Tolbert-PM10	4,700	170	310	450	3.2	3	6	9
	Tolbert-PM10, NO2	4,500	160	300	430	3.1	3	6	9
	Darrow (single pollutant	4,000	140	270	380	2.8	3	6	9
ER-visits (asthma); NYC (Ito et al, 2007)									
	single pollutant model	8,800	400	1,800		19.3	4	17	
	PM2.5	6,900	310	1,400		15.0	4	17	
	NO2	5,700	250	1,100	NA	12.4	4	18	NA
	CO	9,300	430	1,900		20.4	4	17	
	SO2	7,100	320	1,400		15.5	4	17	

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Table 7-11 Short-Term O_3-attributable Morbidity Incidence, Percent of Baseline and Reduction in O_3-attributable Risk – Asthma Exacerbations (2007 and 2009)

Endpoint/Study Area/Descriptor	Air Quality Scenario							
	Absolute Incidence	Change in Incidence			Percent of Baseline	% Change in Ozone-Related Risk		
	75ppb	75-70	75-65	75-60	75ppb	75-70	75-65	75-60
2007 Simulation Year								
Asthma exacerbation (wheeze); Boston (Gent et al., 2003, 2004)								
Chest Tightness	69,000	2,100	5,600	8,600	28.8	2	5	9
Shortness of Breath	49,000	1,400	3,700	5,700	16.2	2	6	10
Chest Tightness (1hr max)	51,000	1,200	3,200	5,000	21.2	2	5	8
Shortness of Breath (1hr max)	59,000	1,300	3,600	5,800	19.6	2	5	8
Chest Tightness (PM2.5)	69,000	2,100	5,600	8,700	29.1	2	5	9
Chest Tightness (PM2.5)	64,000	1,900	5,100	8,000	26.8	2	5	9
Wheeze (PM2.5)	130,000	3,800	10,000	16,000	23.2	2	6	9
2009 Simulation Year								
Asthma exacerbation (wheeze); Boston (Gent et al., 2003, 2004)								
Chest Tightness	63,000	490	2,400	4,800	27.0	0.4	2	5
Shortness of Breath	45,000	330	1,600	3,200	15.1	1	3	6
Chest Tightness (1hr max)	47,000	-180	790	2,200	19.8	-0.4	1	3
Shortness of Breath (1hr max)	54,000	-210	910	2,500	18.3	-0.4	1	4
Chest Tightness (PM2.5)	64,000	500	2,400	4,800	27.2	0.4	2	5
Chest Tightness (PM2.5)	59,000	450	2,200	4,400	25.1	0.4	3	5
Wheeze (PM2.5)	120,000	900	4,300	8,700	21.7	0.5	3	6

Figure 7-5 Plots of Short-Term O_3-attributable Respiratory HA for Meeting Existing standard and Alternative Standards (Medina-Ramon, et al., 2006) (Simulation year 2007 and 2009)

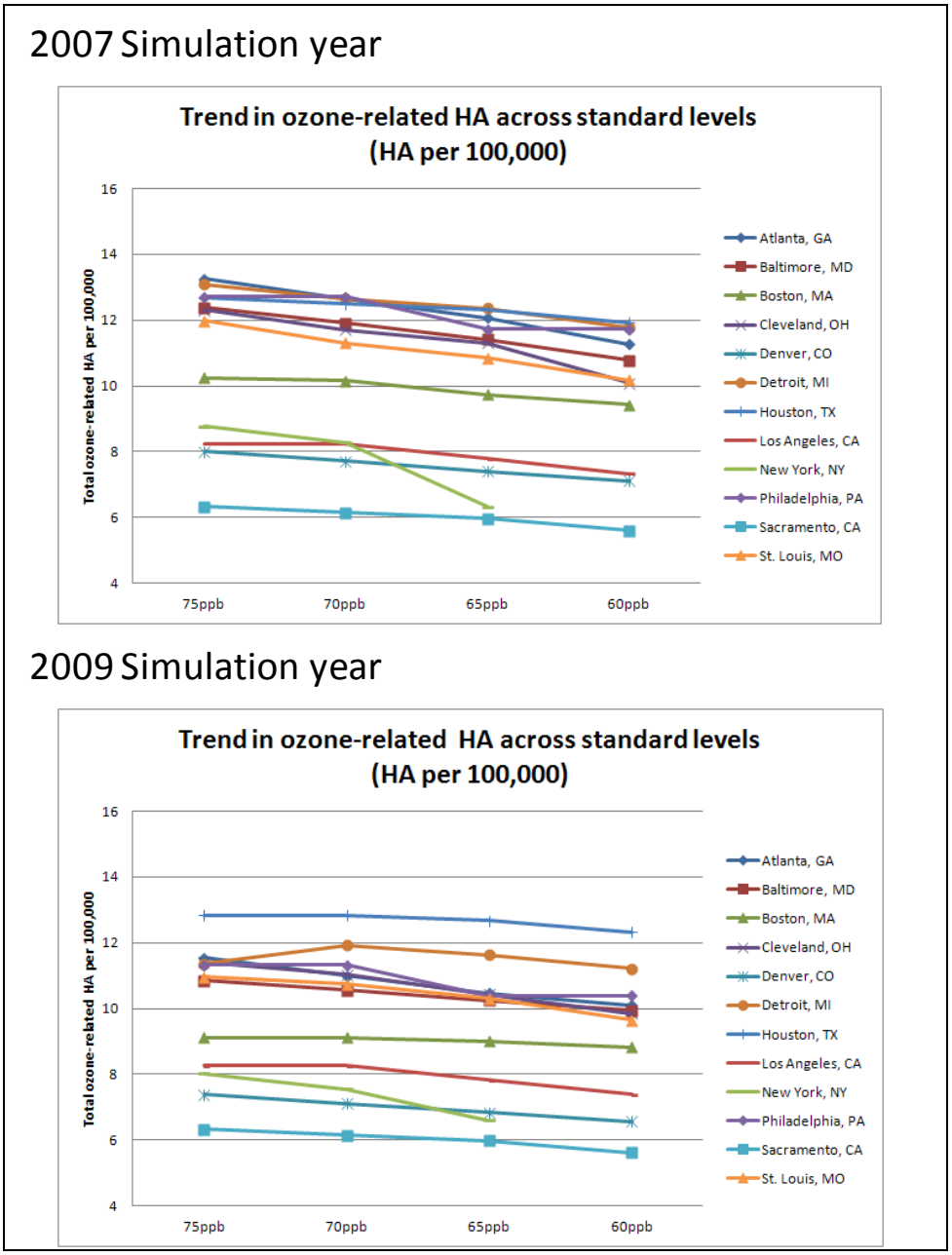


Table 7-12 Long-Term O_3-attributable Respiratory Mortality Incidence (2007 and 2009) (Jerrett et al., 2009 C-R Functions) (CBSA-based study area, no threshold)

Study Area	Air Qualtiy Scenario			
	Absolute Incidence	Change in Incidence		
	75ppb	75-70	75-65	75-60
2007 Simulation Year				
Atlanta, GA	710	43	78	120
	(260 - 1100)	(15 - 71)	(26 - 130)	(41 - 200)
Baltimore, MD	750	33	67	110
	(270 - 1200)	(11 - 55)	(23 - 110)	(37 - 180)
Boston, MA	1,100	35	93	140
	(400 - 1800)	(12 - 58)	(32 - 150)	(49 - 240)
Cleveland, OH	530	26	56	100
	(190 - 820)	(9 - 43)	(19 - 93)	(35 - 170)
Denver, CO	480	19	39	64
	(170 - 740)	(6 - 31)	(13 - 64)	(22 - 100)
Detroit, MI	760	35	63	99
	(270 - 1200)	(12 - 59)	(21 - 100)	(33 - 160)
Houston, TX	550	9.5	19	32
	(190 - 860)	(3 - 16)	(6 - 31)	(11 - 53)
Los Angeles, CA	2,600	140	260	410
	(930 - 4000)	(46 - 230)	(89 - 430)	(140 - 670)
New York, NY	1,800	120	480	NA
	(660 - 2900)	(41 - 200)	(160 - 790)	
Philadelphia, PA	1,300	56	120	170
	(450 - 1900)	(19 - 93)	(40 - 190)	(59 - 290)
Sacramento, CA	680	31	60	100
	(250 - 1100)	(10 - 51)	(20 - 98)	(34 - 170)
St. Louis, MO	600	34	69	100
	(210 - 930)	(230 - 1000)	(23 - 110)	(35 - 170)
2009 Simulation Year				
Atlanta, GA	700	41	76	100
	(250 - 1100)	(14 - 68)	(26 - 120)	(36 - 170)
Baltimore, MD	730	25	54	84
	(260 - 1100)	(8 - 41)	(18 - 90)	(28 - 140)
Boston, MA	1,100	6.8	42	85
	(380 - 1700)	(2 - 11)	(14 - 69)	(29 - 140)
Cleveland, OH	510	24	54	84
	(180 - 800)	(8 - 41)	(18 - 89)	(29 - 140)
Denver, CO	490	9.0	28	70
	(180 - 770)	(3 - 15)	(10 - 47)	(24 - 120)
Detroit, MI	720	-8.9	18	51
	(260 - 1100)	(-3 - -15)	(6 - 30)	(17 - 85)
Houston, TX	610	14	30	49
	(220 - 950)	(5 - 23)	(10 - 49)	(17 - 82)
Los Angeles, CA	2,800	130	280	430
	(1000 - 4300)	(45 - 220)	(94 - 460)	(150 - 710)
New York, NY	1,900	110	390	NA
	(670 - 2900)	(37 - 180)	(130 - 630)	
Philadelphia, PA	1,200	44	94	140
	(430 - 1900)	(15 - 73)	(32 - 160)	(47 - 230)
Sacramento, CA	730	34	66	110
	(260 - 1100)	(12 - 57)	(22 - 110)	(36 - 170)
St. Louis, MO	580	24	53	86
	(210 - 900)	(210 - 910)	(18 - 87)	(29 - 140)

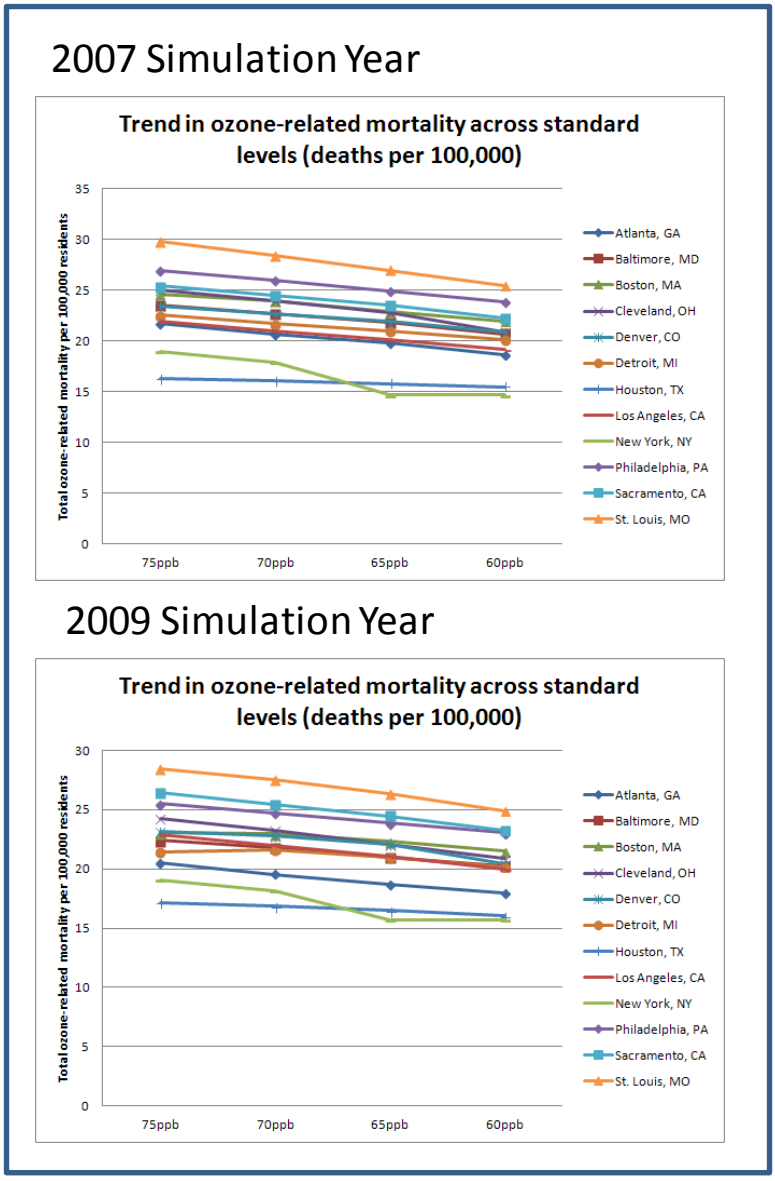
NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Table 7-13 Long-Term O_3-attributable Respiratory Mortality Percent of Baseline Incidence and Percent Reduction in O_3-attributable Risk (simulation years 2007 and 2009) (Jerrett et al., 2009 C-R Functions) (CBSA-based study area, no threshold)

Study Area	Air Quality Scenario			
	% of Baseline Attributable to Ozone	Change in O3-Attributable Risk		
	70ppb	75-70	75-65	75-60
2007 Simulation Year				
Atlanta, GA	17.7	5	9	15
Baltimore, MD	18.1	4	8	12
Boston, MA	16.7	3	7	11
Cleveland, OH	16.9	4	9	17
Denver, CO	20.1	3	7	11
Detroit, MI	17.7	4	7	11
Houston, TX	16.1	1	3	5
Los Angeles, CA	19.6	4	9	13
New York, NY	15.9	6	24	NA
Philadelphia, PA	17.7	4	8	12
Sacramento, CA	17.1	4	7	13
St. Louis, MO	17.9	5	10	15
2009 Simulation Year				
Atlanta, GA	16.1	5	9	13
Baltimore, MD	16.9	3	6	10
Boston, MA	15.9	1	3	7
Cleveland, OH	16.1	4	9	15
Denver, CO	19.7	1	5	12
Detroit, MI	17.1	-1	2	6
Houston, TX	16.6	2	4	7
Los Angeles, CA	19.9	4	8	13
New York, NY	15.9	5	18	NA
Philadelphia, PA	16.7	3	7	10
Sacramento, CA	17.3	4	8	12
St. Louis, MO	17.1	4	8	13

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Figure 7-6 Plots of Long-Term O_3-attributable Respiratory Mortality for Meeting Existing standard and Alternative Standards (Jerrett et al., 2009) (Simulation year 2007 and 2009)



The presentation of key observations drawn from review of the core risk estimates is divided into two sections: 1) the assessment of health risks associated with just meeting the existing standard (section 7.5.1) and 2) the assessment of risk changes from meeting alternative standards relative to meeting the existing standard(section 7.5.2). The presentation of key observations in each of these two sections is further separated into those associated with (a) short-term O_3-attributable mortality, (b) short-term O_3-attributable morbidity and (c) long-term O_3-attributable mortality. Unless otherwise noted, all risk estimates discussed in these three sections are core risk estimates. In some cases we refer to the confidence intervals around risk estimates. When an effect estimate is drawn from a study with low statistical power, confidence intervals can be wide, and can include negative values because of the assumptions of normality in the distribution of the effect estimate. Negative lower-confidence bounds do not imply that additional exposure to O_3 has a beneficial effect, but rather that the estimated O_3 effect estimate in the C-R function was not statistically significantly different from zero, and thus has a higher degree of uncertainty as to the magnitude of the estimated risk. As noted earlier, presentation of sensitivity analysis results and their use in interpreting the core risk estimates is covered in section 7.5.3.

7.5.1 Assessment of Health Risk After Just Meeting the Existing 75 ppb standard

The analysis of risk after simulating just meeting the existing standard focuses on absolute risk, since this is of greatest relevance in evaluating the adequacy of the existing standard.

Short-term O_3-attributable mortality

- After meeting the existing standard, estimates of O_3-related all-cause mortality range across urban areas from 86 to 2,800 deaths (for simulation year 2007) and from 83 to 2,600 deaths (for simulation year 2009) (see Table 7-7). This translates into from 0.8 to 4.1% of baseline all-cause mortality (for simulation year 2007) and from 0.8 to 3.9% (for simulation year 2009) (see Table 7-8) in these study areas. Generally, O_3-attributable all-cause mortality risks continue to be lower for the 2009 simulation year as compared with the 2007 simulation year (with the exception of Houston), reflecting the generally lower ambient O_3 levels associated with 2009 for most of the study areas (see Tables 7-7 and 7-8).

- Confidence intervals (CIs) reflecting the statistical fit of the effect estimates used in modeling risk demonstrate substantial variability across the 12 urban study areas. In

general, the upper 95th percentile CI tends to be from 2-3 times larger than the point estimate for the 12 urban study areas (see Table 7-7). However, some cities have markedly wider confidence intervals (e.g., Denver where the upper CI is ~5 times the point estimate), while others have tighter relative CIs (e.g., New York, where the upper CI is ~1.4 times larger than the point estimate). This variation in the CIs associated with risk estimates can reflect a number of factors including the statistical power of the underlying epidemiological study, which is based on the population size, and differences in the magnitude of such factors as exposure measurement error and correlations between O_3 and other pollutants.

After just meeting the existing O_3 standard, all-cause mortality estimates based on C-R functions from Smith et al., 2009 (for simulation year 2007) continue to be driven largely by days with total O_3 levels falling in the range of 30 to 70 ppb, with 87 to 99% of the mortality estimate across the 12 urban study areas associated with days in this range. A smaller, but still significant fraction (9 to 24%) of the mortality risk is associated with days above 60 ppb (see Figure 7-2, "Existing standard (75)" plot).[32] For 2009, this trend continues although risk distributions are shifted down somewhat (reflecting the lower ambient O_3 levels generally seen in this simulation year compared with 2007) (see Figure 7-3, "current standard (75)" plot). For 2009, a substantial portion (2% to 24%) of O_3-attributable mortality risk is now associated with days having O_3 measurements 55-60 ppb or higher. A relatively smaller fraction (~0% to 2%) of total mortality estimates for the existing standard are associated with days having ambient O_3 levels of 20 ppb or less.[33]

- Estimates of O_3- attributable respiratory-related HA range from 10's to 100's of cases (after just meeting the existing standard) depending on the type of respira$_t$ory HA endpoint modeled and the specific urban study areas evaluated (see Table 7-9). All 12 urban study areas were modeled for one of more respiratory-related HA endpoints.

[32] Houston has a significantly smaller percentage (<1) of its mortality signal associated with days above 60ppb.

[33] In the first draft O_3 REA, we included consideration for surrogate LMLs (based on the lowest composite monitor values used in modeling short-term exposure-related mortality for each urban study area – see Table 7-5, First Draft REA, U.S. EPA, 2012). For the 8hr max monitoring season LML (applicable to the Smith et al.., 2009-based core risk estimate generated for this second draft REA), we have values ranging from 4 to 17 ppb and from 5 to 16 ppb (across the 12 urban study areas for 2007 and 2009, respectively). If we look at heat maps characterizing the distribution of short-term exposure-related mortality for recent conditions (see "recent conditions" heat maps in Figures 7B-1 and 7B-2 in Appendix B) we see that the vast majority of ozone-related mortality falls above these surrogate LML ranges. Consequently we see, that had we integrated consideration for the surrogate LMLs into modeling of short-term exposure-related mortality, there would have been very little change in estimates of risk.

- O_3- attributable ER (for respiratory symptoms) ranged into the thousands for both New York and Atlanta under simulated attainment of the existing standard (these were the only two study areas modeled for this health endpoint) (see Table 7-10).

- Estimates of O_3-attributable asthma exacerbation (wheeze) in Boston are in the tens of thousands to over 100,000 (see Table 7-11). The percent of baseline for this endpoint (after just meeting the existing standard) is generally in the 20-30% range which is markedly higher than other short-term morbidity endpoints modeled for this analysis (see Table 7-11 and compare to values in 7-9 and 7-10).

Long-term O_3-attributable mortality

- After simulating just meeting the existing standard, estimates of O_3-related respiratory mortality range across urban areas from 480 to 2,600 deaths (for 2007) and from 490 to 2,800 deaths (for 2009) (see Table 7-12). This translates into from 16.3 to 20.8% of baseline across the 12 urban study areas (for 2007) and from 15.9 to 20.7% (for 2009) using the single Jerrett et al., 2009 C-R national-scale function applied to each urban area (see Table 7-13). As discussed in section 7.3.2, because of the long-term exposure metric (seasonal mean of daily 8-hour maximum) employed in risk modeling, there is the potential for some degree of overlap between short-term and long-term exposure-related mortality estimates. For that reason, these two categories of mortality estimates cannot be considered distinct and should not be added to estimate total mortality.

- 95th percentile CIs for long-term O_3-attributable respiratory mortality suggest greater power (and potential less heterogeneity) associated with modeling this health endpoint, compared with short-term O_3-attributable mortality. None of the CIs for long-term O_3-attributable mortality include negative estimates as lower bounds (see Table 7-12).

7.5.2 Assessment of Health Risk Associated with Simulating Meeting Potential Alternative Standards of 70, 65, and 60 ppb

As discussed earlier, we have considered three alternative standard levels (70, 65 and 60 ppb), each evaluated using the form and averaging time of the existing standard. In presenting risk estimates associated with the simulated attainment of each of these alternative standard

levels, we focus on the change in risk associated with a comparison of O_3 levels after simulation of the existing standard with levels after simulation of each of the alternative standard levels. This is of greatest relevance in comparing the potential public health benefit associated with each of the alternative standards relative to the level of protection afforded by just meeting the existing standard.

In reviewing these risk estimates, it is important to keep in mind that simulation of alternative standard levels is based on a reaching a peak-based attainment metric. Based on the simulated air quality information for the 12 urban study areas, there is a tendency for O_3 to increase on lower concentration days and decrease on higher concentration days.[34] Therefore, it is not immediately clear that we would expect risk reductions when applying C-R functions that are based on the full distribution of daily 8-hour max values. Specifically, risk reductions are only expected to the extent that the composite monitor daily 8-hour max values decrease as lower alternative standards are simulated. As discussed in Chapter 4 (section ???), after adjustment to alternative standard levels, decreases in O_3 typically occur on higher O_3 days which tend to occur during warmer (summer) months and are concentrated in suburban areas. Conversely, increases in O_3, ttypically occur lower O_3 days which tend to occur in the cooler portions of the year and are focused in core urban areas. In general, variability in predicted daily O_3 concentrations decreases when meeting lower standard levels.

Short-term O_3-attributable mortality

- In our analysis, the mortality risk metric is generally not responsive to meeting the existing and alternative standard levels. This reflects a number of factors all related to 1) how O_3 concentrations respond to reductions in NOx emissions used to meet the standards, and 2) how the risk metrics are associated with temporal and spatial patterns of O_3. As discussed in section 7.1.1, mortality risk is modeled using composite monitor values (i.e., averages of O_3 measurements across monitors in an urban study area) which removes spatial variability in measured O_3 within an urban study area (also removing variability in changes in O_3 across an urban area resulting from NOx reductions). Furthermore, in modeling total mortality risk for the core analysis, we add the risk changes occurring across all days within the monitored O_3 season, including days with low values of O_3 as well as days with high values of O_3. This means that we include both decreases in risk on those days when O_3 is estimated to decrease (generally occurring on days with higher values of O_3) and increases in

[34] This relationship is also observed in ambient air quality measurements as discussed in Chapter 4 and appendices.

risk when O_3 is simulated to increase (generally associated with lower values of O_3). The dampened response of short-term mortality risk can be contrasted with clinical study-based risk estimates. The clinical study-based estimates primarily reflect changes in the upper end of the O_3 distribution where we tend to see more uniform reductions under simulation of alternative standard levels. In addition, clinical-based estimates of risk are based on detailed micro-environmental exposure modeling which uses individual monitor values instead of composite monitor values, thereby resulting is less dampening of spatial variability in O_3 within a given urban study area.

- Generally, the magnitude of risk reduction increases as lower alternative standard levels are simulated. For example, for the lowest alternative standard we evaluated, 60 ppb, across the 12 urban study areas, we predict from 8 to 160 fewer O_3-attributable deaths for simulation year 2007 (relative to risk after just meeting the existing standard) (see Table 7-7). This range is from 7 to 160 deaths for simulation year 2009. These ranges (for the 60 ppb standard level) translate into a 3 to 14% reduction in O_3-attributable risk relative to risk after just meeting the existing standard (see Table 7-8).

- As noted in section 7.1.1, some of the urban study areas are projected to experience increases in O_3 (and hence risk) when attainment with the existing standard and some of the alternative standard levels is simulated. Focusing specifically on the alternative standard levels, we see that, for the core analysis, this potential increase in risk only occurs for the 2009 simulation year and specifically for three of the urban study areas (Boston, Detroit and Houston – see Table 7-7). For example, Detroit is predicted to have an increase of 21 O_3-attributable deaths after meeting the 70 ppb standard (when compared to risk remaining after meeting the existing standard). However, we estimate a net reduction of 15 O_3-attributable deaths after meeting the 60 ppb level (again based on comparison to risk after meeting the existing standard). Furthermore, for all three urban study areas with initial risk increases (based on comparing meeting the existing standard to meeting alternative standards), we see that these increases are offset after meeting the lowest alternative standard simulated (60 ppb) (see Table 7-7). The potential for risk increases is increased somewhat for several of the urban study areas when we simulate how the O_3 distribution shifts from recent conditions to just meeting the existing standard (see Appendix 7B, Tables 7B-1 and 7B-2). Specifically, in simulating estimated risk from moving from recent conditions to attaining the existing standard, we see that for the 2007 simulation year, two of the

study areas (Houston and Los Angeles) have risk increases after meeting the existing standard compared to recent conditions while half of the twelve urban study areas have risk increases for the 2009 simulation year in adjusting air quality to meetthe existing standard relative to recent conditions. It is also important to keep in mind that, for the urban areas of New York and Los Angeles, there are additional uncertainties in the simulation of existing and alternative standards given the limitations in the application of the adjustment methodology to very large emissions perturbations and the fact that the 95th percent confidence interval lower bound estimate of hourly O_3 concentrations was used to capture a scenario in which these cities could meet lower standard levels (65 ppb for New York and 60 ppb for Los Angeles). In five of these eight cases, the initial risk increases (including the increase in going from recent conditions to the existing standard) is fully offset after meeting the lowest alternative standard level (60 ppb).[35]

- Figure 7-4 provides plots of short-term mortality risk for the existing and alternative standards adjusting for total exposed population (i.e., O_3-attributable deaths per 100,000 exposed). From this figure it can be seen that total O_3-attributable risk, even when adjusted for population, varies substantially across the 12 urban study areas, with New York and Philadelphia having the highest risk and Boston and Denver the lowest. This spread in risk (adjusted for population) reflects, to a great extent, differences in the effect estimates used in modeling this endpoint for each study area, which can in turn reflect a number of factors (e.g., differences in behavior such as outdoor activity across cities and differences in exposure measurement error). However, despite considerable variability in absolute O_3-attributable risk, Figure 7-4 also suggests that most of the study areas display relatively limited reduction in O_3-attributable risk across the three alternative standards (with the exception of New York, which has a notable decrease in risk for the 70 to 65 ppb standard level).[36] This suggests that a substantial fraction of O_3-attributable risk would still remain, even after simulated attainment of the lowest alternative standard considered.

- Heat map plots of risk reductions for 2007 suggest that most of the risk reductions associated with simulation of all three alternative standards occur on days with

[35] For both LA and Houston (in 2007) and Houston (2009) a modest net risk increase still persists (compared to risk under recent conditions), even when we have simulated the lowest alternative standard considered (60 ppb) (see Tables 7B-1 and 7B-2).

[36] With the New York City study area, we recognize however that there is significant uncertainty associated with the use of the CBSA-based study area due to significant heterogeneity in short-term O_3-attributable mortality effect estimates (from Smith et al., 2009) falling within that larger urban study area (see discussion in section 7.6.1).

composite O_3 level between 35 and 60 ppb (see Figure 7-4). By contrast, most of the risk increases occur on days with composite O_3 levels between 20 and 35ppb (see Figure 7-4). This is expected given that most of the increases in urban core O_3 are associated with lower O_3 days where NOx titration is prevalent (see Appendix D, section 4.6 Figures 40-54). Very little of the projected change in risk (increases, or decreases) for any of the alternative standards considered occurred on days with O_3 levels below 20 ppb O_3Similar observations hold for risk results generated for simulation year 2009.

Short-term O_3-attributable morbidity

- Generally, because the short-term O_3 exposure-related morbidity endpoints use the same air metrics as used in modeling short-term O_3-attributable mortality (i.e., 8hr maximum and 8hr mean) the pattern of risk reduction seen for these morbidity endpoints are similar to those seen with short-term mortality (see Tables 7-9 though 7-11 and Figure 7-5). However, New York, as mentioned with regard to short-term O_3-attributable mortality, has substantially higher percent reductions (for O_3-attributable risk) compared with the other study areas. For example, with ER visits (asthma), under the lowest alternative standard in simulation year 2007, New York is estimated to have a 22 to 23% reduction in the number of ER-visits associated with O_3 exposure (see Table 7-10).

Long-term O_3-attributable mortality

- Although long-term O_3-attributable mortality is modeled using a different O_3 metric (essentially a long-term trend in the 1hr maximum for the hottest two seasons – see section 7.3.2) the overall magnitude and pattern of reduction in O_3-related risk is similar to that seen with short-term exposure related mortality. Specifically, for the 2007 simulation year, for most urban study areas risk reductions range from 11 to 15% (for the 60 ppb standard) (see Table 7-13). Risk reductions are generally slightly smaller across alternative standard levels for simulation year 2009. For the 2009 simulation year, for Detroit, we see a relatively small risk increase for the 70 ppb alternative standard (compared to risk under the existing standard). However that initial increase is offset by risk reductions for the other (lower) alternative standard levels simulated (see Table 7-13).

7.5.3 Sensitivity Analyses Designed to Enhance Understanding of the Core Risk Estimates

We have completed a number of sensitivity analyses intended to support interpretation of the core risk estimates. These sensitivity analyses, which are described in section 7.4.3, can be divided into two categories: (a) sensitivity analyses exploring factors impacting air quality characterization (specifically composite monitor composition) and (b) sensitivity analyses exploring the impact of alternative C-R function specifications. As noted in section 7.4.3, we also completed an initial influence analysis designed to identify which of the input factors to the risk model (for short-term exposure-related mortality) are primarily responsible for inter-city variability in that risk metric. This section summarizes the results of these sensitivity analyses and presents key observations related to those analyses, beginning with the influence analysis and then proceeding to sensitivity analyses focused on air quality characterization and alternative C-R function specification.

Influence analysis

The influence analysis considered three factors involved in modeling risk for the short-term exposure-related mortality endpoint including: baseline incidence, composite monitor O_3 levels and Bayes-adjusted city-specific effect estimates (recall that the core risk estimate is based on effect estimates derived as part of analyses published in Smith et al., 2009). Each of these input factors displays inter-study area variation and are responsible, collectively, for heterogeneity in risk estimates.[37] In completing the analysis, we first calculated a central tendency estimate of risk based on the mean of each input factor across the 12 urban study areas for the 2009 simulation year (i.e., using the average of the city-specific values for each of the input factors). We then systematically varied each of the three heterogeneity-related factors (effect estimate, composite monitor-based O_3 level and baseline incidence) to one standard deviation (SD) above its mean value (reflecting variance across the 12 urban study area values) and noted the percent increase produced by that perturbation over the initial mean risk estimate. This influence analysis allowed us to explore the impact of both model form – specifically, potential non-linearities in the model – as well as the relative magnitude of variability in each of the three heterogeneity-related input factors on risk. The influence analysis generated the following results: baseline incidence (23%), composite monitor-based O_3 level (8%), and effect estimate (58%). In other words, the 58% result for *effect estimate* means that use of a value 1 SD over the mean (for the effect estimate) in generating risk, resulted in a risk estimate that was 58%

[37] Note, that the demographic count input factor also varies across the study areas and is an important factor in determining total incidence. However, for the influence analysis, we used deaths per 100,000 as the risk metric which standardizes on demographic count and therefore allowed us to exclude this input parameter in conducting the influence analysis.

larger than the risk estimate based on the mean of all input factors. These results clearly show that, of the three input factors considered, the *effect estimate* is primarily responsible for inter-city variability in short-term exposure-related risk.

Interestingly, when we look at the coefficient of variation (CV) for these three heterogeneity-related input factors we see values almost identical to the influence analysis results in terms of relative magnitude to each other (i.e., 0.232, 0.084, and 0.527 for baseline incidence, composite monitor-based O_3 levels and effect estimate, respectively). Given that the CV values only reflect variability in each input factor and not model form (i.e., do not reflect potential non-linearities in the model), the fact that the CV values almost exactly match the influence analysis results in terms of relative magnitude suggest that there is very little if any non-linearity in the model calculations involving these three input factors. Had non-linearity existed to a significant extent, then the influence anlaysis results would have differed substantially from the CV results. The fact that both analyses suggest a primary role for the effect estimate in driving inter-city variability in risk emphasizes the importance of the sensitivity analyses exploring alternative C-R functions specifications that were completed for the REA (see below).

Air quality-related sensitivity analyses

This category of sensitivity analysis covers (a) the use of a smaller study area (the Smith et al., 2009 study areas) as contrasted with the CBSA-based study areas used in the core analysis, and (b) the use of alternative approaches to simulate attainment of the existing and alternative standards (for a subset of the study areas) (see section 7.4.3 for additional detail). This category of sensitivity analysis was applied to short-term O_3-attributable mortality given the importance of the endpoint in the policy-context.[38]

To allow for easier visual comparisons, we have presented the results of this sensitivity analysis category in graphical form (see Figure 7-7, numerical results are presented in Appendix 7C). This figure presents point estimates and 95th percentile confidence ranges for the core model and for two sensitivity analyses: (a) SA1 (use of the smaller Smith et al., 2009 based study area) and (b) SA2 (use of the alternative approach to simulating attainment). SA2 is not presented for all of the study areas, only for the subset included in these alternative simulations (see section 7.4.3). The sensitivity analyses results presented in Figure 7-7 are the changes in O_3-related risk that result from meeting the three alternative standards relative to meeting the existing standard. Furthermore, these changes reflect deaths per 100,000, which standardizes the estimates on

[38] Observations regarding the sensitivity of core short-term O_3-attributable mortality risk to these sensitivity analyses can be applied with care to the core short-term O_3-attributable morbidity endpoints, since many of these used similar air quality metrics in modeling risk.

population. This removes variation in the size of the underlying exposed population as a factor to consider in interpreting these results.

For the sensitivity analysis examining use of the smaller Smith et al., 2009 study area, we have also included heat maps similar to those used in conveying core estimates for short-term exposure related mortality (see section 7.5 for a description of the heat maps used in the core analysis). These heat maps (included in Appendix C – see Figure 7C-1) allow us to consider how changes in risk, including both reductions in risk and increases in risk are distributed across the O_3 air quality distributions for each study area.

Key observations related to the air quality-related sensitivity analyses include:

- **Use of smaller study area reduces magnitude of risk reduction**: For most of the study areas, use of the smaller Smith et al., 2009-based study area resulted in smaller risk reductions (again expressed in terms of changes in deaths per 100,000). For example, in Figure 7-7 (Baltimore plot), we see that estimated change in risk for SA1 (the smaller study area) are lower than estimated change in risk for the core scenario. This likely reflects the mix of monitors in the smaller study areas which results in a smaller change in the composite monitor value (for the existing standard versus alternative standard levels) as compared with composite monitor values based on the larger CBSA study area. However, it is important to keep the relative small magnitude of these risk reductions in mind when considering these sensitivity analysis results. Most of these differences in risk reductions are less than 1 individual per 100,000 which reflects the fact that total risk reduction (for short-term O_3-attributable mortality) across the urban study areas is relatively small (see Table 7-7).
    - **Reductions in risk are focused on higher O_3 days while increases are focused on lower O_3 days**: Figure 7C-1 allows us to consider patterns in risk reductions and increases when using the smaller Smith et al., 2009-based study areas in modeling risk. Figure 7C-1 (particularly the plots of risk decreases) suggests that decreases in risk tend to occur on days with composite monitor O_3 concentrations ranging from 40-70ppb, while increases in risk tend to occur on days with composite monitor values in the range at or below 30-40 ppb (with most risk increases falling in the range of 15ppb to 40ppb). As noted in 7.1.1, there is less confidence in specifying the nature of the C-R function (and therefore less confidence in specifying risk) in the range below 20 ppb.

- **Application of effect estimates derived for smaller study areas to larger CBSA-based study areas:** As noted in section 7.3.2, in those instances where an epidemiological study provides effect estimates for multiple subareas within a larger CBSA-based study area, we are selecting the effect estimates that represent the largest number of individuals to model that CBSA-based study areas. There is uncertainty associated with this approach. Specifically, as illustrated in Table 7-3, effect estimates within some of the CBSA-based study areas can display considerable heterogeneity. For example, consider the Smith et al., 2009-based effect estimates that fall within the CBSA-based New York study areas (these vary from 0.0001 to 0.0009 – almost a 10 fold factor, see Table 7-3). Furthermore, with the CBSA-based New York study area, Smith et al., 2009-based effect estimates only cover about half of the total population, with 8.3 million residents living within portions of the CBSA not covered by Smith et al., 2009-based effect estimates. As noted in section 7.3.2, in these types of situations, we have decided to use the single effect estimates representing the largest number of residents in modeling the larger CBSA-study area. This reflects the observation that, in the case of the New York CBSA, one of the available effect estimates (for the New York study area), represents ~7 times the population of the other effect estimates (see Table 7-3). In the case of the Los Angeles CBSA, there is significantly less difference between the available effect estimates, making the issue of heterogeneity (and the specification of a single effect estimate for this study area) less important. Never the less, we recognize that the issue of heterogeneity does complicate extrapolation of effect estimates for smaller study areas to the larger CBSA study areas modeled in this analysis and does introduce a degree of uncertainty that is difficult to characterize.
- **Use of alternative approach for simulating attainment of existing and alternative standard levels**: Use of an alternative approach to simulate attainment of the existing and alternative standard levels did not produce a consistent trend in terms of changes in risk between existing and alternative standards relative to the core analysis. For example, if we look at Figure 7-7 (plot for Houston), we see that SA2 (reflecting application of the alternative simulation approach) has a larger risk reduction than the core estimate. By contrast, if we look at the plot for Los Angeles, we see that the SA2 risk change is lower than the core estimate. Again, as with the sensitivity analysis results looking at study area size, it is important to keep in mind that the magnitude of these differences is relatively small, reflecting the small magnitude of mortality risk associated with these analyses in general (see Table 7-7). It is also important to note that in the alternative simulation approach, the HDDM-adjustment approach assumed

the same percent reductions of NOx and VOC and did not examine if a different air quality distribution could have been obtained with a different combination of NOx versus VOC reductions. For most of the urban areas, the percent NOx and VOC reductions were very similar to the NOx-only percent reductions. The similarity in the NOx reductions between the two approaches could be the reason for there being little difference in the risk estimates between the core and the alternative approach.

Sensitivity analyses related to specification of C-R functions

This category of sensitivity analysis covers a number of factors related to the specification of C-R functions for both short-term O_3-attributable and long-term O_3-attributable mortality. In the case of short-term O_3-attributable mortality, we consider (a) the use of Bayes adjusted effect estimates using regional priors (as contrasted with the Bayes adjusted values using a national prior applied in the core analysis), (b) the use of a copollutants model considering PM_{10} (as contrasted with the single pollutant model used in the core analysis) and (c) application of effect estimates from Zanobetti and Schwartz 2008 reflecting a summer focused analysis (as contrasted with the Smith et al., 2009-based analysis reflecting the entire monitoring period in each study area, which is used in the core analysis). For long-term O_3-attributable mortality, we consider the use of regionally-differentiated single pollutant effect estimates obtained from Jerrett et al., 2009, as contrasted with the single national copollutants model used in the core analysis (see section 7.1.1). We also present estimates for long-term O_3 attributable mortality based on application of results from a national level single pollutant model.

For sensitivity analyses examining alternative specification of the C-R function for short-term O_3-attributable mortality, we have used the same graphical approach as used in presenting results of the sensitivity analyses examining air quality characterization (i.e., plots of point estimates with 95th percentile C.I.s for the core and sensitivity analyses for each of the study areas – see Figure 7-8). Here we also plot estimates of risk changes using deaths per 100,000 to standardize in terms of total exposed population. For the sensitivity analysis considering alternative C-R functions for long-term O_3-attributable mortality, we present results in tabular form (Table 7-14). Specifically, for both the core and sensitivity analysis, we present (a) the percent of baseline mortality attributable to O_3 (under simulated attainment of the existing standard) and (b) the percent reduction in O_3-attributablerisk for each of the alternative standard levels. Key observations related to sensitivity analyses examining alternative C-R functions specifications include:

- **Use of regional Bayes-adjusted effect estimates in modeling short term O_3-attributable mortality**: The use of Bayes-adjusted effect estimates with regional

priors in modeling short-term O_3-attributable mortality, had a mixed impact across the urban study areas, with some study areas having increased changes in risk and others having smaller changes, relative to the core analysis. For example, in Figure 7-8 (plot for Baltimore), SA1 had a larger change in risk compared with the core analysis. However, as with the sensitivity analyses examining air quality-related factors (discussed above), it is important to keep in mind that the overall magnitude of the O_3-attributable mortality risk is relatively small and that these differences in changes in risk (comparing SA1 to the core analysis) are generally in the fraction of a person per 100,000 exposed population.

- **Use of a copollutants model (with PM_{10}) in modeling short-term O_3-attributable mortality:** The use of the PM_{10} copollutant model in modeling short-term O_3-attributable mortality (as contrasted with the single pollutant model used in the core analysis) tended to have a relatively small effect on estimates of risk changes for the alternative standards considered. For example, in Figure 7-8 (plot for Boston), we see that the estimates of risk changes for SA2 (reflecting application of the PM_{10} copollutant model) is essentially the same as the core risk estimate. It is important to keep in mind that the PM_{10} copollutant model suffers from significantly reduced power due to the 1/3 to 1/6 day sampling frequency used in measuring PM_{10} (this reduces the number of observations available to support epidemiological analysis). This has the impact of greatly increasing the confidence intervals on the SA2 risk estimates relative to the core estimates.
- **Use of Zanobetti and Schwartz 2008 effect estimates in modeling short-term O_3-attributable mortality**: The use of Zanobetti and Schwartz, 2008 effect estimates (reflecting a focus on the warmer summer months) produces a mixed set of results when compared to the core risk estimates. If we look at Figure 7-8 we see that, for Boston, estimates of risk changes for SA3 (reflecting application of the Zanobetti and Schwartz 2008 effect estimates) are significantly larger than core estimates. By contrast, SA3 estimates of risk changes for Houston are significantly smaller than the core estimates. It is important, however to keep in mind that the Zanobetti and Schwartz 2008 effect estimates will tend to under-estimate total risk since they only model impacts during the summer months (while the Smith et al., 2009 effect estimates allow us to model impacts for the entire O_3 monitoring season in each study area). Note that if the O_3 effect were only occurring during the summer months, then the total risk estimated using effect estimates from the two studies would be similar. However, because the risks in many locations are smaller (using the Zanobetti and

Schwartz, 2008 based effect estimates), this suggests that the O_3 effect occurs outside of the summer months evaluated in this study.

- **Use of regional-differentiated effect estimates in modeling long-term O_3-attributable mortality**: Risk estimates generated using regional-specific effect estimates for long-term O_3-attributable mortality differ substantially from the core estimates based on a single national-level effect estimate (see Table 7-14). Furthermore, the risk estimates generated using the regional effect estimates display considerable variability (see Table 7-14) reflecting the significant variability in the underlying effect estimates (see Jerrett et al., 2009, Table 4). The regional effect estimates range from 0.99 (for the Northeast) to 1.21 (for the Southwest) and include 1.00 (no O_3 effect for the Industrial Midwest). As noted earlier in section 7.5, negative risk estimates should not be interpreted as suggesting that O_3 exposure is beneficial. Rather, these suggest that there may be instability in the underlying estimates or that potential confounding has not been fully addressed. Regional effect estimates used in this analysis have considerably larger confidence intervals than the national estimate (compare values in Jerrett et al., 2009 Table 3 with values in Table 4). This suggests that the regional estimates are less stable than the national estimates and are subject to considerably greater uncertainty. For this reason, while the results of this sensitivity analysis point to the potential for regional heterogeneity in the long-term O_3-attributable mortality effect estimate, we do not have significant confidence in the regionally-based risk estimates themselves given the relatively large confidence intervals associated with those estimates.
- **Use of national-based single pollutant model in modeling long-term O_3-attributable mortality**: Risk estimates generated using the national-level O_3-only effect estimate were significantly lower (~30%) than the core risk estimates which utilize a copollutants model (which includes $PM_{2.5}$) (see Table 7-15). In this case, control for another pollutant results in a stronger O_3 signal, possibly due to an association between $PM_{2.5}$ and a confounder or effect modifier associated with the O_3-related effect.

Figure 7-7 Sensitivity Analysis: Short-Term O_3-attributable Mortality (air quality-related factors including study area size and method used to simulate attainment of existing and alternative standard levels) (2009) SA1-smaller (Smith-based) study area, SA2-alternative method for simulating standards.

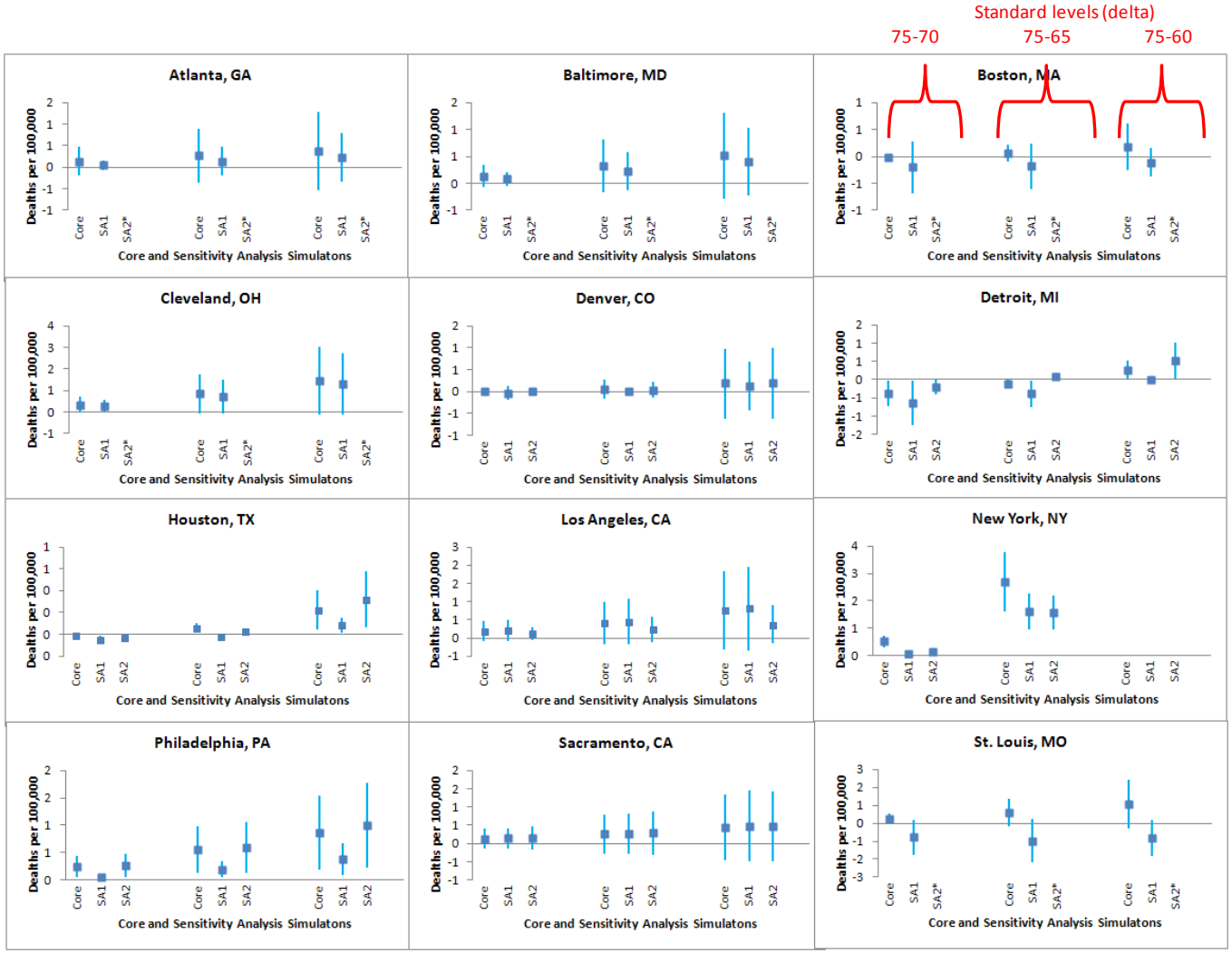


Figure 7-8 Sensitivity Analysis: Short-Term O_3-attributable Mortality (C-R function specification) (2009) SA1-regional Bayes-based adjustment; SA2-copollutant model (PM_{10}); SA3-Zanobetti and Schwartz-based effect estimates.

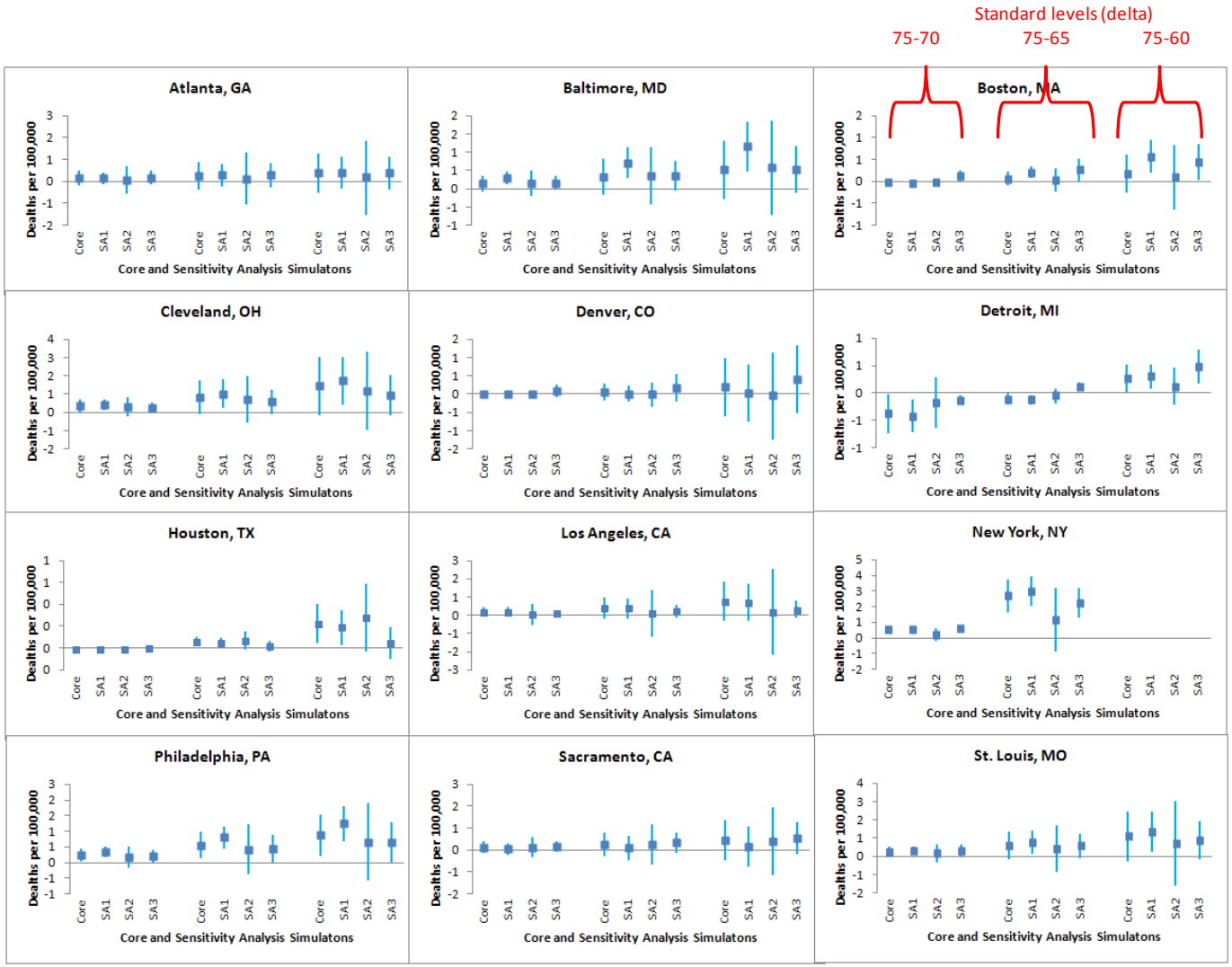


Table 7-14 Sensitivity Analysis for Long-Term O_3-attributable Respiratory Mortality – Alternative C-R Function Specification (regional effect estimates) *% of baseline all-cause mortality and change in O_3-attribuable risk* (2009) (Smith et al., 2009, O_3 season))

Study Area	Air Quality Scenario			
	Baseline Incidence Attributable to Ozone	Change in O_3-Attributable Risk		
	75ppb	75-70	75-65	75-60
Core Simulation				
Atlanta, GA	16.6	5	10	14
Baltimore, MD	17.4	3	6	10
Boston, MA	15.9	1	3	7
Cleveland, OH	16.8	4	9	15
Denver, CO	20.0	1	5	12
Detroit, MI	17.0	-1	2	6
Houston, TX	16.9	2	4	7
Los Angeles, CA	20.7	4	8	13
New York, NY	16.7	5	18	18
Philadelphia, PA	17.2	3	7	10
Sacramento, CA	18.0	4	8	12
St. Louis, MO	17.7	4	8	13
Sensitivity analysis				
Atlanta, GA	41.21	4	8	11
Baltimore, MD	-7.01	4	9	13
Boston, MA	-6.19	1	5	9
Cleveland, OH	0.00	0	0	0
Denver, CO	27.38	1	4	11
Detroit, MI	0.00	0	0	0
Houston, TX	41.15	2	3	6
Los Angeles, CA	4.46	3	6	10
New York, NY	-6.61	7	24	NA
Philadelphia, PA	-6.89	4	9	13
Sacramento, CA	24.79	4	7	11
St. Louis, MO	0.00	0	0	0

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

Table 7-15 Sensitivity Analysis for Long-Term O_3-attributable Respiratory Mortality – Alternative C-R Function Specification (national O_3-only effect estimates) *% of baseline all-cause mortality and change in O_3-attributable risk* (2009) (Smith et al., 2009, O_3 season))

Study Area	Air Quality Scenario			
	Percent of Baseline Incidence	Change in O_3-Attributable Risk		
	75ppb	75-70	75-65	75-60
Core Simulation				
Atlanta, GA	16.6	5	10	14
Baltimore, MD	17.4	3	6	10
Boston, MA	15.9	1	3	7
Cleveland, OH	16.8	4	9	15
Denver, CO	20.0	1	5	12
Detroit, MI	17.0	-1	2	6
Houston, TX	16.9	2	4	7
Los Angeles, CA	20.7	4	8	13
New York, NY	16.7	5	18	18
Philadelphia, PA	17.2	3	7	10
Sacramento, CA	18.0	4	8	12
St. Louis, MO	17.7	4	8	13
Sensitivity analysis				
Atlanta, GA	11.9	5	10	14
Baltimore, MD	12.2	3	7	10
Boston, MA	11.1	1	3	7
Cleveland, OH	11.8	4	10	15
Denver, CO	14.1	2	5	12
Detroit, MI	11.9	-1	2	6
Houston, TX	11.9	2	4	7
Los Angeles, CA	14.6	4	9	14
New York, NY	11.7	5	19	19
Philadelphia, PA	12.0	3	7	10
Sacramento, CA	12.6	4	8	13
St. Louis, MO	12.4	4	8	13

NA: for NYC, the model-based adjustment methodology was unable to adjust O_3 distributions such that they would meet the lower alternative standard level of 60 ppb.

7.6 KEY OBSERVATIONS REGARDING OVERALL CONFIDENCE IN THE RISK ASSESSMENT AND RISK ESTIMATES

This section discusses our overall confidence associated with risk estimates presented in this draft of the REA. We begin by presenting a set of key observations related to overall confidence in the risk assessment. These observations are drawn largely from (a) consideration for the systematic approach used in designing the risk assessment, (b) our assessment of the degree to which we have captured key sources of variability in the analysis (section 7.4.1) (c) our qualitative assessment of uncertainty in the risk assessment (section 7.4.2), and (d) the results of the sensitivity analyses completed (section 7.5.3). Once we present these observations, we provide a synthesis statement reflecting our overall degree of confidence in the risk estimates (at the end of this section). Key observations addressing overall confidence in the analysis include:

- A deliberative process was used in specifying each of the analytical elements comprising the risk model. This is in line with recommendations made by the National Research Council in *Science and Decisions, Advancing Risk Assessment* (NRC, 2009. P. 89-90) for improving risk assessment as applied in the regulatory context. This deliberative process included first identifying specific goals for the analysis, and then designing the analysis to meet those goals, given available information and methods. Specific analytical elements reflected in the design include: selection of urban study areas, characterization of ambient air O_3 levels, selection of health endpoints to model and selection of epidemiological studies (and specification of C-R functions) (see sections 7.1.1 and 7.3). In addition, the design of this draft of the REA reflects consideration for comments provided by the public and by CASAC in their review of the 1st draft REA in letter form (Frey, H. C. 2012.).
- Review of available literature (as specified in the O_3 ISA, U.S. EPA. 2013a), resulted in a decision not to incorporate a true (no effect) threshold into our risk modeling. Conversely, the studies used to develop the C-R functions indicate a range of ambient O_3 (area-wide daily levels, based on averaging across monitors in locations with multiple monitors, of $\leq$ 20ppb) below which there is reduced confidence in specifying the nature of the concentration-response relationship, based on less data in the studies, specifically for short-term O_3-attributable respiratory mortality and morbidity (see section 7.1.1). In any case, only a relatively small fraction of short-term O_3-attributable mortality reflected in the risk estimates is associated with days in this range with the vast majority of the risk estimates reflecting days with peak O_3 measurements well above this level (see section 7.5.1 and 7.5.2). O_3
- Modeling of short-term O_3-attributable mortality utilized Bayes-adjusted city-specific effect estimates (see section 7.1.1 and section 7.3.2). These effect estimates are

considered to have increased overall confidence since they combine elements of the local city-specific signal with a broader scale (national) signal.

- Use of CBSA-based study areas in modeling all health endpoints in order to address known bias associated with using smaller study areas. As discussed in 7.1.1, we have used larger CBSA-based study areas to avoid focusing the risk assessment only on core urban areas (often used in the epidemiological studies providing effect estimates) which can experiences increases in O_3 based on simulated attainment of both existing and alternative standard levels. There is uncertainty in using effect estimates based on smaller study areas to represent larger CBSA-based study areas (see section 7.4.2 and 7.5.3). A key concern is heterogeneity in the effect estimates which may suggest increased uncertainty in applying effect estimates to larger study areas (since larger study areas may display heterogeneity in the nature of the relationship between O_3 exposure and risk). It is possible also that this heterogeneity varies across urban areas, or regionally. For both categories of mortality endpoints (short-term and long-term O_3-attributable), potential heterogeneity in the mortality effect even within larger urban areas remains a potentially important source of uncertainty.
- Specifically in relation to short-term exposure-related mortality and morbidity which depend on time-series studies, there is uncertainty in applying effect estimates derived based on evaluating the longitudinal (in terms of time) relationship between ambient O_3 and a particular health effect to the modeling of a discrete shift in the entire distribution that occurs when you simulate an alternative standard. Specifically, the time-series studies relate unit changes in day to day O_3 with a degree of impact on baseline health effect rates. In the risk assessment, we use this effect estimate to predict risk for a unit shift in daily composite monitor value. There is uncertainty in this application of the effect estimates, although it is not possible at this time to characterize either qualitatively, or quantitatively the magnitude of this uncertainty and the degree of any potential bias that could be introduced into the simulation of risk.
- Use of HDDM-adjustment approach to simulate attainment of both the existing and alternative standard levels provides more refined estimates of ambient O_3 distributions given its ability to characterize the physical and chemical processes of O_3 formation in the atmosphere However, in the case of both the New York and Los Angeles study areas, given the limitations in the application of the adjustment methodology to very large emissions perturbations and the need to use the 95th percent confidence interval lower bound estimate to simulate attainment of these standard levels, we have reduced overall confidence in the simulation of the O_3

concentrations for these study areas and consequently all health endpoints modeled for risk for these two study areas (see section 7.4.2 and 7.5.3).

- Sensitivity analyses exploring alternative C-R functions for modeling short-term O_3-attributable mortality (e.g., Bayes regional prior based estimates, copollutants models) suggested that alternative models can have a moderate impact on risk (see section 7.5.3). This modest impact reflects primarily the relatively small magnitude of short-term O_3-attributable mortality reductions simulated for the alternative standard levels.
- The use of alternative C-R functions for modeling long-term O_3-attributable mortality (specifically the regional-based estimates referenced earlier) was shown to have a significant impact on risk (see section 7.5.3). However, concerns over the power and hence stability of the regional effect estimates used in this simulation limit our ability to draw firm conclusions regarding the potential magnitude of that regional heterogeneity.

Based on the key observations regarding confidence presented above, we draw the following conclusions regarding overall confidence in the risk estimates generated for this draft of the REA. We have a reasonable degree of confidence in short-term O_3-attributable mortality and morbidity estimates for ten of the twelve study areas. This confidence is tempered somewhat by concerns over potential heterogeneity in effect estimates for mortality which can impact the risk assessment given our use of larger CBSA-based study areas. Our confidence in risk estimates generated for both New York and Los Angeles is considerably lower than for the remaining ten study areas due to (a) concerns over air quality modeling (specifically the use of lower-bound fits to the DDM model) and (b) specifically in the case of New York, evidence for significant heterogeneity in the mortality effect estimates for subareas within the CBSA. For long-term O_3-attributable mortality, we also have a reasonable degree of confidence in our risk estimates. However, as with short-term O_3-attributable mortality, this confidence is also tempered by concerns over regional heterogeneity in the O_3 effect. If we had regionally-differentiated effect estimates for this endpoint that had sufficient power and stability, we would consider using these as the basis for generating core risk estimates (rather than the national-level effect estimate used in the current analysis).

7.7 REFERENCES

Abt Associates Inc. 1996. *A Particulate Matter Risk Assessment for Philadelphia and Los Angeles*. Prepared for Office of Air Quality Planning and Standards. Research Triangle Park, NC: EPA Office of Air and Radiation, OAQPS, July 1996. Available at: <http://www.epa.gov/ttnnaaqs/standards/pm/data/jly3amd.pdf>.

Abt Associates, Inc. 2010. *Environmental Benefits and Mapping Program (Version 4.0)*. Prepared for U.S. Environmental Protection Agency, Bethesda, MD. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. Available on the Internet at: <http://www.epa.gov/air/benmap>.

Akinbami, LJ; C. D. Lynch; J. D. Parker and T. J. Woodruff. 2010. "The Association Between Childhood Asthma Prevalence and Monitored Air Pollutants in Metropolitan Areas, United States, 2001-2004." *Environmental Research,*110: 294-301.

Bell, M. L. and F. Dominici. 2008. "Effect Modification by Community Characteristics on the Short-term Effects of O_3 Exposure and Mortality in 98 U.S. Communities." *American Journal of Epidemiology*, 167: 986-997.

Bell, M. L.; A. McDermott; S. L. Zeger; J. M. Samet; F. Dominici. 2004. "O_3 and Short-term Mortality in 95 U.S. Urban Communities, 1987-2000." *JAMA*, 292: 2372-2378.

Darrow, L. A.; M. Klein; J. A. Sarnat; J. A. Mulholland; M. J. Strickland; S. E. Sarnat, et al. 2011. "The Use of Alternative Pollutant Metrics in Time-series Studies of Ambient Air Pollution and Respiratory Eemergency Department Visits." *Journal of Exposure Science and Environmental Epidemiology*, 21, 10-19.

Frey, H. C. 2012. "Letter from Clean Air Scientific Advisory Committee to the Honorable Lisa P. Jackson, Administrator, US EPA. CASAC Review of the EPA's Health Risk and Exposure Assessment for O_3 (First External Review Draft - Updated August 2012) and Welfare Risk and Exposure Assessment for O_3 (First External Review Draft - Updated August 2012)," November, 19, 2012.

Gent, J. F.; E. W. Triche; T. R. Holford; K. Belanger; M. B. Bracken; W. S. Beckett, et al. 2003. "Association of Low-level O_3 and Fine Particles with Respiratory Symptoms in Children with Asthma." *Journal of the American Medical Association*, 290(14), 1859-1867.

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." *Journal of Exposure Science and Environmental Epidemiology*, 17 Suppl 2, S45-60.

Jerrett, M.; R. T. Burnett; C. A. Pope, III; K. Ito; G. Thurston; D. Krewski; Y. Shi and E. Calle; M. Thun. 2009. "Long-term O_3 Exposure and Mortality." *New England Journal of Medicine*, 360:1085-1095.

Katsouyanni, K.; J. M. Samet; H. R. Anderson; R. Atkinson; A. L. Tertre; S. Medina, et al. 2009. "Air Pollution and Health: A European and North American Approach (APHENA)," Health Effects Institute.

Lin, S.; E. M. Bell; W. Liu; R. J. Walker; N. K. Kim and S. A. Hwang. 2008a. "Ambient O_3 Concentration and Hospital Admissions Due to Childhood Respiratory Diseases in New York State, 1991– 2001." *Environmental Research*, 108, 42-47.

Lin, S; X. Liu; L. H. Le; S. A. Hwang. 2008. "Chronic Exposure to Ambient O_3 and Asthma Hospital Admissions Among Children." *Environonmental Health Perspective*, 116: 1725-1730.

Linn, W. S.; Y. Szlachcic; H. Gong, Jr.; P. L. Kinney and K. T. Berhane. 2000. "Air Pollution and Daily Hospital Admissions in Metropolitan Los Angeles." *Environmental Health Perspective*, 108(5), 427-434.

Medina-Ramon, M.; A. Zanobetti and J. Schwartz. 2006. "The Effect of O_3 and PM_{10} on Hospital Admissions for Pneumonia and Chronic Obstructive Pulmonary Disease: A National Multicity Study." *American Journal of Epidemiology*, 163(6), 579-588.

Meng, Y. Y.; R. P. Rull; M. Wilhelm; C. Lombardi; J. Balmes and B. Ritz. 2010. "Outdoor Air Pollution and Uncontrolled Asthma in the San Joaquin Valley, California." *Journal of Epidemiology Community Health*, 64: 142-147.

Moore, K; R. Neugebauer; F. Lurmann; J. Hall; V. Brajer; S. Alcorn; I. Tager. 2008. "Ambient O_3 Concentrations Cause Increased Hospitalizations for Asthma in Children: An 18-year Study in Southern California." *Environmental Health Perspective*, 116: 1063-1070.

National Research Council (NRC). 2009. *Science and Decisions, Advancing Risk Assessment. Committee on Improving Risk Analysis Approaches*. Washington, DC: The National Academies Press.

Silverman, R. A.; and K. Ito. 2010. "Age-related Association of Fine Particles and O_3 with Severe Acute Asthma in New York City." *Journal of Allergy Clinical Immunology*, 125(2), 367-373 e365.

Smith, R.L.; B. Xu and P. Switzer. 2009. "Reassessing the Relationship Between O_3 and Short- term Mortality in U.S. Urban Communities." *Inhalation Toxicology*, 21: 37-61.

Strickland, M. J.; L. A. Darrow; M. Klein; W. D. Flanders; J. A. Sarnat; L. A. Waller, et al. 2010. "Short-term Associations between Ambient Air Pollutants and Pediatric Asthma Emergency Department Visits." *American Journal of Respiratory Critical Care Medicine*, 182, 307-316.

Tolbert, P. E.; M. Klein; J. L. Peel; S. E. Sarnat and J. A. Sarnat. 2007. "Multipollutant Modeling Issues in a Study of Ambient Air Quality and Emergency Department Visits in Atlanta." *Journal of Exposure Science and Environmental Epidemiology*, 17 Suppl 2, S29-35.

U.S. Center for Disease Control. 2010. "Centers for Disease Control and Prevention, Behavioral Risk Factor Surveillance System (BRFSS), 2010, Table "Table C1 Adult Self-Reported Current Asthma Prevalence Rate (Percent) and Prevalence (Number) by State or Territory." Available at: <http://www.cdc.gov/asthma/brfss/2010/current/tableC1.htm>.

U.S. Environmental Protection Agency. 2001. *Risk Assessment Guidance for Superfund. Vol. III, Part A. Process for Conducting Probabilistic Risk Assessment (RAGS 3A)*. Washington, DC: EPA. (EPA document number EPA 540-R-02-002; OSWER 9285.7-45; PB2002 963302). Available at: <http://www.epa.gov/oswer/riskassessment/rags3adt/index.htm)>.

U.S. EPA. 2004. *EPA's Risk Assessment Process for Air Toxics: History and Overview*. In: Air Toxics Risk Assessment Reference Library, Technical Resource Manual, Vol. 1., pp. 3-1 – 3-30. (EPA document number EPA-453-K-04-001A). Washington, DC: EPA. Available at: <http://www.epa.gov/ttn/fera/data/risk/vol_1/chapter_03.pdf)>.

U.S. EPA. 2007. O_3 *Health Risk Assessment for Selected Urban Areas*. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. (EPA document number EPA 452/R-07-009). Available at: <http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_cr.html>.

U.S. EPA. 2011. O_3 *National Ambient Air Quality Standards: Scope and Methods Plan for Health Risk and Exposure Assessment*. Research Triangle Park, NC: EPA. (EPA document number EPA-452/P-11-001).

U.S. EPA. 2012. *Health Risk and Exposure Assessment for* O_3 *First External Review Draft*. Research Triangle Park, NC: U.S. Environmental Protection Agency, Research Triangle Park, NC. (EPA document number EPA 452/P-12-001).

U.S. EPA. 2013a. *Integrated Science Assessment for* O_3 *and Related Photochemical Oxidants: Final*. Research Triangle Park, NC: U.S. Environmental Protection Agency. (EPA document number EPA/600/R-10/076F).

U.S. EPA. 2013b. *Environmental Benefits Mapping and Analysis Program—Community Edition (Version 0.63), 2013a*. Research Triangle Park, NC: U. S. Environmental Protection Agency. Available at: <hhtp://www.epa.gov/air/benmap/beta.html>.

World Health Organization. 2008. *Part 1: Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment, Harmonization Project Document No. 6*. Published under joint sponsorship of the World Health Organization, the International Labour Organization and the United Nations Environment Programme. WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 2476).

Zanobetti, A; J. Schwartz. 2008. "Mortality Displacement in the Association of O_3 with Mortality: An Analysis of 48 Cities in the United States." *American Journal of Respiratory and Critical Care Medicine*, 177: 184-189.

8 NATIONAL SCALE MORTALITY RISK BURDEN BASED ON APPLICATION OF RESULTS FROM EPIDEMIOLOGICAL STUDIES

As described in Chapter 2, the O_3 ISA (U.S. EPA 2013) concluded that there is likely to be a causal relationship between short-term O_3 exposure and all-cause mortality and that there is likely to be a causal relationship between long-term O_3 exposure and respiratory effects, including respiratory mortality. Chapter 7 estimated health risks associated with recent O_3 concentrations and meeting the current and alternative O_3 standards in 12 selected urban study areas. In this chapter we estimate nationwide premature mortality attributable to recent short-term and long-term exposures to ambient O_3 (Section 8.1); and assess the degree to which the selected urban case study areas represent the full national distribution of risk-related attributes and air quality dynamics (Section 8.2). Compared with the urban scale analysis in Chapter 7, this analysis includes full spatial coverage across the U.S. but has less geographic specificity in the concentration-response functions that are used to calculate O_3-attributable mortality. The national scale analysis is therefore intended as a complement to the urban scale analysis, providing both a broader assessment of O_3-related health risks across the U.S. as well as an evaluation of how well the urban study areas examined in Chapter 7 represent the full distribution of O_3-related health risks and air quality dynamics in the U.S.

8.1 NATIONAL-SCALE ASSESSMENT OF MORTALITY RELATED TO O_3 EXPOSURE

This section estimates the total annual deaths for 2007 populations associated with average 2006-2008 O_3 levels across the continental U.S. We first describe the methods and inputs used to estimate O_3-attributable risk across the continental U.S., including O_3 exposure estimates, population and baseline mortality rate estimates, and epidemiologically derived O_3-mortality effect estimates. Results for the estimation of O_3-attributable risk are then discussed in terms of the magnitude and percent of total mortality attributable to O_3 exposure. We provide two analyses to give perspective on the confidence in the estimates of O_3-related mortality: (1) risk estimated only within the urban areas for which O_3 mortality effect estimates are available; and (2) the distribution of O_3-related deaths across the range of observed 2006-2008 average O_3 concentrations fused with modeled 2007 concentrations. These results are then synthesized and compared with previous estimates of the burden of O_3 exposure on mortality in the U.S. from the literature in a discussion section.

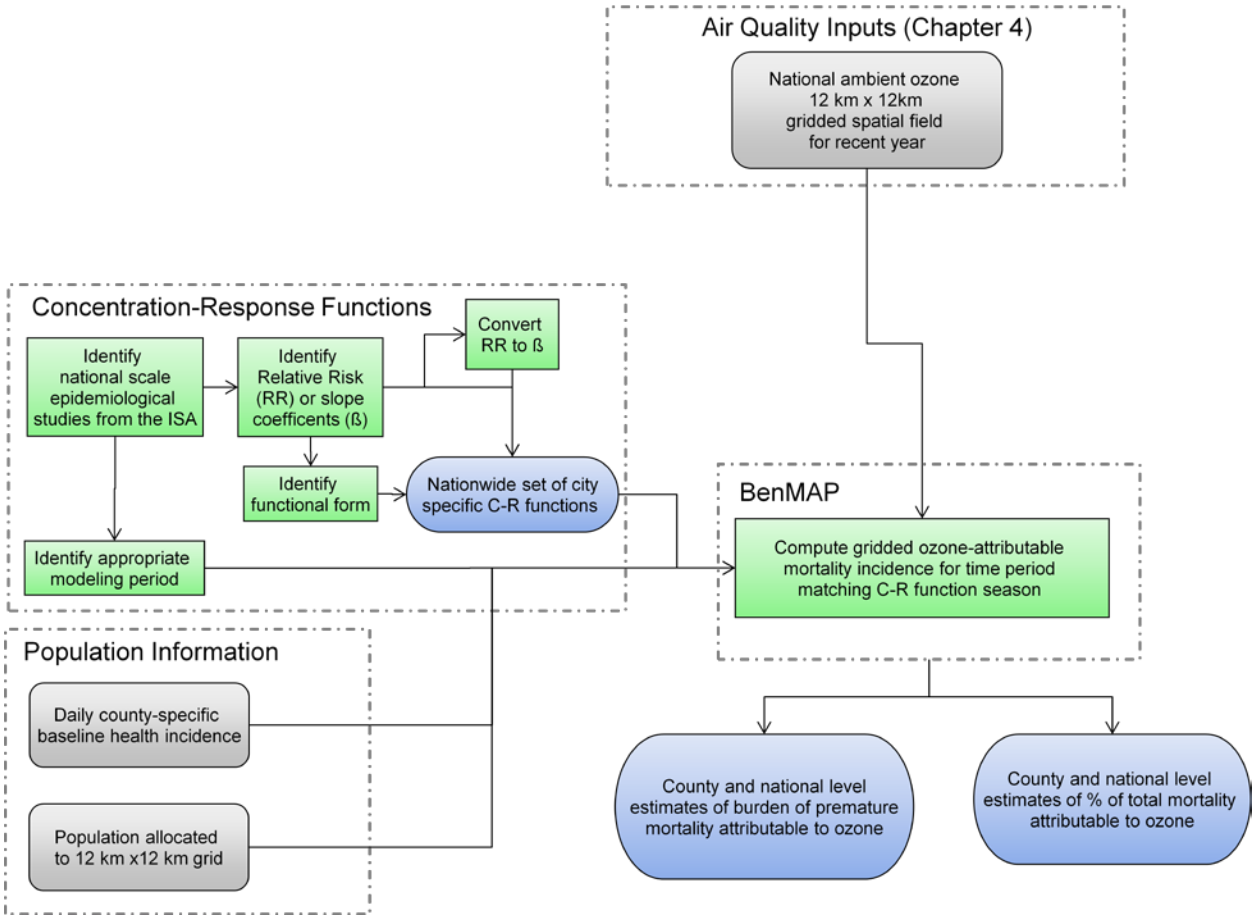


Figure 8.1 Conceptual Diagram for National-scale Mortality Risk Assessment

8.1.1 Methods

This section describes the inputs and datasets used to conduct the national-scale assessment of O_3-attributable risk. As shown in the conceptual diagram in Figure 8-1, we conduct this analysis using the BenMAP software, which uses projections of the size and geographic distribution of the potentially exposed population along with estimates of the ambient O_3 concentrations to estimate O_3-attributable health risks. In general, this analysis uses the same analytical structure and many of the same inputs as are used in the epidemiology-based assessment of O_3-attributable risk in the selected urban case study areas in Chapter 7. We refer back to Chapter 7 for details on these shared inputs, and describe where the urban-scale and national-scale analyses use divergent methods.

8.1.1.1 Ambient O_3 Concentrations

Air quality inputs to this analysis are described in detail in Chapter 4. In contrast to the urban study areas analysis in Chapter 7, the national-scale analysis employs a data fusion approach that takes advantage of the accuracy of monitor observations and the comprehensive spatial information of the CMAQ modeling system to create national-scale "fused" spatial surfaces of seasonal average O_3. Measured O_3 concentrations from 2006-2008 were fused with modeled concentrations from a 2007 CMAQ model simulation, run for a 12 km domain covering the contiguous U.S. In the first draft of the REA, the spatial surfaces were created using the enhanced Voronoi Neighbor Averaging (eVNA) technique (Timin et al, 2010), using the EPA's Model Attainment Test Software (MATS; Abt Associates, 2010b). In this draft, the spatial surfaces are created using EPA's Downscaler software (Berrocal et al, 2012). More details on the ambient measurements, the 2007 CMAQ model simulation, the Downscaler fusion technique, and a technical justification for changing from eVNA to Downscaler can be found in Chapter 4.

Three "fused" spatial surfaces were created for: (1) the May-September mean of the 8-hr daily maximum (consistent with the metric used by Smith et al. 2009); (2) the June-August mean of the 8-hr daily mean from 10am to 6pm (consistent with the metric used by Zanobetti and Schwartz 2008); and (3) the April-September mean of the 1-hr daily maximum (consistent with the metric used by Jerrett et al. 2009) O_3 concentrations across the continental U.S. These fused spatial surfaces each represent one seasonal average across 2006-2008, rather than three separate years of concentrations. Section 4.3.2 presents maps, distributions, and statistical characterizations of these O_3 concentrations metrics across the U.S., including how they compare to 2006-2008 design values.

8.1.1.2 Concentration-Response Functions

While Chapter 7 assessed both mortality and morbidity risks associated with O_3 concentrations, due to limitations in baseline morbidity incidence rates, the national scale assessment focuses on mortality risks only. To quantify the impact of O_3 concentrations on mortality, we apply risk estimates drawn from two major short-term epidemiological studies and one long-term epidemiological study. These studies are consistent with those used in the analysis of O_3-related risk in selected urban areas (Section 7.2) and those mortality endpoints concluded to have a causal or suggestive causal relationship with O_3 exposure by the 2013 Integrated Science Assessment for O_3 and Related Photochemical Oxidants (U.S. EPA 2013).

For short-term mortality, we use city-specific and national average risk estimates drawn from the Smith et al. (2009) study of O_3 and mortality in 98 U.S. urban communities between 1987 and 2000 as our main results, and the Zanobetti and Schwartz (2008) study of O_3 and

mortality in 48 U.S. cities between 1989 and 2000 as a sensitivity analysis, consistent with the urban case study analysis in Chapter 7. City-specific effect estimates for both studies are provided in Appendix 4-A.

Smith et al. (2009) found that the average non-accidental mortality increase across all 98 urban areas was 0.32% ± 0.08 (95% posterior interval [PI], 0.41%-0.86%) for a 10 ppb increase in the 8-hr daily maximum O_3 concentration, based on April to October O_3 observations. Since the national-scale analysis requires a single modeling period definition but some monitors only collect data from May to September, the corresponding city-specific effect estimates are applied to each day from May to September in BenMAP using May to September average 8-hr daily maximum O_3 concentration based on 2006-2008 observed concentrations fused with 2007 modeled concentrations. The length of the O_3 season can affect the magnitude of mortality effect estimates – a longer season may yield higher effect estimates per unit O_3 concentration since O_3 concentrations over the longer season may be lower than the O_3 concentrations over the warmest months only. Conversely, if the longer period captures periods of lower O_3-related mortality incidence, the effect estimates may be lower than effect estimates for the warmest months only. Our application of the Smith et al. (2009) April to October effect estimates to May to September O_3 concentrations likely introduces some bias in the results, but it is unclear in which direction.

Zanobetti and Schwartz (2008) found that the average total mortality increase across all 48 cities was 0.53% (95% confidence interval, 0.28%-0.77%) for a 10 ppb increase in June-August 8-hr daily mean O_3 concentration from 10 am to 6 pm, using a 0-3 day lag. We apply the city-specific effect estimates that correspond to this national average effect estimate each day from June to August in BenMAP using the June to August, mean 8-hr daily mean O_3 concentration based on 2006-2008 observed concentrations fused with 2007 modeled concentrations. Consistent with Chapter 7, these results are presented as a sensitivity analysis.

As in Chapter 7, we use city-specific risk estimates from the short-term epidemiology studies, but apply them here only to the counties that were included in the epidemiology studies rather than to the entire core-based statistical area (CBSA). Chapter 7 estimated risk across entire CBSAs to more completely capture expected O_3 changes across broader areas and avoid bias resulting from including only those areas where O_3 is expected to increase under alternative standards. The inclusion of the entire CBSA in that analysis required the application of a single effect estimate to the entire CBSA. However, the national-scale assessment is a gridded analysis, which allows greater spatial resolution in the application of effect estimates. In addition, eight CBSAs nationwide included multiple cities defined separately by Smith et al. (2009), some of which showed considerable heterogeneity in effect estimates within the same CBSA. Heterogeneity among effect estimates within a single CBSA implies that effect estimates from one county may not be accurate representations of effect estimates in nearby

counties. However, since city-specific effect estimates often have low power due to small population size, we are unable to draw a strong conclusion regarding how well one county's effect estimates represents those in nearby counties. For this national-scale assessment, we apply effect estimates from each city as defined in the epidemiology studies to retain the full set of information available from those studies. In addition, for counties not included by the epidemiology studies, we apply the average effect estimate derived from all the urban areas included in each of the studies ("national average") as it takes advantage of a wider and more diverse population.

Since both national average estimates from these studies are based on urban areas only, we have higher confidence in their application to other U.S. urban areas than to rural areas. To demonstrate the magnitude of the results for which we have the highest confidence, we present the percentage of estimated deaths occurring within the urban areas included in the epidemiological studies and within all urban areas across the U.S. Lower confidence in the results for rural areas does not indicate that the mortality risk among populations living in such areas is unaffected by O_3 pollution. Rather, the level of understanding for the O_3-mortality relationship in these areas is simply lower due to a lack of available epidemiological data at these levels. We also examine the effect of varying the effect estimate applied between the cities included by the epidemiology studies in a sensitivity analysis.

We quantify long-term O_3-related respiratory mortality in this REA since the Integrated Science Assessment for O_3 and Related Photochemical Oxidants (O_3 ISA) concluded that the evidence supports a likely to be causal relationship between long-term O_3 exposure and respiratory effects, including respiratory morbidity and respiratory-related mortality (U.S. EPA, 2013). As detailed in Chapter 7, we quantify long-term O_3-related mortality using the respiratory mortality effect estimates from the Jerrett et al. (2009) two-pollutant model that controlled for $PM_{2.5}$ concentrations, applied to each gridcell across the entire United States. This model found that a 10 ppb increase in the April-September average of the 1-hr daily maximum O_3 concentration was associated with a 4% (95% confidence interval, 1.0%-6.7%) increase in respiratory mortality.

8.1.1.3 Demographic Inputs

This analysis uses the same baseline mortality rates and population estimates as were used in the urban case study area analysis in Chapter 7. We derive baseline incidence rates for mortality by age, cause, and county from the CDC Wonder database (CDC, 2004-2006). As this database only provides baseline incidence rates in 5-year increments, we use data for the year 2005, the closest year to the analysis year 2007 used for the population and air quality modeling.

We use 2007 population because it matches both the year of the emissions inventory and meteorology used for the air quality modeling.

The starting point for estimating the size and demographics of the potentially exposed population is the 2010 census-block level population, which BenMAP aggregates up to the same grid resolution as the air quality model. BenMAP back-casts this 2010 population to the analysis year of 2007 using county-level growth factors based on economic projections (Woods and Poole Inc., 2012).

8.1.2 Results

Table 8.1 summarizes the estimated O_3-related premature mortality associated with 2006-2008 average O_3 concentrations under various assumptions for the health impact function. Applying Smith et al. (2009) effect estimates for May-September, we estimate 15,000 (95% CI, 1,400-28,000) premature O_3-related non-accidental deaths annually for 2007. As a sensitivity analysis, we apply Zanobetti and Schwartz (2008) effect estimates for June-August, finding 16,000 (95% CI, 6,000-25,000) premature O_3-related all-cause deaths annually for 2007. Figure 8.2 Figure 8.4 show that estimated O_3-related mortality is most concentrated in highly populated counties or those counties with urban areas found to have high effect estimates by Smith et al. (2009) or Zanobetti and Schwartz (2008). For the application of Jerrett et al. (2009) national average effect estimate for April-September, we estimate 45,000 (95% CI, 17,000-70,000) premature O_3-related respiratory deaths among adults age 30 and older.

Because the epidemiological studies included only selected urban areas, we are more confident in the magnitude of the estimated O_3-related deaths occurring within those urban areas. As shown in Table 8.1, approximately 43% of the O_3-related deaths estimated using Smith et al. (2009) effect estimates occur in the 98 urban locations included in that study, and 30% of the O_3-related deaths estimated using Zanobetti and Schwartz (2008) effect estimates occur in the 48 urban areas included in that study. We are also more confident in extrapolating the national average effect estimates to other urban areas than we are to rural areas, as the national average estimates are based on all urban areas included by the study. To estimate the percentage of total O_3-attributable deaths occurring within all urban areas across the continental U.S., we sum the results for the 12km gridcells that have a total population greater than 12,000 (approximately equal to the 95th percentile of gridcell populations across the continental U.S.). The percentage of O_3-attributable deaths occurring within urban areas defined in this way is 65% for results based on Smith et al. (2009) effect estimates and 64% for results based on Zanobetti and Schwartz (2008) effect estimates. While our confidence is lower when the national average effect estimates

are extrapolated to rural areas, less certainty in the magnitude of O_3-related deaths in rural areas does not imply that O_3 has no effect on health in these areas.

Table 8-1 Estimated annual O_3-related premature mortality in 2007 associated with 2006-2008 average O_3 concentrations (95th percentile confidence interval)

Source of risk estimate and modeling period	Exposure duration	Age	City-specific effect estimates[1]	National average effect estimate[2]
Smith et al. (2009), May-September	Short-term	>0	15,000	16,000
95% confidence interval			(1,400-28,000)	(7,200-22,000)
% occurring within the 98 cities			43%	
Zanobetti and Schwartz (2008), June-August	Short-term	>0	16,000	15,000
95% confidence interval			(6,000-25,000)	(8,300-22,000)
% occurring within the 48 cities			30%	
Jerrett et al. (2009), April-September	Long-term	≥30 years	-	45,000
95% confidence interval				(17,000-70,000)

[1] City-specific effect estimates are applied to the gridcells lying within the cities defined in the epidemiological studies. Average effect estimates across all cities included in the epidemiological studies (national average) are applied to all other gridcells. For the application of Smith et al. (2009) effect estimates, city-specific effect estimates were applied to 2,227 gridcells and the national average to 44,064 gridcells. For the application of Zanobetti and Schwartz (2008) effect estimates, city-specific effect estimates were applied to 925 gridcells and the national average to 45,366 gridcells.

[2] National average effect estimates are based on the average of all cities included in the epidemiological studies applied to all 12km gridcells nationally.

Table 8.1 also shows O_3-related deaths estimated by applying the national average risk estimate from the epidemiological studies to all gridcells in the U.S. Compared with applying city-specific effect estimates to the gridcells corresponding to each urban area, using the national average effect estimate for all gridcells yields equivalent central estimates. However, applying the national average also results in tighter confidence intervals since the national average effect estimates had higher statistical power and thus tighter confidence bounds compared with the effect estimates for individual cities.

Table 8.2 shows the mean, median, 2.5 percentile and 97.5 percentile of the estimated percentage of mortality attributable to ambient O_3 across all counties in the U.S. Using Smith et al. (2009) effect estimates, O_3-attributable mortality contributes an average of 1.5% (95% confidence interval, 1.1%-1.8%) to county-level May-September non-accidental mortality (all ages) and 0.6% (0.4%-0.7%) to all year all-cause mortality (all ages). For results using Zanobetti

and Schwartz (2008) effect estimates, O_3-attributable mortality contributes an average of 2.5% (95% confidence interval, 1.7%-3.0%) to county-level June-August all-cause mortality (all ages) and 0.6% (0.4%-0.8%) to all year all-cause mortality (all ages). For the results using Jerrett et al. (2009) effect estimates, O_3-attributable mortality contributes an average of 18.5% (95% confidence interval, 15.2%-21.5%) to county-level April-September adult (age 30+) respiratory mortality and 1.9% (1.3%-2.6%) to all year all-cause mortality (all ages). Figure 8.5 through Figure 8.7 show that the counties with the highest percentage of mortality attributable to O_3 are typically those with the highest O_3 levels.

Figure 8.8 displays the cumulative distribution of the percent of county-level all-cause, all-age, and all-year mortality attributable to ambient O_3 using effect estimates from all three epidemiological studies. For the results based on Smith et al. (2009) and Zanobetti and Schwartz (2008) effect estimates, 0.8% of all-cause, all-age, and all-year mortality is attributable to O_3 for approximately 99% of U.S. counties. For the results based on Jerrett et al. (2009) effect estimates, 2.8% of all-cause, all-age, and all-year mortality is attributable to O_3 for approximately 99% of U.S. counties.

Table 8-2 Mean, median, 2.5 percentile, and 97.5 percentile of the estimated percentage of mortality attributable to ambient O_3 for all 3087 counties in the continental U.S.[1]

Source of risk estimate, modeling period, and mortality endpoint used to generate percentage	Total incidence (2005)	Percentage of total incidence attributable to O_3			
		Mean (%)	Median (%)	2.5 Percentile (%)	97.5 Percentile (%)
Smith et al. (2009), May-September					
Non-accidental mortality, all ages	964,837	1.5	1.5	1.1	1.8
All-cause mortality, all ages	1,028,334	1.4	1.4	1.0	1.7
All-cause mortality, all ages, all year	2,454,896	0.6	0.6	0.4	0.7
Zanobetti and Schwartz (2008), June-August					
All-cause mortality, all ages	618,345	2.5	2.5	1.7	3.0
All-cause mortality, all ages, all year	2,454,896	0.6	0.6	0.4	0.8
Jerrett et al. (2009), April-September					
Respiratory mortality, ages 30+	236,756	18.5	18.7	15.2	21.5
All-cause mortality, all ages	1,229,968	3.8	3.7	2.6	5.2
All-cause mortality, all ages, all year	2,454,896	1.9	1.9	1.3	2.6

[1] For the mortality endpoints matching the epidemiology studies as a percentage of incidence of the same endpoint for the same seasonal definition, and as a percentage of all-cause mortality for all age groups (both seasonal and all year).

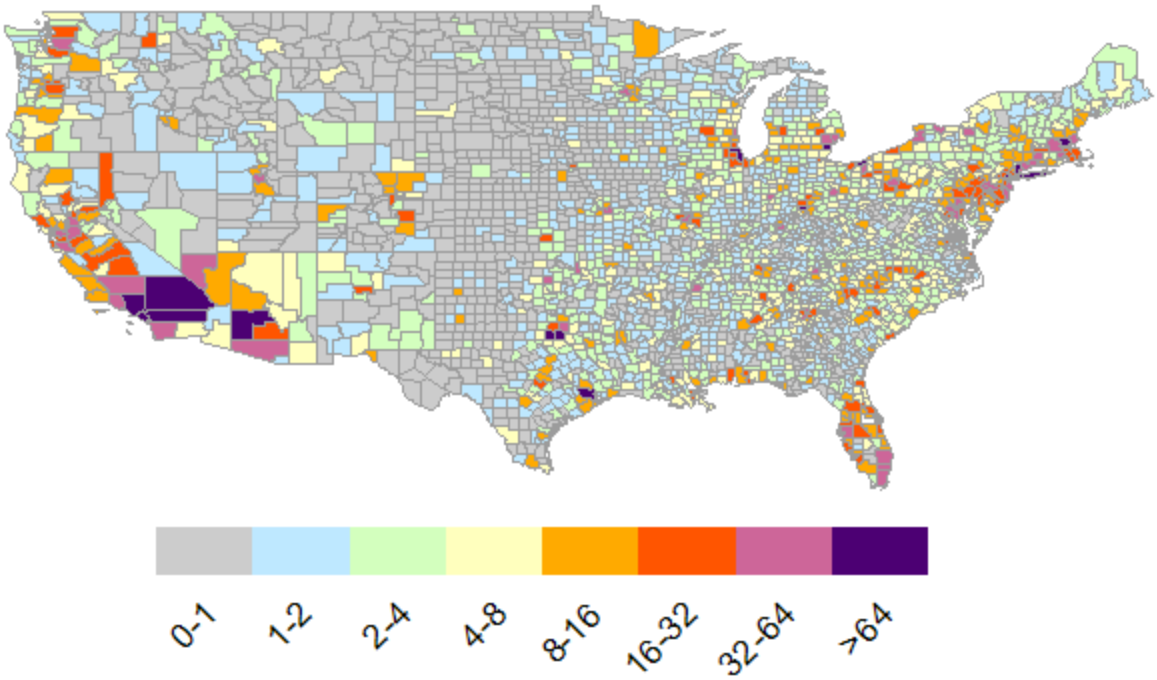


Figure 8.2 Estimated annual non-accidental premature deaths (individuals) in 2007 associated with average 2006-2008 May-September average 8-hr daily maximum O_3 levels by county using Smith et al. (2009) effect estimates

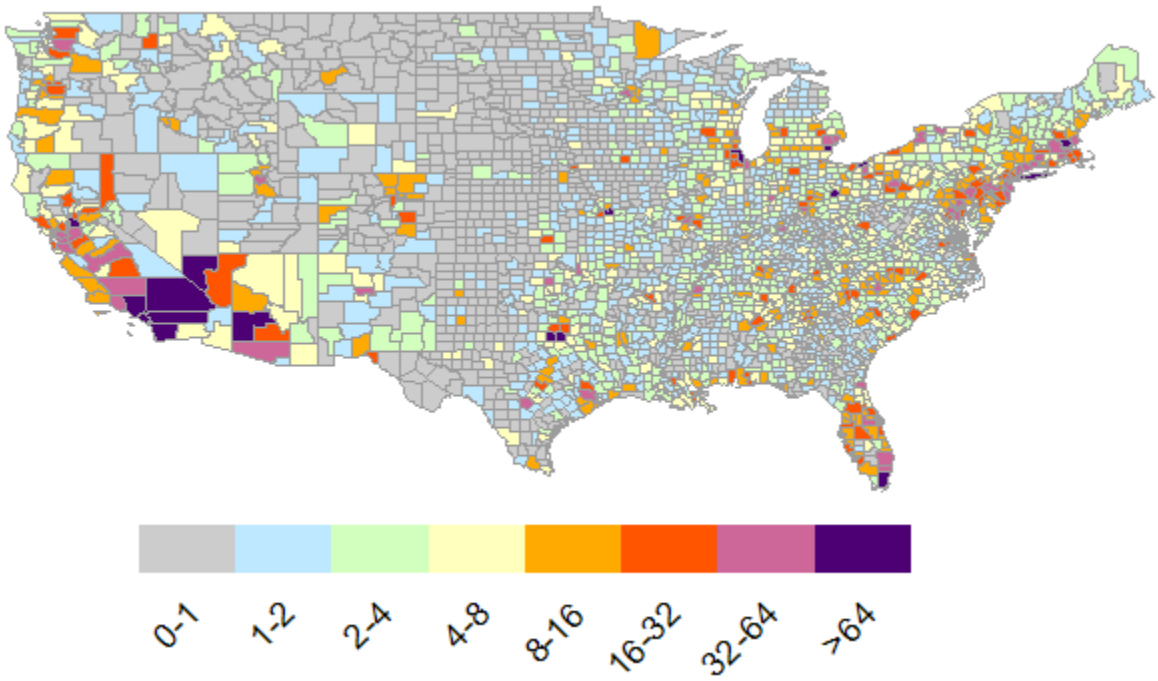


Figure 8.3 Estimated annual all-cause premature deaths (individuals) in 2007 associated with average 2006-2008 June-August average 8-hr daily mean (10am-6pm) O_3 levels by county using Zanobetti and Schwartz (2008) effect estimates

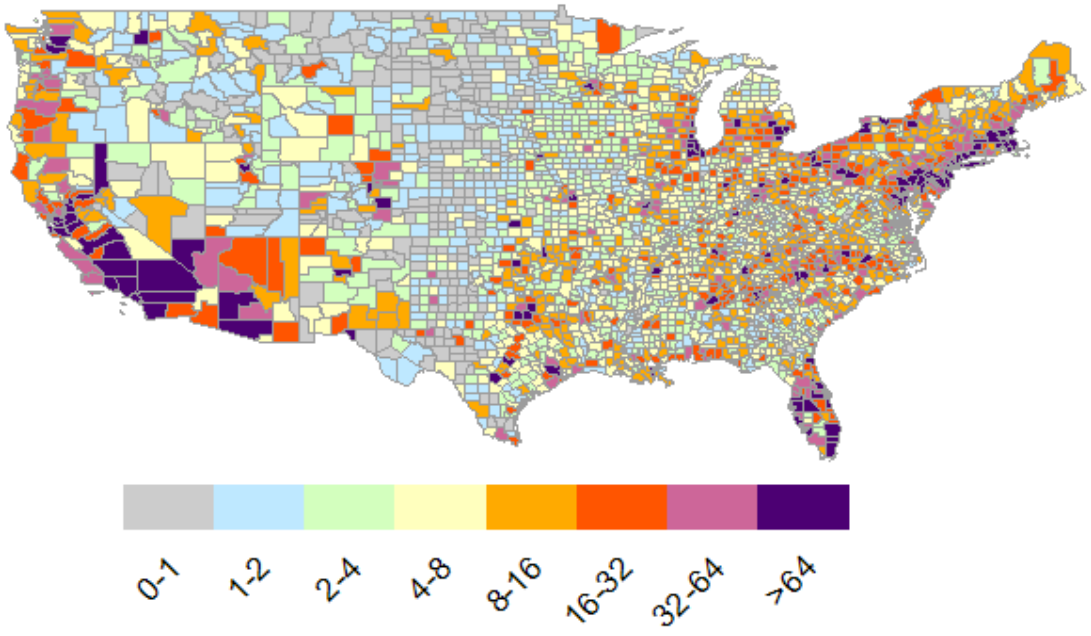


Figure 8.4 Estimated annual adult (age 30+) respiratory premature deaths (individuals) in 2007 associated with average 2006-2008 April-September average 1-hr daily max O_3 levels by county using Jerrett et al. (2009) effect estimates

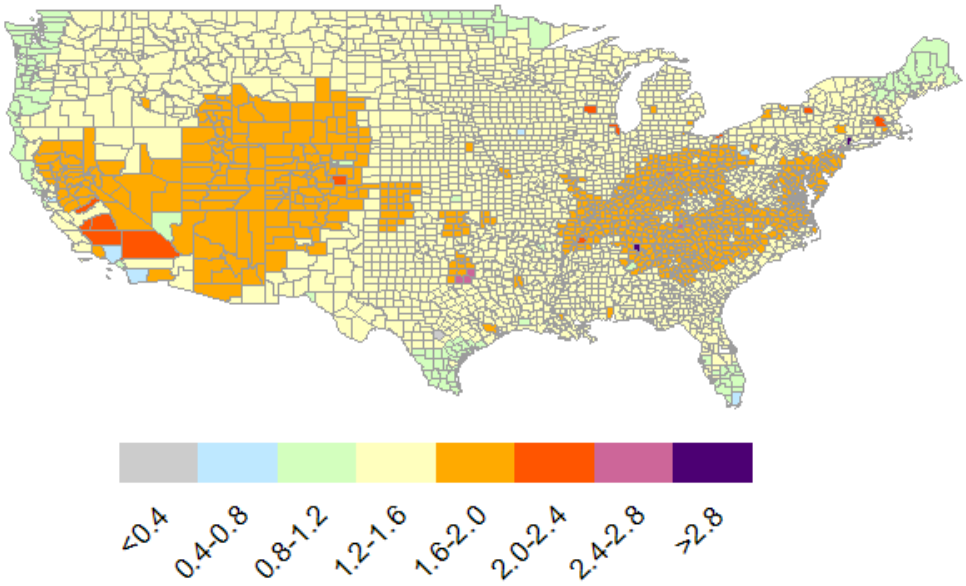


Figure 8.5 Estimated percentage of May-September total non-accidental mortality (all ages) attributable to 2006-2008 average O_3 levels by county using Smith et al. (2009) effect estimates

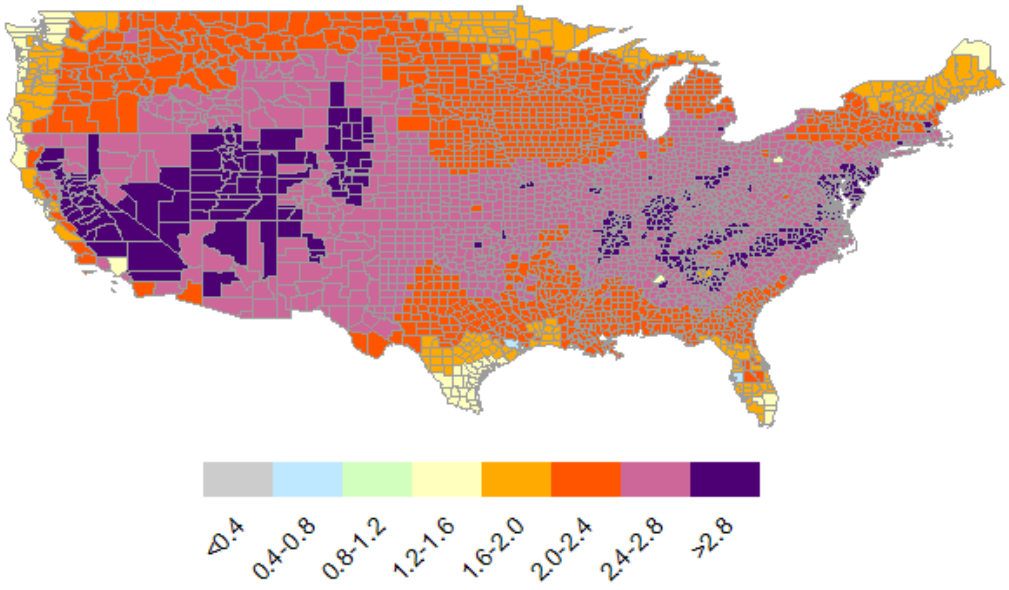


Figure 8.6 Estimated percentage of June-August total all-cause mortality (all ages) attributable to 2006-2008 average O_3 levels by county using Zanobetti and Schwartz (2008) effect estimates

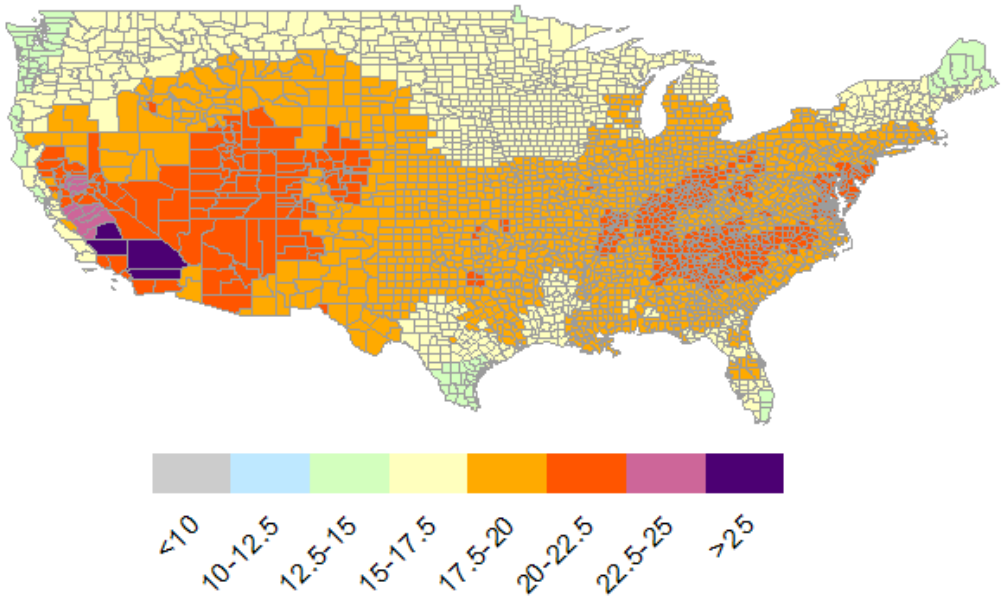


Figure 8.7 Estimated percentage of April-September respiratory mortality among adults age 30+ attributable to 2006-2008 average O_3 levels by county using Jerrett et al. (2009) effect estimates

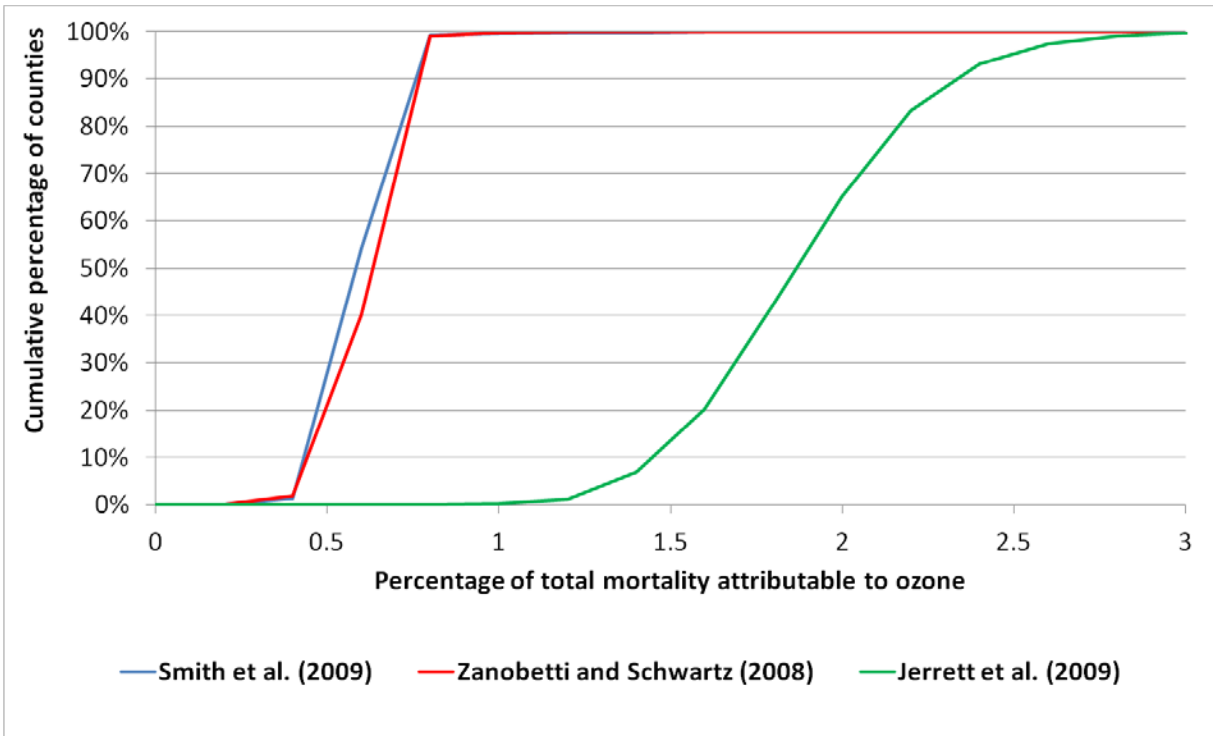


Figure 8.8 Cumulative distribution of county-level percentage of all-cause, all-year, and all-age mortality attributable to 2006-2008 average O_3 for the U.S.[2]

[2] Estimated O_3-attributable deaths are based on the mortality cause, age group, and season inherent to the epidemiological study upon which it is based (May-September non-accidental mortality for all ages for results based on Smith et al. (2009) effect estimates, June-August all-cause mortality for all ages for results based on Zanobetti and Schwartz (2008) effect estimates, and April-September respiratory mortality for ages 30+ for results based on Jerrett et al. (2009) effect estimates).

Figure 8.9 shows the cumulative distribution of the county-level percent of total O_3-related deaths by O_3 concentration. The mortality results based on Smith et al. (2009) concentration-response functions are compared with the May-September average of the 8-hr daily maximum O_3 concentration, those based on Zanobetti and Schwartz (2008) concentration-response functions are compared with the June-August average of the 8-hr mean O_3 concentration from 10am to 6pm, and those based on Jerrett et al. (2009) concentration-response functions are compared with the April-September average of the 1-hr daily maximum O_3 concentration, consistent with the O_3 concentration metrics used in each study. The mortality results based on Zanobetti and Schwartz (2008) effect estimates are shifted to the right of the mortality results based on the Smith et al. (2009) concentration response functions because the seasonal averaging time for the results based on Zanobetti and Schwartz (2008) is limited to the summer months when O_3 tends to be highest. Similarly, the mortality results based on Jerrett et al. (2009) effect estimates are shifted to the right of the mortality results based on Zanobetti and Schwartz (2008) and Smith et al. (2009) because Jerrett et al. (2009) results use the seasonal average of the 1-hr daily maximum, which tends to be higher than the seasonal average of 8-hr daily maximum and seasonal average of 8-hr daily mean metrics (see Figure 4-18). For all three epidemiology studies, we find that 90-95% of O_3-related deaths occur in locations where the May to September average 8-hr daily maximum, June to August average 8-hr daily mean (10am-6pm), or April to September average 1-hr daily maximum O_3 concentrations are greater than 40 ppb. A seasonal average concentration of 40 ppb corresponds to 2006-2008 design values ranging from approximately 50 to 90 ppb, depending on the seasonal average concentration metric (see Figure 4-19).

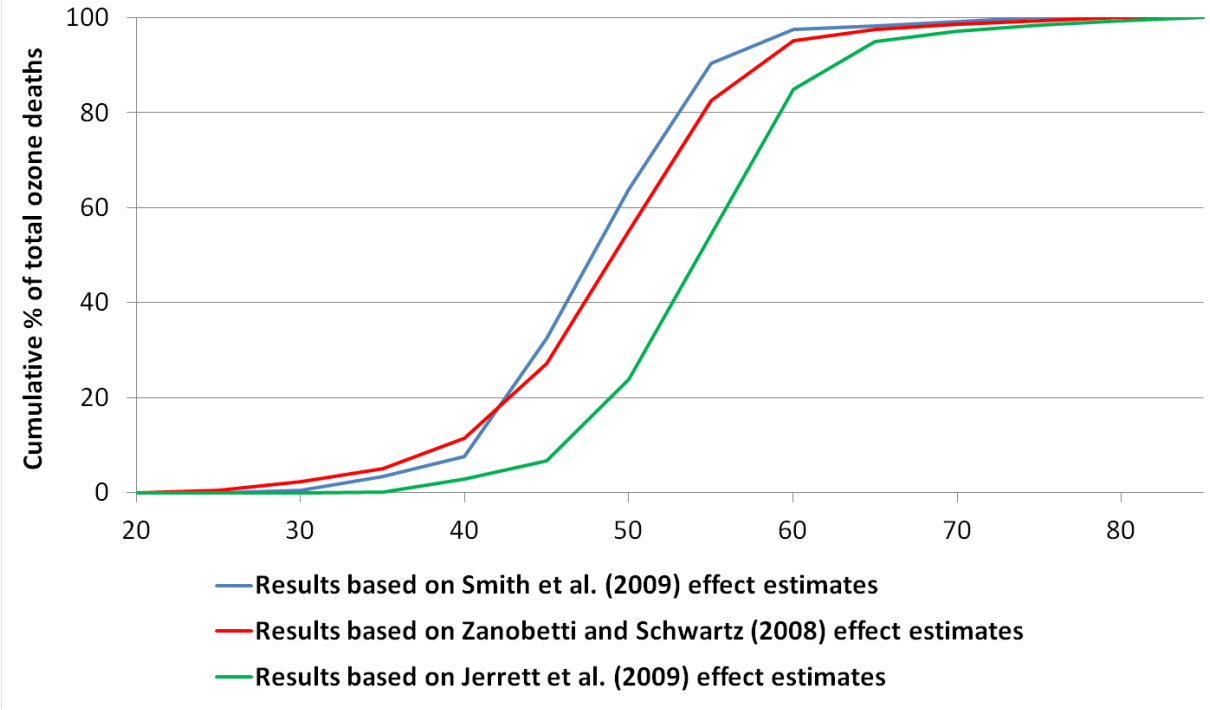


Figure 8.9 Cumulative percentage of total O_3 deaths by baseline O_3 concentration. O_3 concentrations are reported as May-September average 8-hr daily maximum for results based on Smith et al. (2009) effect estimates, June-August average 8-hr mean (10am to 6pm) for results based on Zanobetti and Schwartz (2008) effect estimates, and April-September average 1-hr daily maximum for results based on Jerrett et al. (2009) effect estimates.

8.1.3 Sensitivity analysis

For the results presented above, the national average effect estimate for results based on Smith et al. (2009) and Zanobetti and Schwartz (2008) was applied to all gridcells between the cities included in the studies. However, O_3-mortality effect estimates have been shown to exhibit significant regional variability across the U.S. (e.g. Smith et al. 2009). Smith et al. (2009) found that using the national average effect estimate may overestimate risk in cities that have low effect estimates, including Los Angeles and Denver, but may underestimate risk in cities that have high effect estimates, including New York City and Chicago. We conduct two sensitivity analyses aimed at characterizing the sensitivity of estimated O_3-attributable premature deaths to the use of national average effect estimates between the cities that were included by Smith et al. (2009) and Zanobetti and Schwartz (2008).

First, we examine the sensitivity of estimated O_3-attributable premature deaths to the application of the 5th highest and 5th lowest effect estimates of all the cities included in the Smith et al. (2009) and Zanobetti and Schwartz (2008) studies to the gridcells between the cities

included in these studies (Table 8.3). As in the main results, city-specific effect estimates are applied to the gridcells in which the cities lie. Applying the 5th highest effect estimate from Las Vegas to the gridcells between the cities included by Smith et al. (2009) yields a 36% lower estimate of O_3-attributable deaths as compared with the main results. Applying the 5th lowest effect estimate from Dallas/Ft. Worth yields a 42% higher estimate of O_3-attributable deaths as compared with the main results. Applying the 5th lowest effect estimate from Los Angeles to the gridcells between the cities included by Zanobetti and Schwartz (2008) yields a 37% lower estimate of O_3-attributable deaths as compared with the main results. Applying the 5th highest effect estimate from Columbus, OH, yields a 30% higher estimate of O_3-attributable deaths as compared with the main results.

Table - 8.3 Sensitivity of estimated O_3-attributable premature deaths to the application of the 5th lowest and 5th highest city-specific risk estimates found by Smith et al. (2009) and Zanobetti and Schwartz (2008) to the gridcells between the cities included in those studies.

Source of risk estimate and sensitivity study	Beta and city	O_3-attributable mortality	Percent change from main results
Smith et al. (2009), May-September			
5th lowest city beta	0.00014 Las Vegas, NV	9,600 (-20,000 – 38,000)	-36%
5th highest city beta	0.000538 Dallas/Ft. Worth, TX	21,000 (-1000 – 43,000)	+42%
Zanobetti and Schwartz (2008), June-August			
5th lowest city beta	0.000274 Los Angeles, CA	9,800 (-3,700 – 23,000)	-37%
5th highest city beta	0.000739 Columbus, OH	20,000 (100 – 40,000)	+30%

Second, we examine the sensitivity of estimated O_3-attributable premature deaths to the application of Smith et al. (2009) Bayesian-shrunken city-specific estimates using regional average priors rather than the national average prior (Table 8.4). For gridcells between the cities included by Smith et al. (2009), we apply the regional average effect estimate, rather than the national average effect estimate as in the main results. Regional definitions are shown in Figure

8.10. Estimated O_3-attributable deaths using the regional prior city-specific effect estimates and the regional average effect estimates between the 98 cities included by Smith et al. (2009) are approximately 20% larger than the main results, with 38% of estimated deaths occurring in the 98 cities rather than 43%. The 95% confidence interval for the results using the regional prior spans zero, whereas the 95% confidence interval for the results using the national prior does not. Since the regional average effect estimates are all based on fewer data points (in some regions, the regional average is based on only seven cities; see Appendix 8-A) than is the national average, the confidence interval for each regional average effect estimate is large and sometimes spans zero. The large confidence intervals for the regional average effect estimates drive the confidence interval that spans zero for O_3-attributable mortality estimated using regional prior effect estimates. Confidence intervals that span zero do not imply that higher O_3 is associated with decreased mortality, as there is no biologically plausible mechanism for such an effect, and in no case do we see a significant negative central estimate. Rather, confidence intervals spanning zero indicate a lack of statistical power to precisely determine the magnitude of an effect.

Figure 8.11 shows estimated O_3-attributable deaths by region using the national average prior compared with using the regional average priors from Smith et al. (2009). Results generally follow conclusions made by Smith et al. (2009) based on the magnitude of the regional effect estimates. For example, using the national average effect estimate may substantially underestimate O_3-attributable deaths in the North East and Industrial Midwest where regional effect estimates are large. Using the national average effect estimate may also overestimate O_3-attributable deaths in the Upper Midwest, Southern California, and South West, which were found to have small effect estimates. However, these three regions have very large confidence intervals which all span zero, since these regional averages are based on few cities (7, 7, and 9, respectively, compared with 26 in the South East, 19 in Industrial Midwest, 16 in North East, and 12 in North West; see Appendix 8-A).

Table - 8.4 Sensitivity of estimated O_3-attributable premature deaths to the application of Smith et al. (2009) regional prior Bayes-shrunken city-specific and regional average effect estimates, as compared with the national prior Bayes-shrunken city-specific and national average effect estimates as in the main results.

Risk estimate	O_3-attributable premature deaths	Percent O_3-attributable deaths in 98 cities
City-specific, national prior with national average	15,000 (1,400 – 28,300)	43%
City-specific, regional prior with regional averages	18,000 (-2,000 – 24,000)	38%

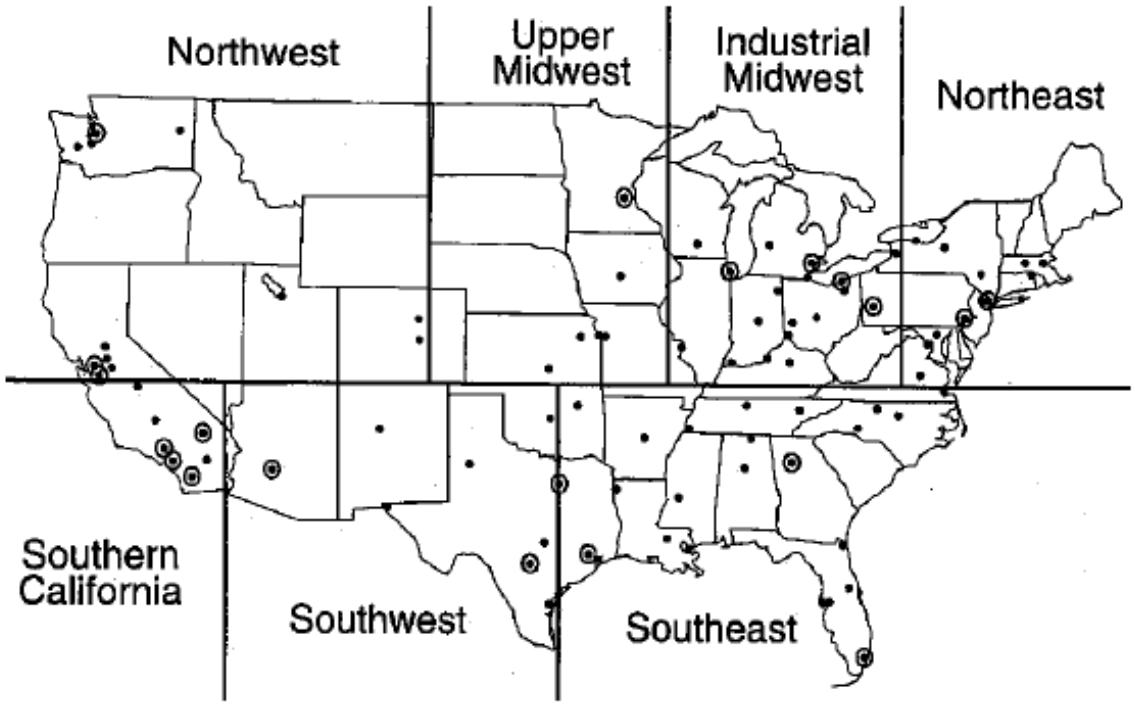


Figure 8.10 Regions used in the sensitivity analysis based on the Smith et al. (2009) regional-prior Bayes-shrunken city-specific and regional average effect estimates (Source: Samet et al. 2000).

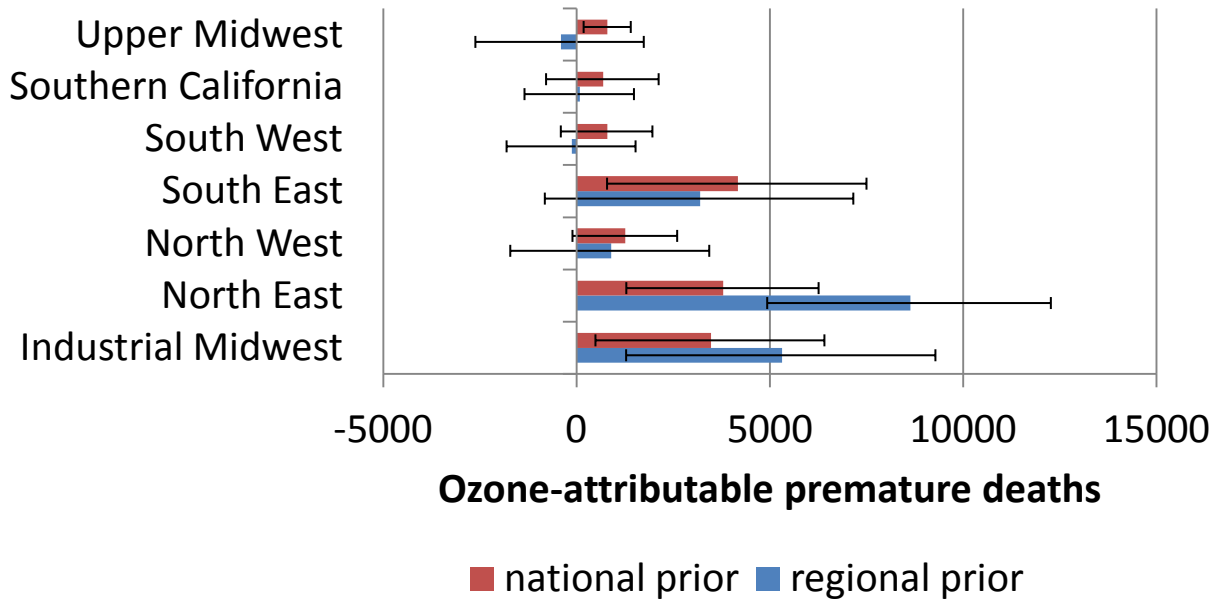


Figure 8.11 O_3-attributable premature deaths by region as calculated by applying Smith et al. (2009) regional prior Bayes-shrunken and regional average effect estimates, as compared with the national prior Bayes-shrunken and national average effect estimates as in the main results

8.1.4 Discussion

We estimated the total all-cause deaths associated with short-term exposure to recent O_3 levels across the continental U.S., using average 2006-2008 observations from the O_3 monitoring network fused with a 2007 CMAQ simulation and city-specific O_3-mortality effect estimates from two short-term epidemiology studies. Applying Smith et al. (2009) effect estimates for May-September, we estimate 15,000 (95% CI, 1,400-28,000) premature O_3-related non-accidental deaths (all ages) annually for 2007. Using Smith et al. (2009) effect estimates, O_3-attributable mortality contributes an average of 1.5% (95% confidence interval, 1.1%-1.8%) to county-level May-September non-accidental mortality (all ages) and 0.6% (0.4%-0.7%) to all year all-cause mortality (all ages). As a sensitivity, we apply Zanobetti and Schwartz (2008) effect estimates for June-August, finding 16,000 (95% CI, 6,000-25,000) premature O_3-related all-cause deaths (all ages) annually for 2007. For results using Zanobetti and Schwartz (2008) effect estimates, O_3-attributable mortality contributes an average of 2.5% (95% confidence interval, 1.7%-3.0%) to county-level June-August all-cause mortality (all ages) and 0.6% (0.4%-0.8%) to all year all-cause mortality (all ages). For the application of Jerrett et al. (2009) effect estimates for April-September, we estimate 45,000 (95% CI, 17,000-70,000) premature O_3-related adult (age 30 and older) respiratory deaths. For the results using Jerrett et al. (2009) effect

estimates, O_3-attributable mortality contributes an average of 18.5% (95% confidence interval, 15.2%-21.5%) to county-level April-September adult (age 30+) respiratory mortality and 1.9% (1.3%-2.6%) to all year all-cause mortality (all ages). For all three epidemiology studies, we find that 90-95% of O_3-related deaths occur in locations where the May to September average 8-hr daily maximum, June to August average 8-hr daily mean (10am-6pm), or April to September average 1-hr daily maximum O_3 concentrations are greater than 40 ppb. A seasonal average concentration of 40 ppb corresponds to 2006-2008 design values ranging from approximately 50 to 90 ppb, depending on the seasonal average concentration metric.

A previous analysis estimated that short-term O_3 exposure was associated with 4,700 (95% CI, 1,800-7,500) premature deaths nationwide annually, based on 2005 O_3 concentrations and Bell et al. (2004) national average effect estimates (Fann et al., 2012). The results estimated here are higher, resulting mainly from two important differences. First, Fann et al. (2012) estimated risk only above North American background, simulated O_3 concentrations in the absence of North American anthropogenic emissions, which was set to 22 ppb in the east and 30 ppb in the west. Fann et al. (2012) also used a national average mortality effect estimate for 8-hr daily maximum O_3 during the warm season only, calculated using ratios of 24-hr mean concentrations to 8-hr daily maximum concentrations (see Abt Associates 2010). The Smith et al. (2009) national average beta used here, 0.000322, is based on April-October O_3 data and is approximately 23% larger than that used by Fann et al. (2012), 0.000261. Since the risk modeling period (and the seasonal definition for the seasonal average 8-hr daily maximum concentration) was May to September for both studies, the higher beta used here yields a larger O_3 mortality estimate. These two differences in methods explain the larger O_3 mortality estimates of this analysis compared with the previous estimate by Fann et al. (2012).

Estimated O_3-attributable premature deaths based on Jerrett et al. (2009) effect estimates are approximately three times larger than results based on Smith et al. (2009) and Zanobetti and Schwartz (2008) effect estimates. The mean estimated county-level percent of all-cause, all-year, and all-age mortality is also three times larger for results based on Jerrett et al. (2009) effect estimates, indicating that the larger estimate does not simply result from a longer modeling period or different population subset (e.g. adult respiratory disease for Jerrett et al. (2009) effect estimates versus all-age non-accidental or all-cause mortality for Smith et al. (2009) and Zanobetti and Schwartz (2008) effect estimates). Recent studies using long-term O_3-mortality relationships found by Jerrett et al. (2009) to quantify the burden of mortality due to anthropogenic O_3 globally (Anenberg et al. 2010, 2011) and for the U.S. specifically (Fann et al. 2012) have also found that using Jerrett et al. (2009) long-term effect estimates yields O_3-related mortality burden estimates that are approximately two to four times larger than estimates based on short-term effect estimates. Since long-term mortality relationships include both acute and

chronic exposure effects, the significantly larger mortality estimates calculated using long-term concentration-mortality relationships suggest that considering only short-term mortality may exclude a substantial portion of O_3-related risk. However, since the short-term mortality relationships include a larger population (all ages versus adults ages 30 and older only) and all mortality causes, the short-term mortality relationships may capture some O_3 effects that are not captured by Jerrett et al. (2009). It is likely that some portion of the estimated premature deaths attributable to short-term O_3 exposure is captured by estimated premature deaths attributable to long-term O_3 exposure, but the extent of the overlap between these estimates is unknown.

8.2 EVALUATING THE REPRESENTATIVENESS OF THE URBAN STUDY AREAS IN THE NATIONAL CONTEXT

To further support interpretation of risk estimates generated in Section 7.2, this section presents three analyses that assess the representativeness of the 12 urban study areas in the national context. First, we assess the degree to which the urban study areas represent the range of air quality levels and key O_3 risk-related attributes that vary spatially across the nation. We have partially addressed this issue by selecting urban study areas in different geographical regions of the country (see Section 7.2). In this section, we evaluate how well the selected urban areas represent the overall U.S. for a set of spatially-distributed O_3 risk related variables (e.g. weather, demographics including socioeconomic status, baseline health incidence rates; Section 8.2.1). Section 8.2.2 identifies where our 12 urban study areas fall along the distribution of O_3-attributable mortality risk across the U.S. This analysis allows us to assess the degree to which the 12 urban study areas capture locations within the U.S. likely to experience elevated levels of risk related to ambient O_3. Finally, we give a national context to the estimated O_3 responses to emission changes in the urban study areas by assessing how well these 12 areas and the 3 additional exposure areas represent air quality trends and responses to emissions across the entire U.S. (Section 8.2.3).

We do not attempt to assess the representativeness of the 15 urban study areas considered in the exposure assessment for O_3 related risk because data limitations preclude us from being able to characterize individual-level exposure across the U.S. However, the urban study areas considered in both the exposure and risk assessments shared common selection criteria, including consideration of O_3 concentrations, availability of adequate monitoring data, demographics, and exposure factors. Therefore, conclusions from this analysis of the representativeness of the 12 urban study areas for risk would also apply to those areas for exposure.

8.2.1 Analysis Based on Consideration of National Distributions of Risk-Related Attributes

This section evaluates how well the urban study areas reflect national-level variability in a series of O_3 risk-related variables. For this analysis, we first generate distributions for risk-related variables across the U.S. and for the specific urban study areas considered in Section 7.2 from generally available data (e.g. from the 2000 Census, Centers for Disease Control (CDC), or other sources). We then plot the specific values of these variables for the selected urban study areas on these distributions, and evaluate how representative the selected study areas are of the national distributions for these individual variables.

Estimates of risk (either relative or absolute, e.g. number of cases) within our risk assessment framework are based on four elements: population, baseline incidence rates, air quality, and the coefficient relating air quality and the health outcome (i.e. the O_3 effect estimates). Each of these elements can contribute to heterogeneity in risk across urban locations, and each is variable across locations. In addition, there may be other identifiable factors that contribute to the variability of the four elements across locations. In this assessment, we examine the representativeness of the selected urban area locations for the four main elements, as well as factors that have been identified as influential in determining the magnitude of the C-R function across locations.

While personal exposure is not incorporated directly into O_3 epidemiology studies, city-specific O_3 effect estimates are affected by differing levels of exposure which in turn are related to variability in exposure determinants. The correlation between monitored O_3 and personal O_3 exposure also varies between cities. The O_3 ISA has comprehensively reviewed epidemiological and toxicological studies to identify variables which may affect the O_3 effect estimates used in the city-specific risk analysis in Section 7.2 and the national-scale risk analysis in Section 8.1 (U.S. EPA 2013). Determinants of the O_3 effect estimates used in risk assessment can be grouped into four broad areas:

- Demographics: education, income, age, unemployment rates, race, body mass index and physical conditioning, public transportation use, and time spent outdoors.
- Baseline health conditions: asthma, chronic obstructive pulmonary disease, cardiovascular disease (atherosclerosis, congestive heart disease, atrial fibrillation, stroke), diabetes, inflammatory diseases, and smoking prevalence.
- Climate and air quality: O_3 levels, co-pollutant levels (annual mean $PM_{2.5}$), temperatures (days above 90 degrees, mean summer temp, 98th percentile temp).
- Exposure determinants: air conditioning prevalence.

Although data limitations preclude our ability to conduct a national-scale exposure assessment as we have done for O_3-attributable risk in Section 8.1, we assess the representativeness of the

urban study areas across the national distribution of climate, air quality, and air conditioning prevalence, factors which influence individual exposure. As discussed in detail in Chapter 5, no available data base is sufficient to assess the national representativeness of time spent outdoors, another important personal exposure determinant, among persons residing in each of the urban case study areas. However, previous analyses suggest that children's time spent outdoors varies little across U.S. regions (section 8.10.2 of U.S. EPA, 2009). In addition, as discussed in Section 5.1.1, time spent outdoors and the percent of person-days having at least one minute outdoors (participation rate) does not appear to vary much over the past few decades based on analyses using the CHAD database, nor does there appear to be a temporal trend over the past decade based on analyses using the American Time Use Survey (ATUS). In considering that many of the activity pattern studies in CHAD were from national surveys conducted in metropolitan areas and that the evaluation results indicate little difference in time expenditure over broad geographic areas and survey collection years, it is likely that the distribution of time spent outdoors generated for the simulated persons in the 15 urban study areas (Chapter 5) reasonably reflects the most important elements of a national distribution of time spent outdoors.

Based on these identified potential risk determinants, we identify datasets that could be used to generate nationally representative distributions for each parameter. We are not able to identify readily available national datasets for all variables. In these cases, if we are able to identify a broad enough dataset covering a large enough portion of the U.S., we use that dataset to generate the parameter distribution. In addition, we are not able to find exact matches for all of the variables identified through our review of the literature. In cases where an exact match is not available, we identify proxy variables to serve as surrogates. For each parameter, we report the source of the dataset, its degree of coverage, and whether it is a direct measure of the parameter or a proxy measure (Table 8.5). Summary statistics for the most relevant variables are provided in Table 8.6.

Figure 8.12 through Figure 8.18 show the cumulative distribution functions (CDF) plotted for the nation for the four critical risk function elements (population, air quality, baseline incidence, and the O_3 effect estimate), as well as where the urban study areas fall on the distribution. While the urban-scale analysis in Chapter 7 includes the full core-based statistical area for the selected cities, we consider here only the counties included in each city as defined by the epidemiological studies, since we only have information on O_3 effect estimates for these counties. This approach is consistent with the national-scale assessment of O_3-attributable risk in Section 8.1, from which we draw county-level O_3-attributable risk estimates for the representativeness analysis in Section 8.2.3. These figures focus on critical variables representing each type of risk determinant, e.g. we focus on all-cause and non-accidental mortality rates, but we also have conducted analyses for cardiovascular and respiratory mortality separately. The

vertical black lines in each graph show the values of the variables for the individual urban study areas. The city-specific values that comprise the national CDF for mortality risks found by Zanobetti and Schwartz (2008) are also displayed on the graphs of those attributes, as the number of cities included in that study is smaller (48 cities). The complete set of analyses is provided in Appendix 4-A.

These figures show that the selected urban study areas represent the upper percentiles of the distributions of population and do not represent the locations with lower populations (urban study areas are all above the 90th percentile of U.S. county populations). This is consistent with the objectives of our case study selection process, e.g. we are characterizing risk in areas that are likely to be experiencing excess risk due to O_3 levels above alternative standards. The urban study areas span the full range of seasonal average 8-hr daily maximum O_3 concentrations in monitored U.S. counties and the full distribution of O_3 risk coefficients across the cities included by Smith et al. (2009) and Zanobetti and Schwartz (2008). The urban study area analysis includes the two cities with the highest risk coefficients found by Smith et al. (2009) – New York City and Philadelphia – as well as the two highest found by Zanobetti and Schwartz (2008) – New York City and Detroit. In Table 8.6, respiratory and cardiovascular mortality have higher concentration-response relationships than non-accidental and all-cause mortality because they are based on a smaller baseline population and are the diseases most affected by O_3 exposure. The urban study areas do not capture the upper end of the distribution of baseline mortality, including all-cause (Figure 8.15) and non-accidental mortality (Figure 8.16), as well as cardiovascular and respiratory mortality (see Appendix 8-B). The interpretation of this is that the case study risk estimates may not capture the additional risk that may exist in locations that have the highest baseline mortality rates.

Table - 8.5 Data Sources for O_3 risk-related Attributes

Potential risk determinant	Metric	Year	Source	Degree of national coverage
Demographics				
Age	Percent age 85 years and older	2005	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Age	Percent age 65 years and older	2005	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Age	Percent age 14 years and younger	2005	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Education	Population with less than high school diploma	2000	USDA/ERS, http://www.ers.usda.gov/Data/Education/	All counties
Unemployment	Percent unemployed	2005	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Income	Per capita personal income	2005	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Race	Percent nonwhite	2006	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties

Potential risk determinant	Metric	Year	Source	Degree of national coverage
Population	Total population	2008	Cumulative Estimates of Resident Population Change for the United States, States, Counties, Puerto Rico, and Puerto Rico Municipios: April 1, 2000 to July 1, 2008, Source: Population Division, U.S. Census Bureau	All counties
Population density	Population/square mile	2008	Cumulative Estimates of Resident Population Change for the United States, States, Counties, Puerto Rico, and Puerto Rico Municipios: April 1, 2000 to July 1, 2008, Source: Population Division, U.S. Census Bureau (calculated "as the crow flies")	All counties
Urbanicity	ERS Classification Code	2003	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Climate and Air Quality				
O_3 levels	Monitored 4th high 8-hr daily maximum	2007	EPA Air Quality System (AQS)	725 Monitored counties
O_3 levels	Seasonal mean 8-hr daily maximum	Avg. 2006-2008	AQS	671 Monitored counties
O_3 levels	Seasonal mean 1-hr daily	Avg. 2006-2008	AQS	671 Monitored

Potential risk determinant	Metric	Year	Source	Degree of national coverage
	maximum			counties
O_3 levels	Seasonal mean	Avg. 2006-2008	AQS	671 Monitored counties
$PM_{2.5}$ levels	Monitored annual mean	2007	AQS	617 Monitored counties
Temperature	Mean July temp	1941-1970	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Relative Humidity	Mean July RH	1941-1970	County Characteristics, 2000-2007 Inter-university Consortium for Political and Social Research	All counties
Exposure Determinants				
Ventilation	Percent residences with no air conditioning	2004	American Housing Survey	76 cities
Baseline Health Conditions				
Baseline mortality	All Cause		CDC Wonder 1999-2005	All counties
Baseline mortality	Non Accidental		CDC Wonder 1999-2006	All counties
Baseline mortality	Cardiovascular		CDC Wonder 1999-2007	All counties
Baseline mortality	Respiratory		CDC Wonder 1999-2008	All counties
Baseline morbidity	Acute myocardial infarction prevalence	2007	Behavioral Risk Factor Surveillance System (BRFSS)	184 metropolitan statistical areas

Potential risk determinant	Metric	Year	Source	Degree of national coverage
				(MSA)
Baseline morbidity	Diabetes prevalence	2007	BRFSS	184 MSA
Baseline morbidity	Stroke prevalence	2007	BRFSS	184 MSA
Baseline morbidity	Congestive heart disease prevalence	2007	BRFSS	184 MSA
Obesity	Body Mass Index	2007	BRFSS	184 MSA
Level of exercise	Vigorous activity 20 minutes	2007	BRFSS	184 MSA
Level of exercise	Moderate activity 30 minutes or vigorous activity 20 minutes	2007	BRFSS	184 MSA
Respiratory risk factors	Current asthma	2007	BRFSS	184 MSA
Smoking	Ever smoked	2007	BRFSS	184 MSA
C-R Estimates				
Mortality risk	Non Accidental	2009	Smith et al. (2009)	98 cities
Mortality risk	All Cause	2008	Zanobetti and Schwartz (2008)	48 cities
Mortality risk	Cardiovascular	2008	Zanobetti and Schwartz (2008)	48 cities
Mortality risk	Respiratory	2008	Zanobetti and Schwartz (2008)	48 cities

Table - 8.6 Summary Statistics for Selected O_3 Risk-related Attributes

Risk Attribute	Average		Standard Deviation		Maximum		Minimum		Sample Size (# of counties or cities)	
	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset
Demographics										
Population per county	1,642,198	97,020	1,972,403	312,348	9,862,049	9,862,049	354,361	42	23	3143
Population density (Pop/sq mile)	10,378	258	16,550	1,757	71,758	71,758	1,313	0	23	3143
Median age (Years)	35.7	38.6	2.3	4.4	40.0	55.3	32.1	20.1	23	3141
% Age 0 to 14 years	20.7	19.0	2.4	2.9	24.6	36.8	14.7	0.0	23	3141
% Age 65+ years	11.3	14.9	2.5	4.1	15.2	34.7	5.8	2.3	23	3141
% Age 85+ years	1.7	2.1	0.6	0.9	2.5	7.7	0.5	0.1	23	3141
Unemployment rate (%)	5.7	5.4	1.2	1.8	8.6	20.9	4.1	1.9	23	3133
% with less than high school diploma	20.9	22.6	7.9	8.8	37.7	65.3	8.7	3.0	23	3141
Income per capita ($)	40305	27367	14238	6604	93377	93377	23513	5148	23	3086
% Non-white	36.4	13.0	15.3	16.2	86.7	95.3	31.7	0.0	23	3141
% Commute by public transportation*	7.1	1.6	8.1	2.5	30.7	30.7	1.5	0.0	12	366
Health Conditions										
Prevalence of CHD (%) *	3.6	4.3	0.8	1.3	4.6	8.7	2.6	1.8	11	184
Prevalence of asthma (%) *	8.5	8.1	1.3	1.9	11.2	13.2	6.0	3.6	11	184
Prevalence of diabetes (%) *	8.1	8.5	1.2	2.1	10.6	16.5	5.4	2.2	11	184
Prevalence of AMI (%) *	3.6	4.1	0.6	1.3	4.8	10.2	2.8	1.7	11	184
Prevalence of obesity (%) *	24.7	26.0	4.0	4.1	32.7	35.7	18.7	14.0	11	182
Prevalence of stroke (%) *	2.6	2.7	0.7	1.0	3.7	6.5	1.5	0.7	11	184
Prevalence of ever smoked (%)*	18.3	19.6	3.1	4.0	23.1	34.4	14.2	6.5	11	184
Prevalence of exercise (20 minutes, %)*	29.5	28.0	2.7	4.8	33.8	44.1	23.7	15.4	11	183
Prevalence of exercise (30 minutes,%)*	50.2	49.7	2.3	5.4	55.3	67.1	47.4	37.3	11	182
Non-accidental mortality (deaths per 100,000 people)	756.2	950.6	204.1	249.6	1139.5	1958.4	361.6	117.7	23	3142

Risk Attribute	Average		Standard Deviation		Maximum		Minimum		Sample Size (# of counties or cities)	
	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset	Urban Study Areas	U.S. Dataset
All cause mortality (deaths per 100,000 people)	810.1	1022.3	217.4	258.6	1257.8	2064.2	402.5	176.8	23	3142
Cardiovascular mortality (deaths per 100,000 people)	310.5	392.1	93.9	121.0	459.6	970.4	122.4	37.5	23	3142
Respiratory mortality (deaths per 100,000 people)	66.2	97.3	17.0	32.3	90.1	351.0	34.8	13.3	23	3136
Air Quality and Climate										
O_3 4th high maximum 8-hr average (ppm)	0.087	0.077	0.009	0.010	0.105	0.126	0.072	0.033	23	725
O_3 seasonal mean (ppb)	33.9	34.5	5.4	6.6	51.0	64.8	25.8	8.6	22	671
O_3 seasonal mean of maximum 8-hr average (ppb)	50.7	48.6	7.5	7.2	70.2	79.7	40.8	13.3	22	671
O_3 seasonal mean of 1-hr daily maximum (ppb)	58.8	54.7	7.5	8.0	85.1	92.4	46.5	17.6	22	671
$PM_{2.5}$ annual mean (µg/m3)	14.1	11.7	2.6	3.1	16.9	22.5	8.4	3.4	23	617
$PM_{2.5}$ 98th %ile daily average (µg/m3)	35.8	30.7	8.1	9.3	59.0	81.1	21.2	9.1	23	617
Average temperature (°F)	57.2	57.2	5.0	7.9	70.3	76.2	50.1	39.0	23	202
July temperature long term average (°F)	76.0	75.9	3.4	5.4	83.3	93.7	68.5	55.5	23	3104
July Relative Humidity long term average (%)	61.5	56.2	10.2	14.6	70.0	80.0	28.0	14.0	23	3104
Exposure Determinants										
% No air conditioning*	15.5	16.6	85.7	79.1	42.9	86.7	0.4	0.0	12	76
C-R Estimates										
Non-accidental mortality O_3 risk*	0.000388	0.000322	0.000217	0.000131	0.000917	0.000917	0.000148	-0.000033	12	98
All Cause mortality O_3 risk*	0.000627	0.000527	0.000314	0.000205	0.001092	0.001092	0.000163	0.000096	12	48
Respiratory mortality O_3 risk*	0.000877	0.000800	0.000282	0.000186	0.001424	0.001424	0.000307	0.000307	12	48
Cardiovascular mortality O_3 risk*	0.000898	0.000825	0.000173	0.000124	0.001064	0.001064	0.000418	0.000418	12	48

*Attribute for which only city-specific data were available.

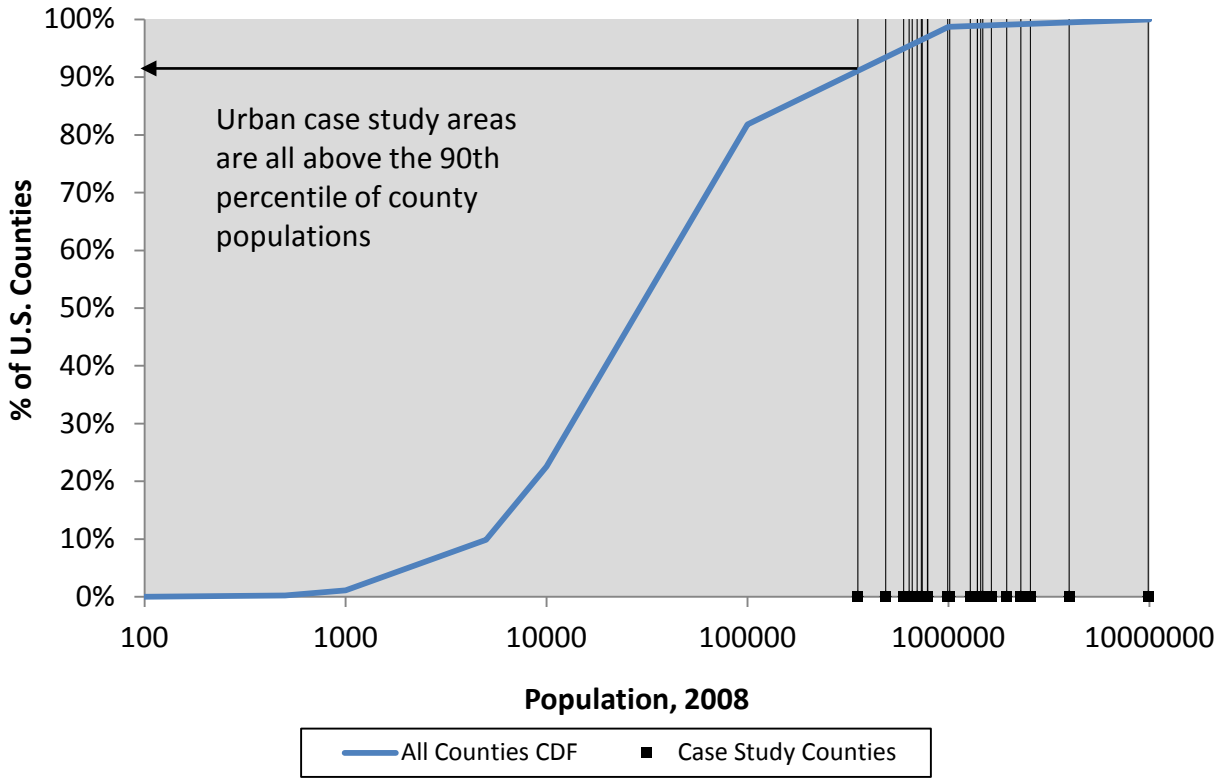


Figure 8.12 Comparison of county-level populations of urban case study area counties to the frequency distribution of population in 3,143 U.S. counties.

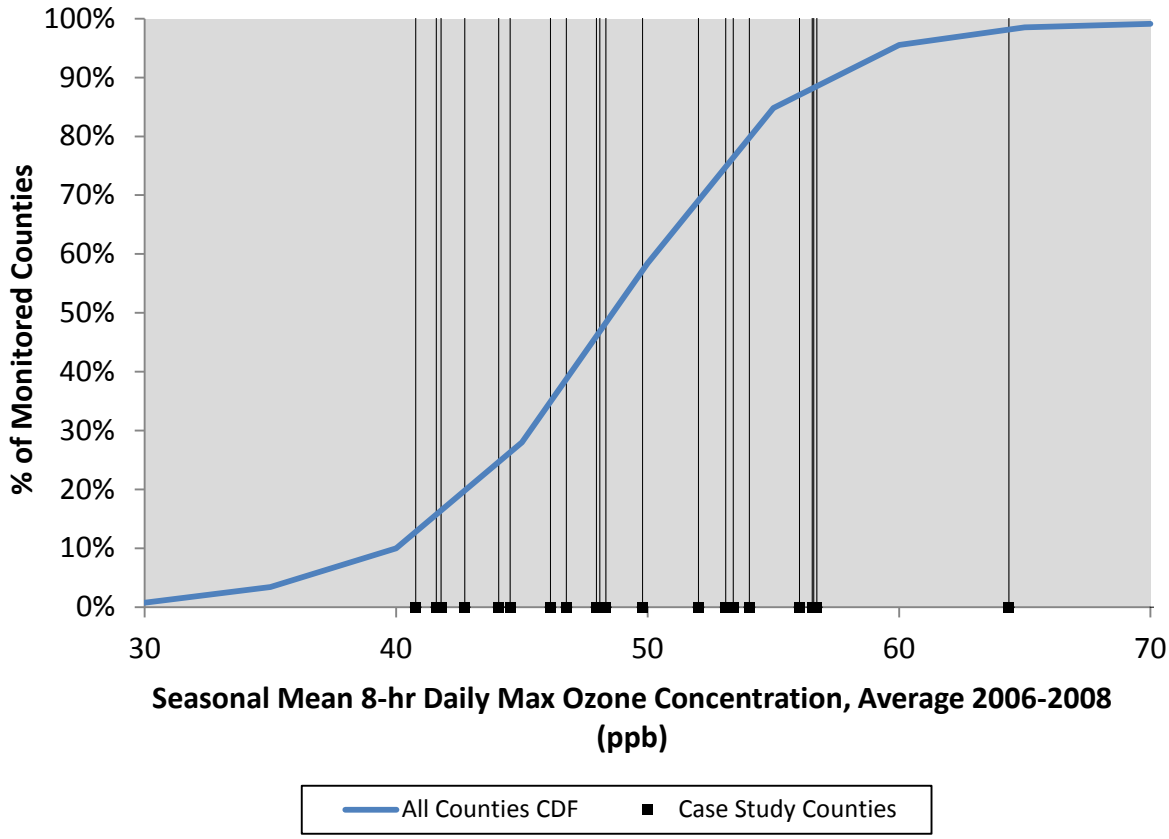


Figure 8.13 Comparison of county-level seasonal mean 8-hr daily maximum O_3 concentrations in urban case study area counties to the frequency distribution of seasonal mean 8-hr daily maximum O_3 concentrations in 671 U.S. counties with O_3 monitors.

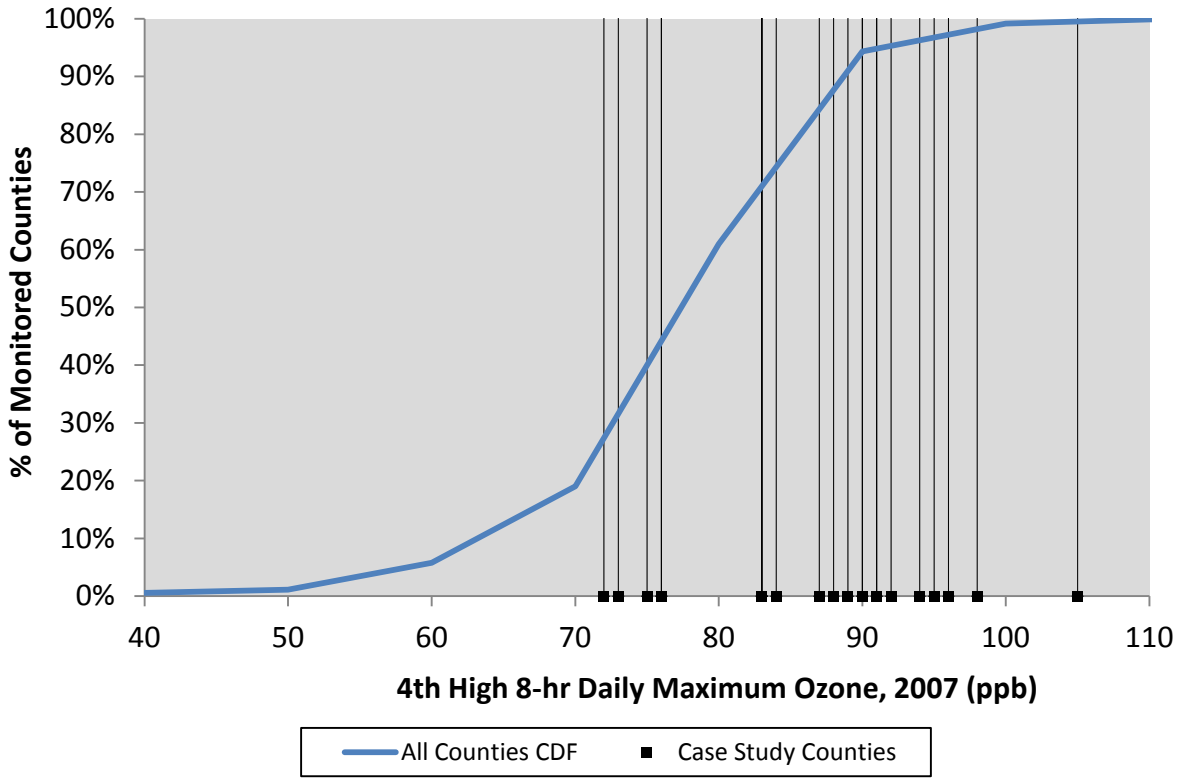


Figure 8.14 Comparison of 2007 county-level 4th high 8-hr daily maximum O_3 concentrations in urban case study area counties to the frequency distribution of 2007 4th high 8-hr daily maximum O_3 concentrations in 725 U.S. counties with O_3 monitors.

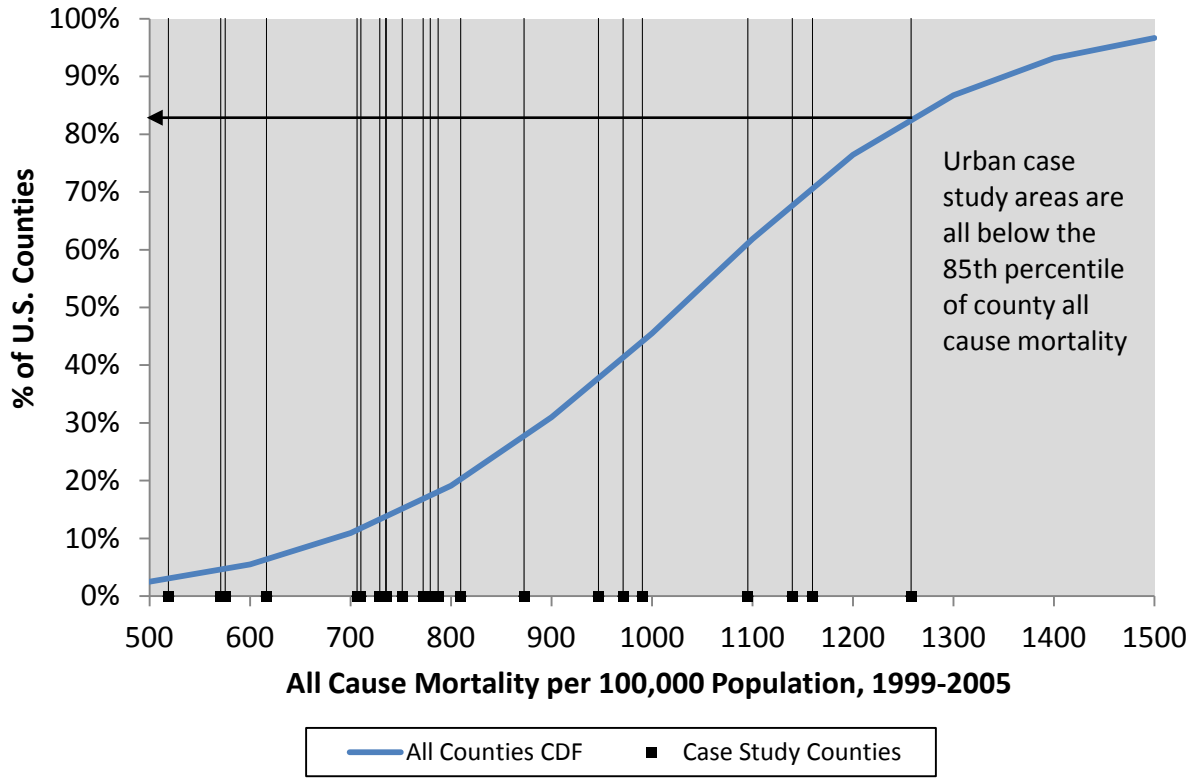


Figure 8.15 Comparison of county-level all-cause mortality in urban case study area counties to the frequency distribution of all-cause mortality in 3,137 U.S. counties.

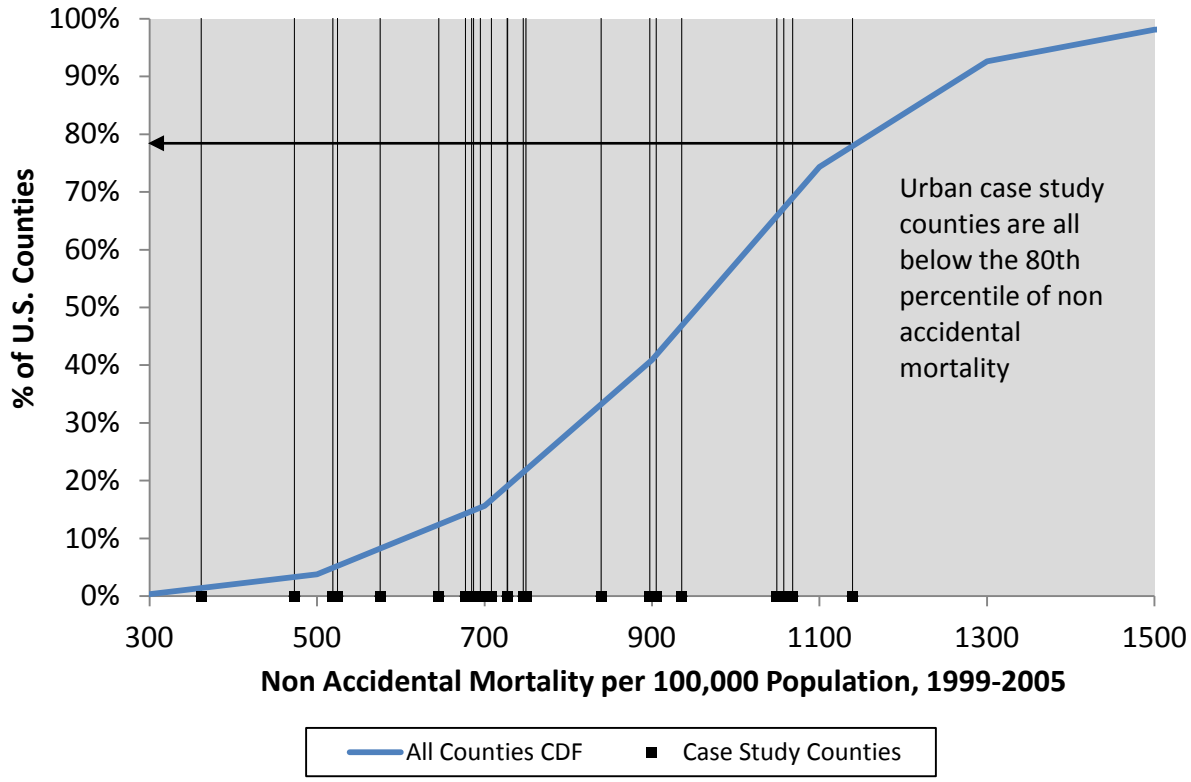


Figure 8.16 Comparison of county-level non-accidental mortality in urban case study area counties to the frequency distribution of non-accidental mortality in 3,135 U.S. counties.

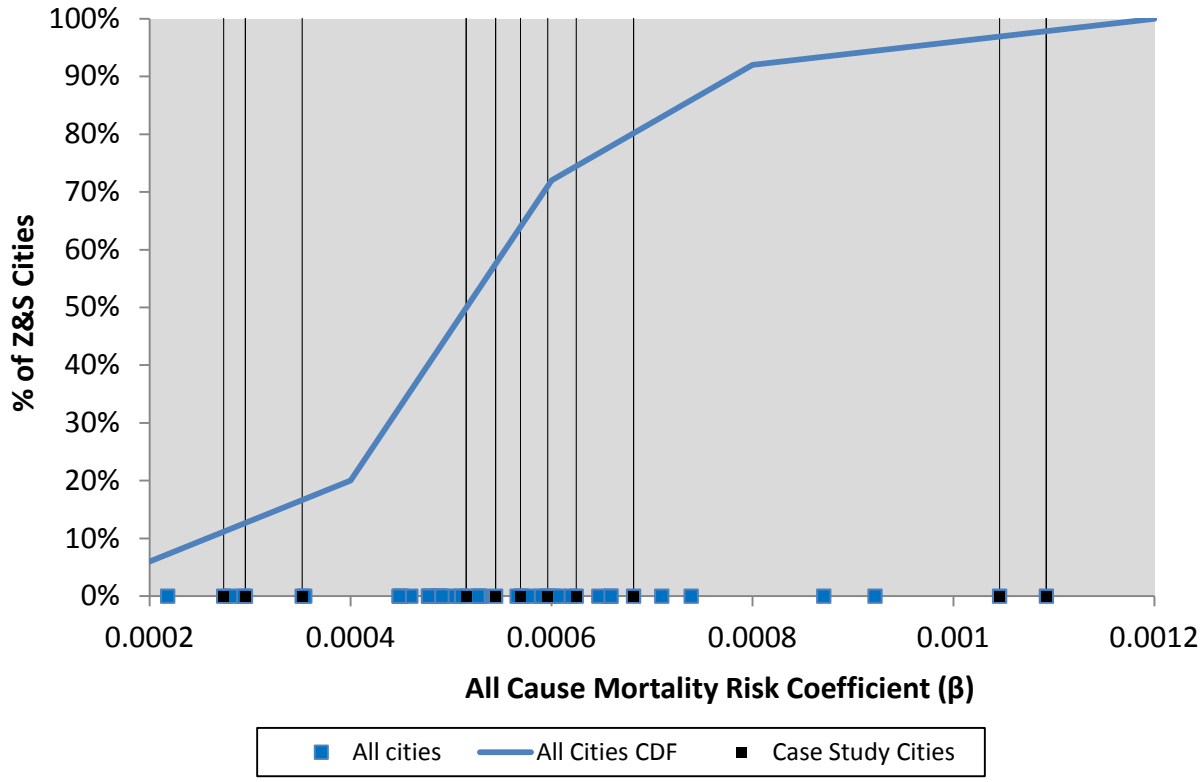


Figure 8.17 Comparison of city-level all-cause mortality risk coefficients from Zanobetti and Schwartz (2008) in urban case study areas to the frequency distribution of all-cause mortality risk coefficients from Zanobetti and Schwartz (2008) in 48 U.S. cities.

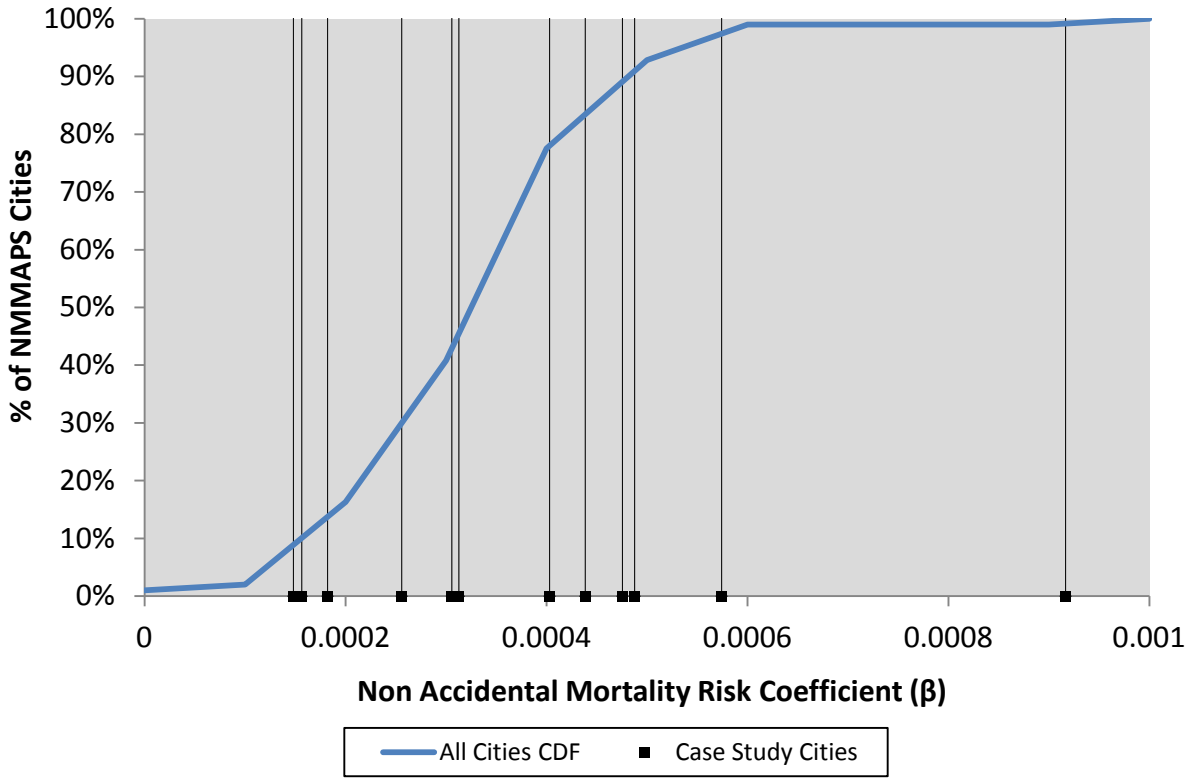


Figure 8.18 Comparison of city-level national prior Bayes-shrunken non-accidental mortality risk coefficients from Smith et al. (2009) in urban case study areas to the frequency distribution of national prior Bayes-shrunken non-accidental mortality risk coefficients from Smith et al. (2009) in 98 U.S. cities.

Figure 8.19 through Figure 8.24 show national CDFs and the urban study area values for several selected potential risk attributes. These potential risk attributes do not directly enter the risk equations, but have been identified in the literature as potentially affecting the magnitude of the O_3 C-R functions reported in the epidemiological literature. Comparison graphs for other risk attributes are provided in Appendix 4-A. The selected urban study areas do not capture the higher end percentiles of several risk characteristics, including populations 65 years and older, baseline cardiovascular disease prevalence, baseline respiratory disease prevalence, and smoking prevalence. Summarizing the analyses of the other risk attributes, we conclude that the urban study areas provide adequate coverage across population, population density, O_3 levels (seasonal mean, seasonal mean 8-hr daily maximum, and seasonal mean 1-hr daily maximum), $PM_{2.5}$ co-pollutant levels, temperature and relative humidity, unemployment rates, percent non-white population, asthma prevalence obesity prevalence, income, and less than high school education.

We also conclude that while the urban study areas cover a wide portion of the distributions, they do not provide coverage for the upper end of the distributions of percent of population 65 and older (below 60th percentile), percent of population 85 years and older (below 75th percentile), prevalence of angina/coronary heart disease (below 70th percentile), prevalence of diabetes (below 85th percentile), stroke prevalence (below 90th percentile), prevalence of heart attack (below 80th percentile), prevalence of smoking (below 85th percentile), all-cause mortality rates (below 85th percentile), non-accidental mortality rates (below 80th percentile), cardiovascular mortality rates (below 75th percentile) and respiratory mortality rates (below 50th percentile), and percent of residences without air conditioning (below 90th percentile). In addition, the urban study areas do not capture the highest or lowest ends of the distribution of exercise prevalence and do not capture the low end of the distribution of public transportation use (above the 65th percentile).

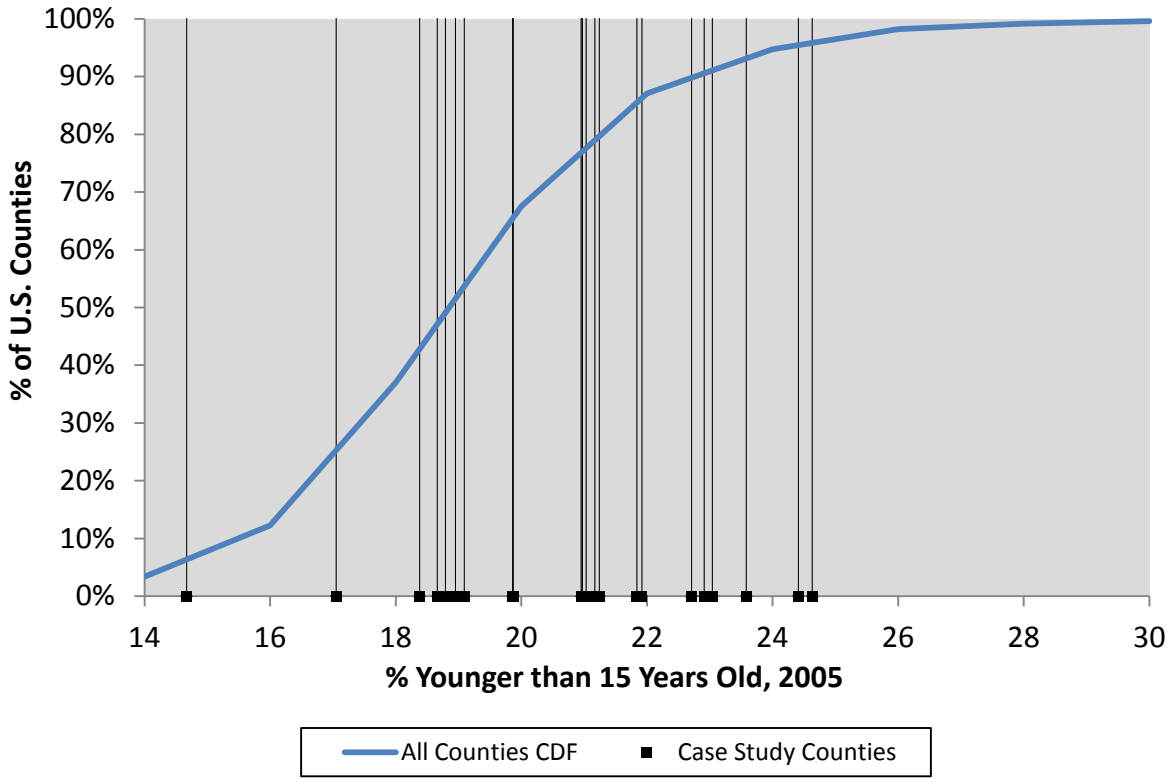


Figure 8.19 Comparison of county-level percent of population 0 to 14 years old in urban case study area counties to the frequency distribution of percent of population 0 to 14 years old in 3,141 U.S. counties.

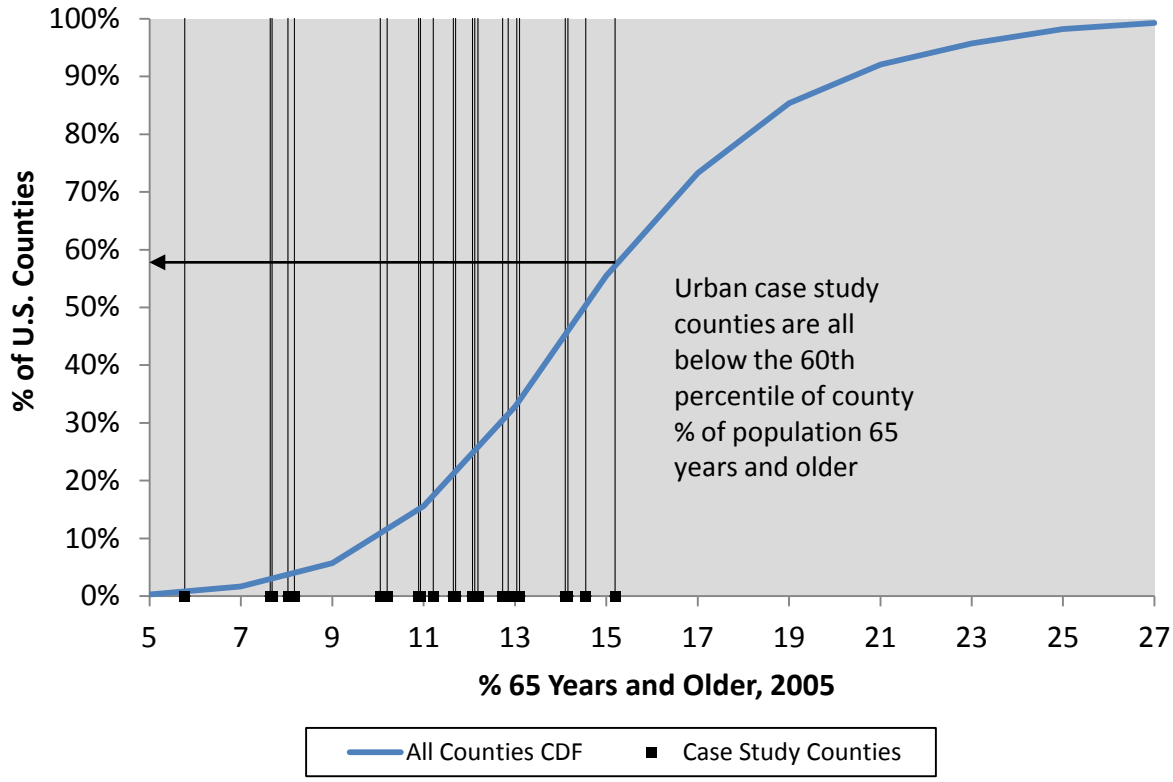


Figure 8.20 Comparison of county-level percent of population age 65 years old and older in urban case study area counties to the frequency distribution of percent of population age 65 and older in 3,141 U.S. counties.

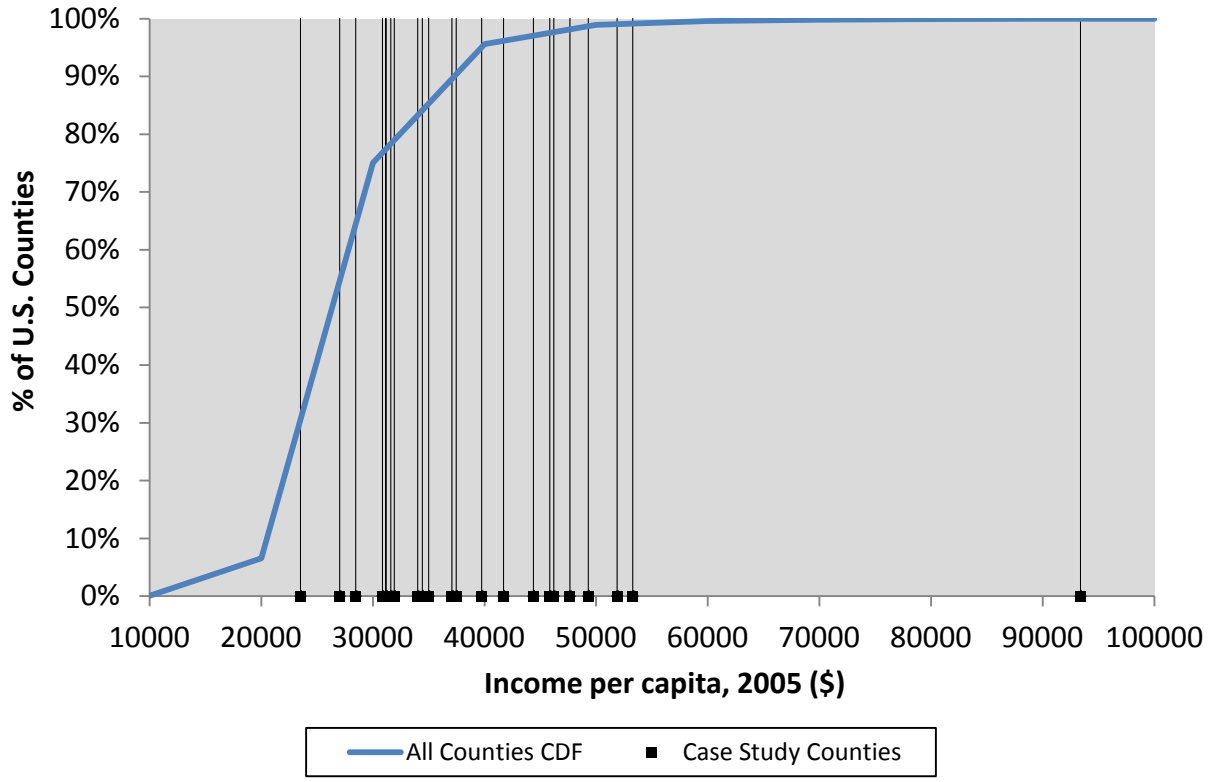


Figure 8.21 Comparison of county-level income per capita in urban case study areas to the frequency distribution of income per capita in 3,141 U.S. counties.

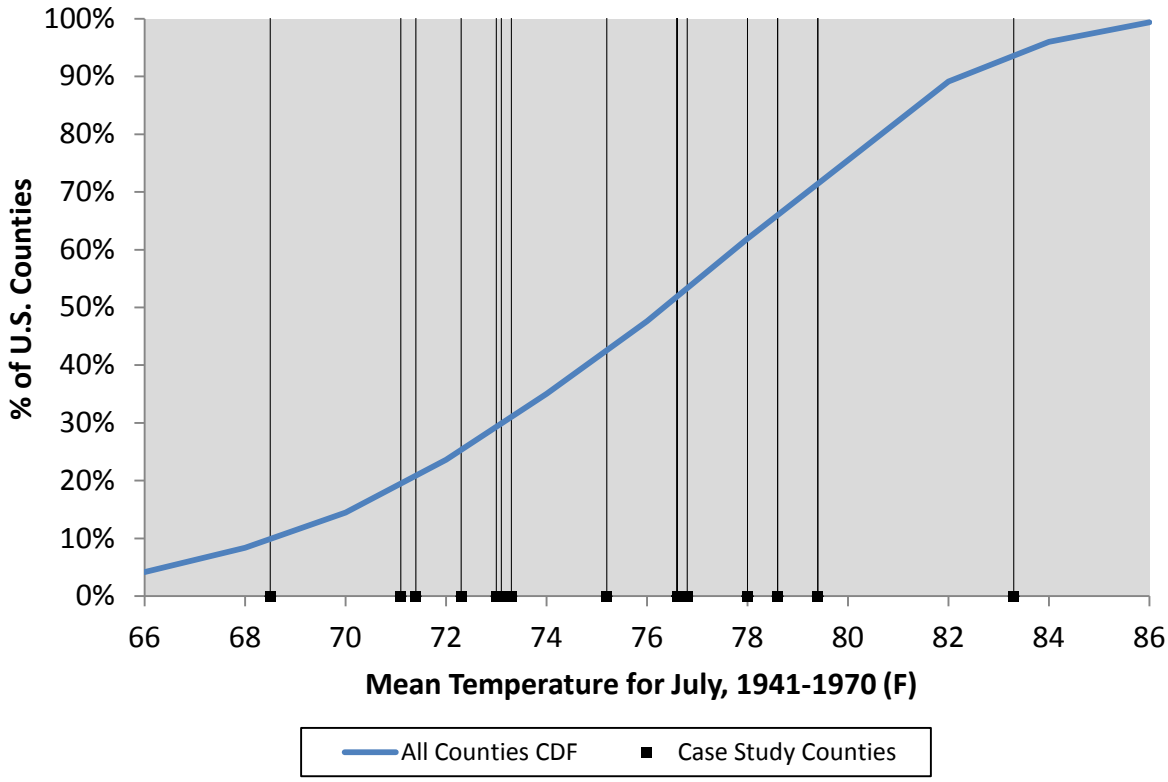


Figure 8.22 Comparison of county-level July temperature in urban case study area counties to the frequency distribution of July temperature in all U.S. counties.

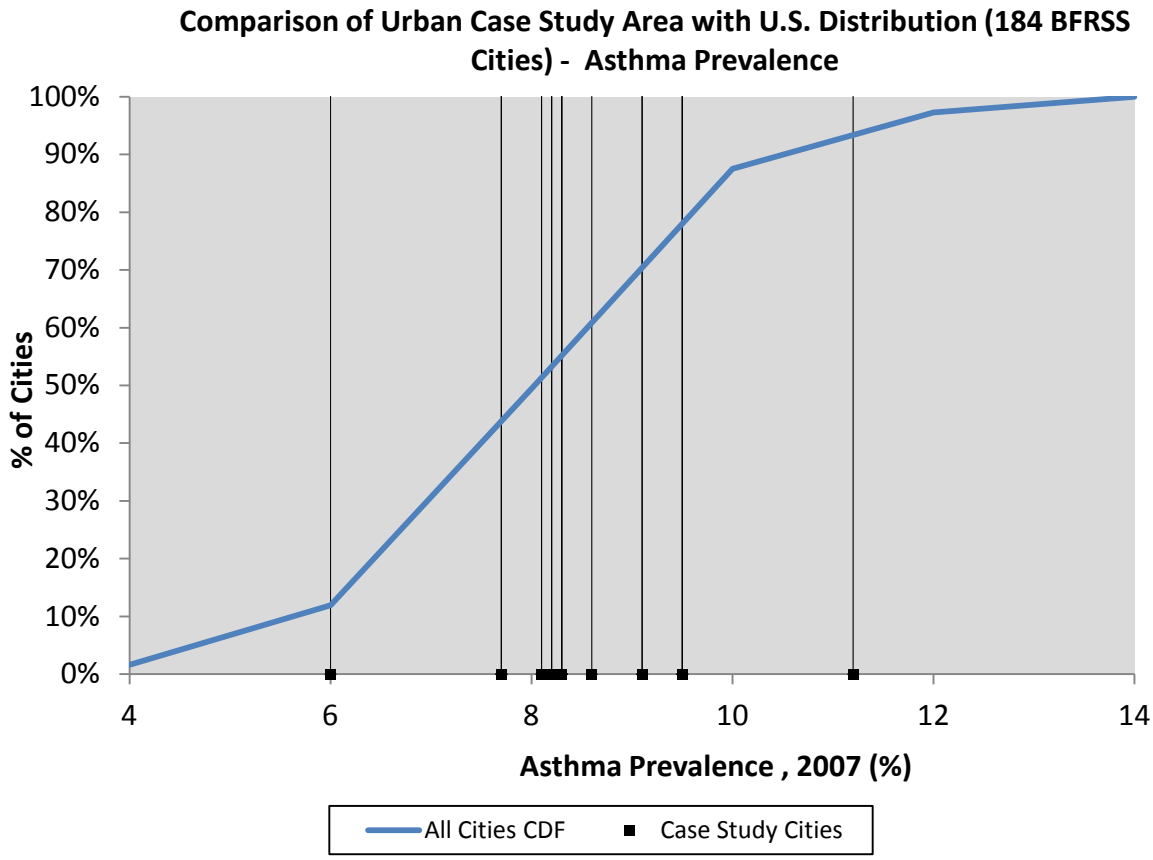


Figure 8.23 Comparison of city-level asthma prevalence in urban case study areas to the frequency distribution of asthma prevalence in 184 U.S. cities.

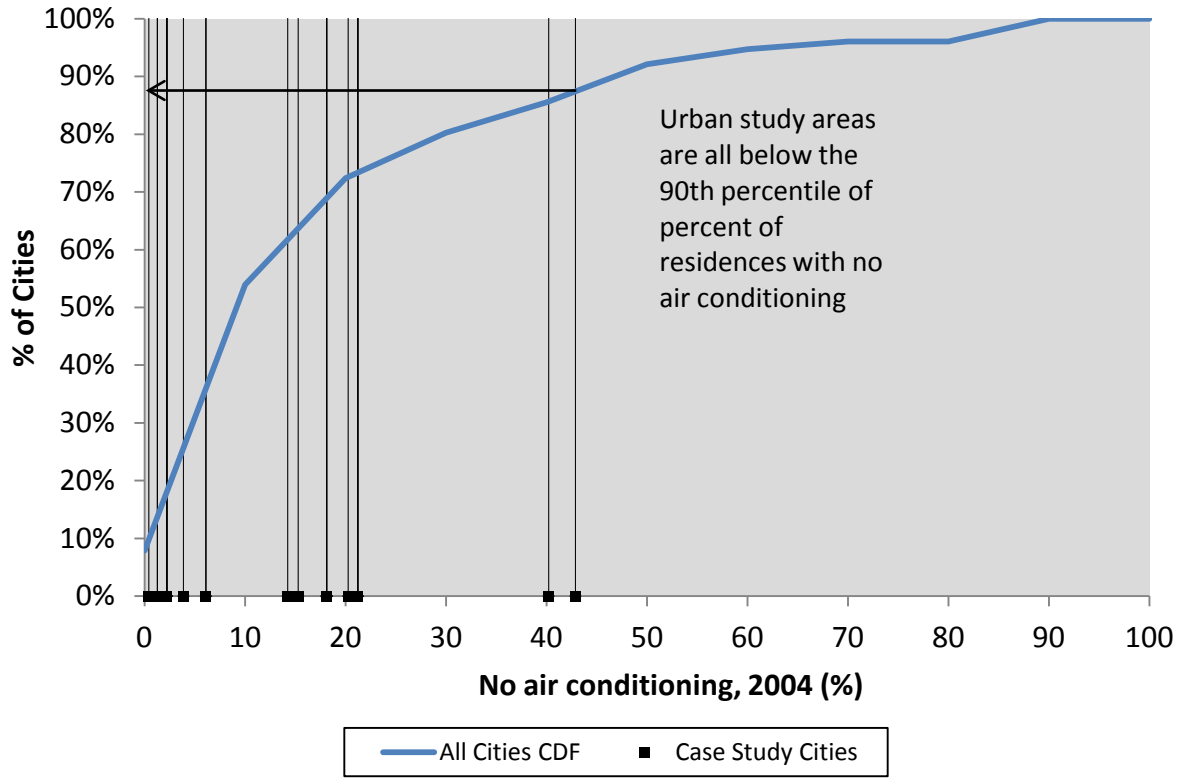


Figure 8.24 Comparison of city-level air conditioning prevalence in urban case study areas to the frequency distribution of air conditioning prevalence in 76 U.S. cities.

Based on the above analyses, we can draw several inferences regarding the representativeness of the urban case studies. First, the case studies represent urban areas that are among the most populated in the U.S. Second, they represent areas with relatively high levels of O_3 (4th high 8-hr daily maximum, seasonal mean 8-hr daily maximum, seasonal mean 1-hr daily maximum, and seasonal mean). Third, they capture well the range of city-specific effect estimates found by Smith et al. (2009) and Zanobetti and Schwartz (2008) studies. These three factors would suggest that the urban study areas should capture well overall risk for the nation, with a potential for better characterization of the high end of the risk distribution.

However, there are several other factors that suggest that the urban study areas may not be representing areas that may have a high risk per ppb of O_3. Several of the factors with underrepresented tails, including age and baseline mortality are spatially correlated (R=0.81), so that certain counties which have high proportions of older adults also have high baseline mortality and high prevalence of underlying chronic health conditions. Because of this, omission

of certain urban areas with higher percentages of older populations, for example, cities in Florida, may lead to underrepresentation of high risk populations. However, with the exception of areas in Florida, most locations with high percentages of older populations have low overall populations, less than 50,000 people in a county. And even in Florida, the counties with the highest O_3 levels do not have a high percent of older populations. This suggests that while the risk per exposed person per ppb of O_3 may be higher in these locations, the overall risk to the population is likely to be within the range of risks represented by the urban case study locations.

Due to data limitations, we were only able to assess the representativeness of the urban study areas in terms of one exposure-related attribute, air conditioning prevalence. Assessing the representativeness of the urban study areas in terms of air conditioning prevalence, we found that the urban study areas do not capture the highest end of percent of residences without air conditioning. If the cities with the lowest air conditioning prevalence also have high O_3 levels, we could be missing a high risk portion of the population that is exposed to O_3 indoors as air infiltrates indoors from outdoors. However, 4th highest 8-hr daily maximum O_3 levels in the cities in the top 10th percentile of percentage of residences without air conditioning (mainly in northern California and Washington) are approximately average (0.08 ppm) or lower than average. Since these concentrations are not the highest found across the U.S., we are likely not excluding a high risk population that has both low air conditioning prevalence and high O_3 concentrations, and the overall risk to the population is likely to be within the range of risks represented by the urban case study locations.

8.2.2 Analysis Based on Consideration of National Distribution of O_3-Related Mortality Risk

In this section we discuss the second representativeness analysis which identifies where the 12 urban study areas examined in Chapter 7 fall along the distribution of estimated national-scale mortality risk. This assessment reveals whether the baseline O_3 mortality risks in the 12 urban case study areas represent more typical or higher end risk relative to the national risk distribution presented in Section 8.1. For consistency, we compare the national O_3 mortality risk distribution to the O_3 mortality risk results for the urban study areas that were generated from the national-scale assessment in Section 8.1, rather than the results from the urban study area analysis in Chapter 7 which uses different methods. To be consistent with the national-scale assessment, we define the urban study areas here as they were defined in the epidemiology studies, rather than including full core-based statistical areas as in Chapter 7. The results of this representativeness analysis are presented graphically in Figure 8.25 through Figure 8.27, which display the cumulative distribution of total mortality attributable to ambient O_3 at the county level developed as part of the national-scale analysis. Values for the 23 counties included in the

urban case study areas as defined in the epidemiology studies are then superimposed on top of the cumulative distribution to assess the representativeness of the urban case study areas.

For the results based on Smith et al. (2009) effect estimates, New York City and Philadelphia have the highest percentage of May-September non-accidental mortality attributable to ambient O_3 of the 12 urban study areas and are located at the highest end of the distribution of U.S. O_3-related mortality risk (Figure 8.25). Of the 12 urban study areas, these two cities had the highest effect estimates found by Smith et al. (2009; See Appendix 4-A). Boston and Los Angeles had the lowest O_3-related mortality risk of the 12 urban study areas and are located at the lowest end of the U.S. distribution. Overall, O_3 mortality risk in the 12 urban study areas are representative of the full distribution of U.S. O_3-related mortality risk, with the mean percentage of May-September non-accidental mortality for all ages of 1.5% (95% confidence interval, 1.1-1.8%).

For the results based on Zanobetti and Schwartz (2008) effect estimates, Detroit and New York City are at the very highest end of the U.S. distribution of county-level risk of June-August all-cause mortality due to ambient O_3 (Figure 8.26). These two cities had the highest effect estimates of the 48 cities included in the study (see Appendix 4-A). The high effect estimates in Detroit and New York City could be due to high rates of public transportation use (for New York City), low air conditioning prevalence, high smoking prevalence (in Detroit), high incidence of mortality and other adverse health outcomes (e.g. diabetes, stroke, acute myocardial infarction, etc.), and high unemployment rates. Houston and Los Angeles had the lowest risk and were located at the very lowest end of the U.S. distribution of county-level risk of mortality due to ambient O_3. These two cities had the lowest effect estimates found by Zanobetti and Schwartz (2008), possibly because they cover a large spatial extent and have high rates of time spent driving, which could lead to exposure misclassification in the underlying epidemiologic study. Houston also has a very high rate of air conditioning use (nearly 100% of residences) and Los Angeles has been shown to have high rates of adaptive behavior on high ambient O_3 days (i.e. more time spent indoors as a result of high ambient O_3 concentrations; Neidell 2009, 2010), both of which would lead to lower personal O_3 exposure relative to other cities. Overall, O_3 mortality risk in the 12 urban study areas are representative of the full distribution of U.S. O_3-related mortality risk, with the mean percentage of June-August all-cause mortality for all ages of 2.5% (95% confidence interval, 1.7-3.0%).

For the results based on Jerrett et al. (2009) effect estimates, the 12 urban study areas are centered more in the middle of the distribution of U.S. county-level risk of adult (ages 30 and older) respiratory mortality due to ambient O_3 exposure. These results are based on the application of a single national average effect estimate to all gridcells across the U.S., rather than city-specific effect estimates as were applied for the results based on Smith et al. (2009) and

Zanobetti and Schwartz (2008) effect estimates. Therefore, the location of the urban study areas on the distribution of county-level risk is driven mainly by O_3 concentration and not by the effect estimate. While Denver, Atlanta, Sacramento, and Los Angeles are at the highest end of the U.S. distribution, Figure 8.27 shows that some counties have a higher percentage of mortality attributable to O_3 than these four cities. Overall, O_3 mortality risk in the 12 urban study areas are representative of the full distribution of U.S. O_3-related mortality risk, with the mean percentage of April-September respiratory mortality for adults ages 30 and older of 18.7% (95% confidence interval, 15.2-21.5%). However, we are not capturing the very highest end of O_3-related risk based on Jerrett et al. (2009) effect estimates in the 12 urban study areas.

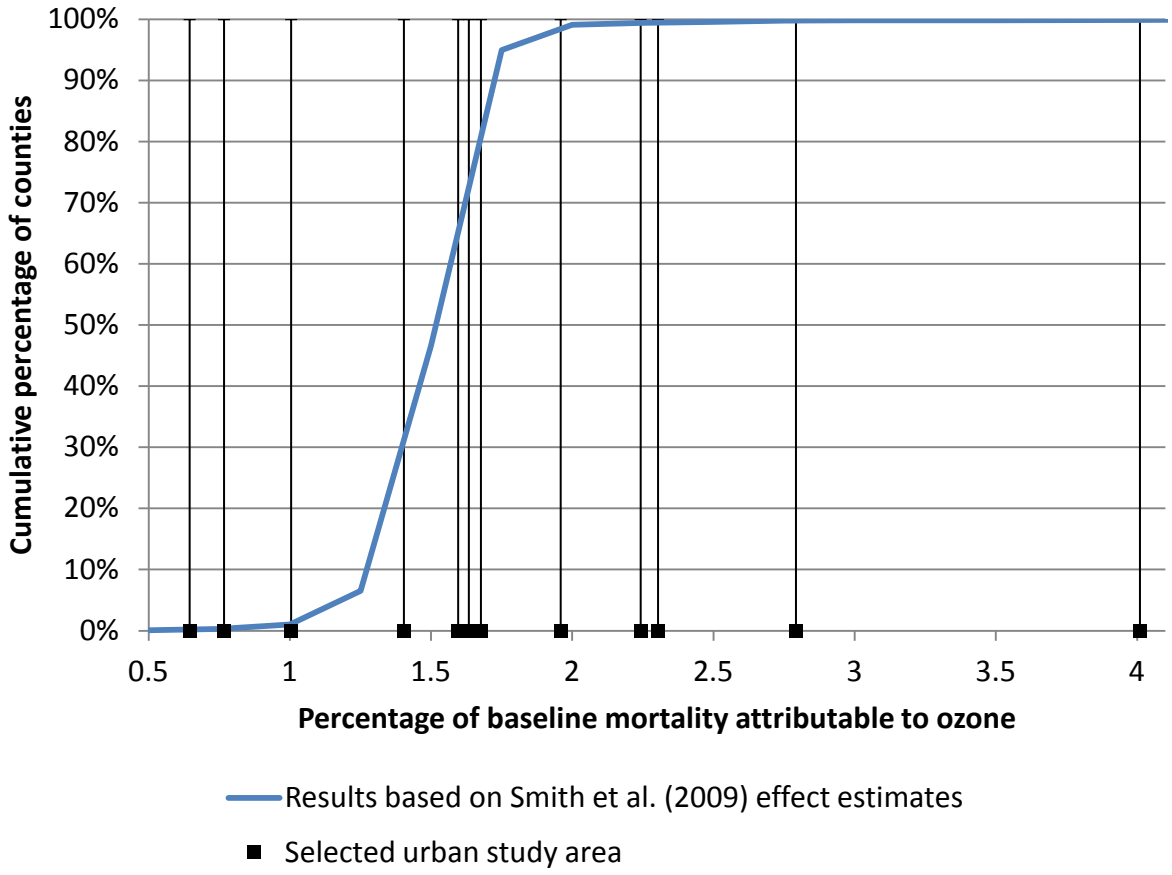


Figure 8.25 Cumulative distribution of county-level percentage of May-September non-accidental mortality for all ages attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Smith et al. (2009) effect estimates.

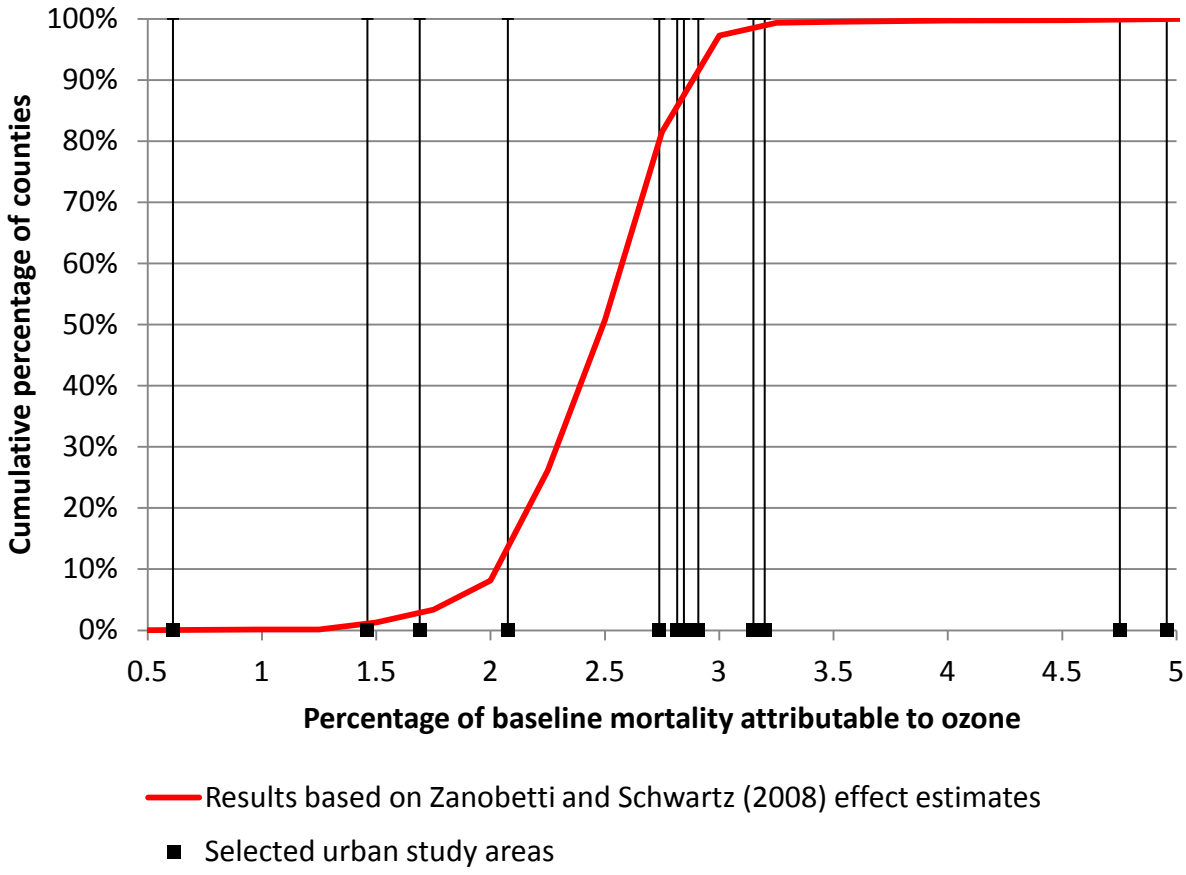


Figure 8.26 Cumulative distribution of county-level percentage of June-August all-cause mortality for all ages attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Zanobetti and Schwartz (2008) effect estimates.

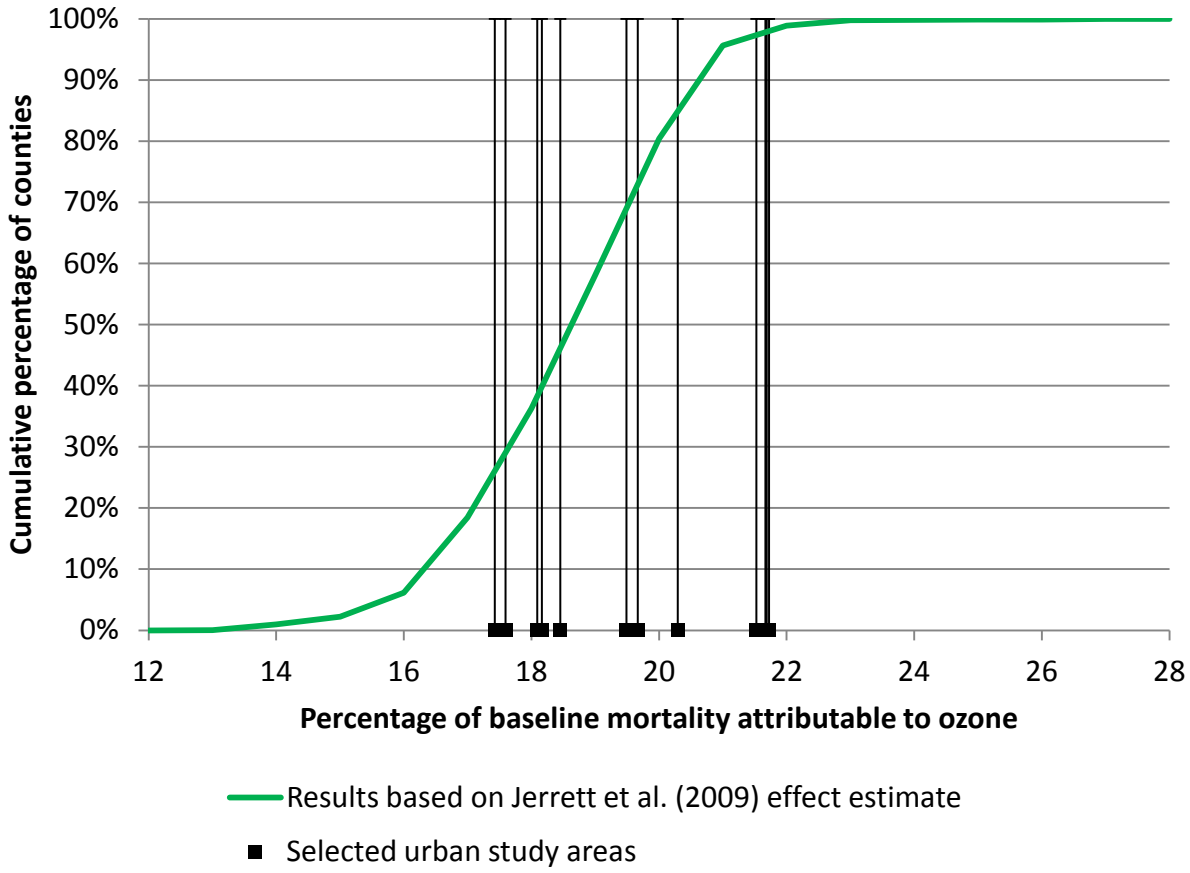


Figure 8.27 Cumulative distribution of county-level percentage of April-September adult (age 30+) respiratory mortality attributable to 2006-2008 average O_3 for the U.S. and the locations of the selected urban study areas along the distribution, using Jerrett et al. (2009) effect estimates and city definitions from Zanobetti and Schwartz (2008).

8.2.3 Analysis Based on Consideration of National Responsiveness of O_3 Concentrations to Emissions Changes

Estimates of O_3 response to precursor emissions reductions (NOx and VOC) are important inputs to estimation of risk for scenarios of just meeting existing and alternative O_3 standards. To evaluate the national representativeness of O_3 responses to decreases in precursor emissions in the 15 urban study areas, we examine two different types of air quality data. In section 8.2.3.1 we examine ambient O_3 trends that have been measured at monitor locations across the country over a recent period of decreasing NOx emissions. This analysis provides real-world observations but does not isolate the effects of emissions changes alone and can only characterize past phenomena. In section 8.2.3.2, we look at air quality model predictions of temporal and spatial patterns of O_3 changes in response to further NOx reductions from 2007

levels. This analysis is subject to typical model limitations but has the advantage of isolating the effects of precursor emissions changes and has the ability to simulate how O_3 would change in response to NOx (and VOC) emissions reductions (relative to recent 2007 levels) similar to those used in the HDDM adjustment scenarios for just meeting existing and alternative standards. These two complimentary analyses give qualitatively similar results, building confidence that the overarching conclusions are robust across the US as a whole.

8.2.3.1 Ambient patterns in trends of measured O_3 concentrations

This section describes how annual distributions of O_3 measurements collected by EPA's national monitoring network have changed between 1998 and 2011. These years were chosen because large reductions in anthropogenic NOx emissions have occurred over this time period especially in the Eastern half of the United States. From 2000 to 2011 nationwide NOx emissions were cut almost in half (from 22.6 to 12.9 million tons per year)[3]. However, it should be noted that these reductions did not occur uniformly across the country. Improvements in vehicle emissions standards helped reduce NOx emissions in many locations throughout the country. In contrast, EPA rules like the NOx SIP call were focused on controlling emissions from power plants in the Eastern US and consequently there have been relatively larger reductions in NOx emissions in the East. In addition, some urban areas which have traditionally had high O_3 levels, like Los Angeles and Houston, have substantially cut local NOx and VOC emissions to improve their air quality. Also, there may be some localized areas in which NOx emissions have increased due to population growth, new sources such as oil and gas development, or increased wildfire activity. Appendix 8-C provides plots of emissions trends by region of the U.S. These plots show that each of nine regions of the U.S. have experienced decreasing NOx emissions ranging from approximately a 20% decrease to a 45% decrease from 2002 to 2011 depending on the region. Conversely, VOC emissions have increased in some regions since 2002 (the South, the Southwest, and the West-North-Central) and decreased in others. Due to non-linear O_3 formation chemistry and the potential for changes in local chemical regimes resulting from these emissions reductions, past trends may not reflect the ambient changes which will occur from future emissions reductions. Nonetheless, these ambient data provide information on actual O_3 changes in response to emissions reductions and can give insight into the types of changes in O_3 that have occurred both within and outside the urban study areas.

First, we look at national maps which show changes in 50th percentile and 95th percentile summer season (April-October)[4] 8-hour daily maximum O_3 values (Figure 8.28,Figure 8.29).

[3] Data were accessed from EPA's emission trend website on August 15, 2013: http://www.epa.gov/ttn/chief/trends/trends06/national_tier1_caps.xlsx

[4] The April-October time period corresponds to the required monitoring season for most of the 12 urban areas. Therefore, in selecting a consistent time period that could be analyzed for the urban case study areas, we chose to

These maps reflect the absolute (ppb) difference between O_3 percentiles from two three-year periods (2001-2003 and 2008-2010)[5]. Figure 8.28 shows that increases in median O_3 concentrations occurred in many large urban areas including both study area locations in Chapter 7[6] and non-study area locations[7]. Only a few monitors with increasing median O_3 appear outside of cities, most notably in southwestern Colorado and central Kansas. The increases in urban areas are likely explained by O_3 "disbenefits" to NOx reductions which were described in Chapter 4, Appendix 4-C and in the following section of this chapter. Widespread decreases of median O_3 in suburban and rural locations suggest the efficacy of large NOx emissions reductions on reducing O_3 over large regions of the country. Finally, the less frequently observed cases of median O_3 increases in rural areas are likely caused by different phenomena. Cooper et al. (2012) suggested that increasing rural O_3 in the Western US may be due to increasing oil and gas development, wildfires and O_3 transport from Asia. Conversely, Figure 8.29 shows that 95th percentile O_3 values for these two sets of years decrease in almost all urban as well as rural areas of the country. Only a few sites in Colorado, Nevada, and California show any increases in 95th percentile O_3 between 2001-2003 and 2008-2010. The consistent decreases across most of the United States indicate that the large NOx reductions from power plants and mobile sources have been quite successful in reducing O_3 on the highest O_3 days. These results suggest that many of the urban case study areas may show O_3 responses that are typical of other large urban areas in the U.S. However, decreasing O_3 in large non-urban portions of the country may not be fully captured in the urban case studies.

use the April-October time period in Chapter 4 for composite monitor distributions to summarize ozone values relevant to the epidemiology-based risk assessment.

[5] These two three-year periods were chosen to represent years before and after most NOx emission reductions were in place. In addition the 2001-2003 period was used to designate areas for the 1997 8-hour ozone standard and the 2008-2010 period was used to designate areas for the 2008 8-hour ozone standard. Data from these two time periods have undergone extensive quality checks.

[6] Los Angeles, Denver, Houston, Atlanta, Chicago, Detroit, Cleveland, New York, Philadelphia, and Washington D.C.

[7] San Francisco, Reno, Phoenix, New Orleans, Birmingham, Miami, and Cincinnati

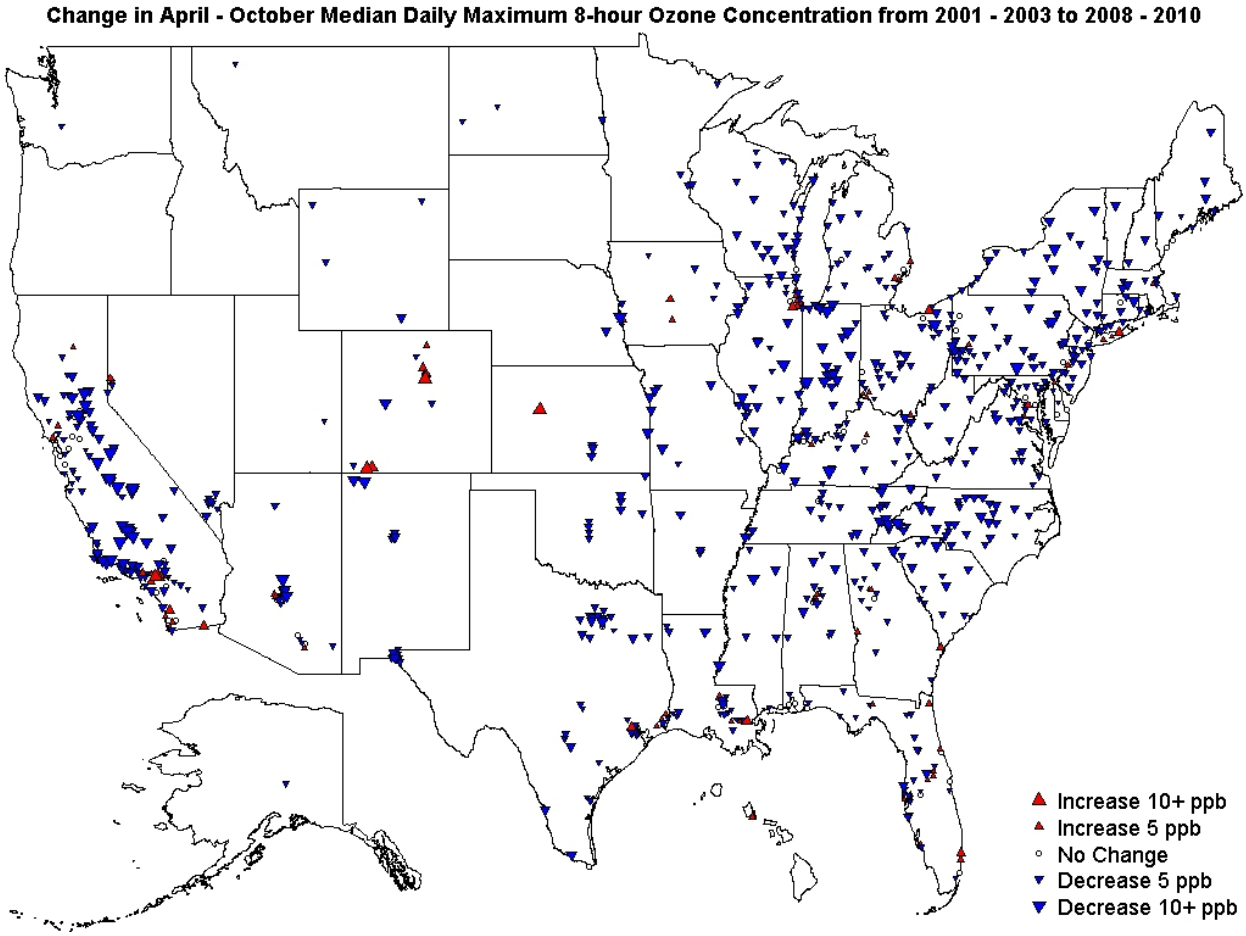


Figure 8.28 Change in 50th percentile summer season (April-October) daily 8-hr maximum O_3 concentrations between 2001-2003 and 2008-2010.

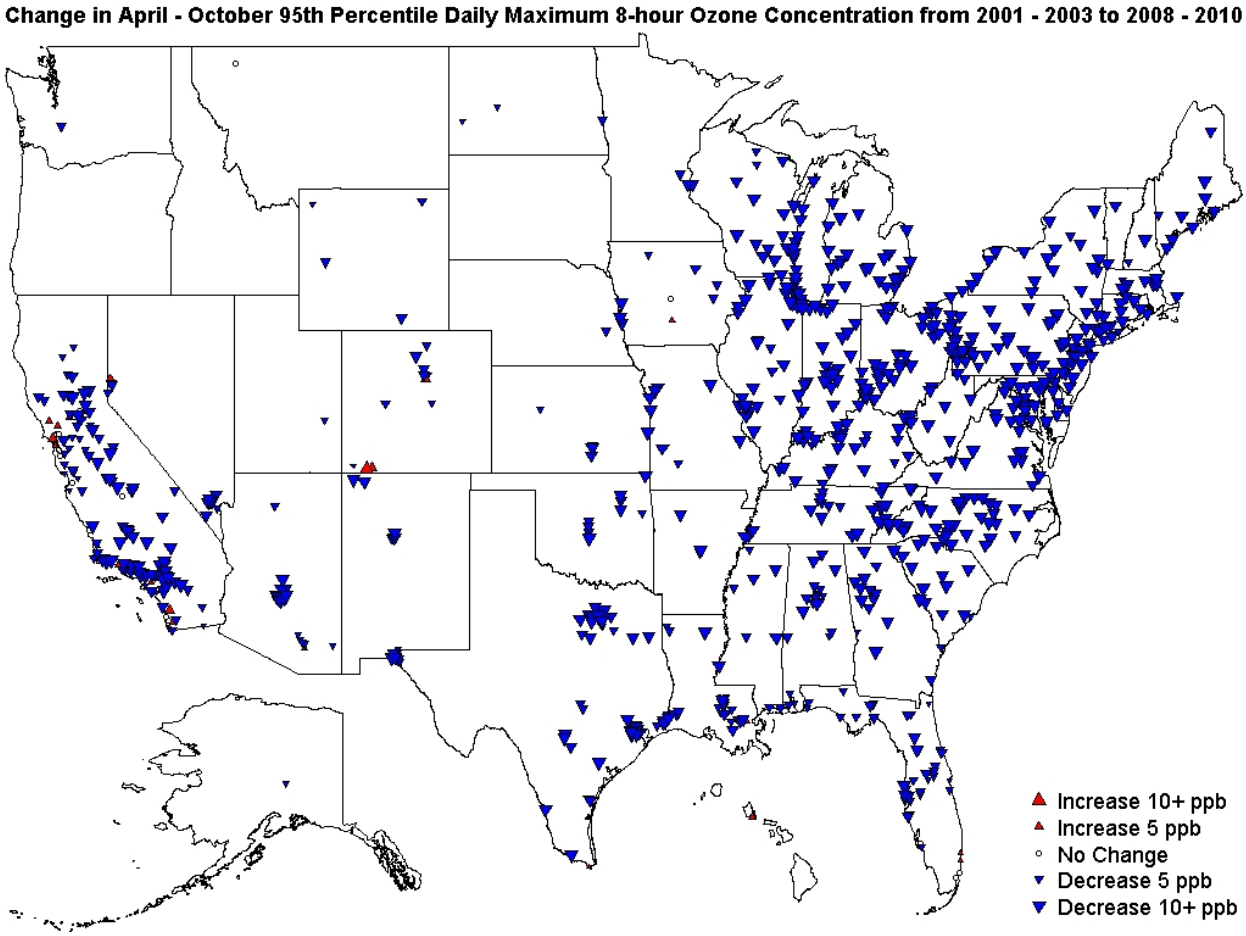


Figure 8.29 Change in 95th percentile summer season (April-October) daily 8-hr maximum O_3 concentrations between 2001-2003 and 2008-2010.

To examine these trends further, we evaluate the 1998-2011 data from the 15 case-study areas. Only monitors within the 15 study areas[8] were analyzed, and within each study area, monitors were put into three groups based on the degree of urbanization. The degrees of urbanization were determined by the population density of the census tract containing the monitor (plotted in Figure 8.30). Population data were obtained from the U.S. Census Bureau[9], and the classes were determined by breaks in the population density calculated from those data: "high population density" (> 1000 people/km^2), "medium population density" (between 400 and 1000 people/km^2), and "low population density" (< 400 people/km^2). Data were additionally split out into three different time periods (all months, warm months: May through September, and cool months: October through April). These warm and cool season categorizations were chosen to isolate effects that are observed at different times of year. The April-October time period

[8] These 15 areas are the 12 urban case study areas in the epidemiological-based risk assessment and the 3 additional exposure urban case study areas.

[9] Obtained from: http://www2.census.gov/geo/tiger/TIGER2010DP1/Tract_2010Census_DP1.zip

which was examined in Figure 8.28 and Figure 8.29 include all warm season and two cold season months and thus show behavior that has influences from both. Summaries were thus calculated for groups of monitors specific to 1) Study Area, 2) Month subset, and 3) Urban class.

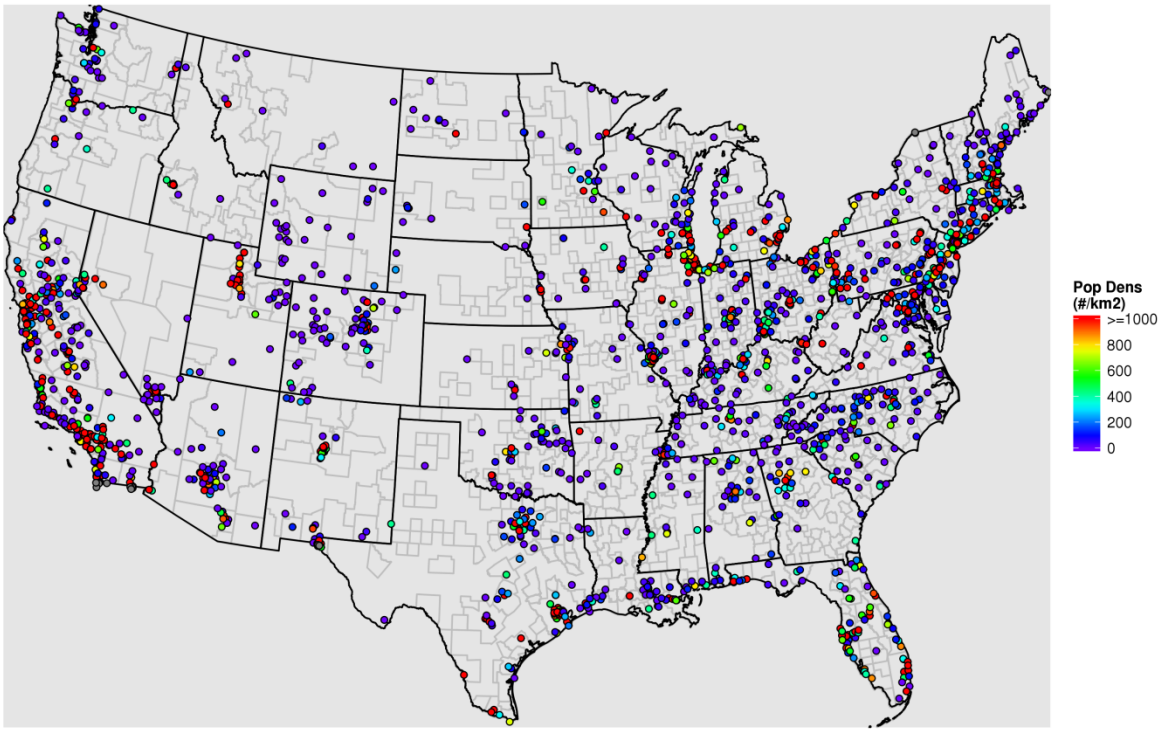


Figure 8.30 Population density at each O_3 monitor.

Figure 8.31, Figure 8.32, and Figure 8.33 display the data described above in ribbon plots for high, medium, and low population density monitor locations in each case study area. The lines bordering the dark and light red ribbons in this plot are (from top to bottom) the 95th, 75th, 25th, and 5th percentiles of the annual data indicated by each panel, and the median (i.e. 50th percentile) is shown by the line in the middle of the central lighter ribbon. The colors of the lines separating the ribbons depict significant trends (dark blue for decreasing and light blue for increasing) or no significant trend (white). Statistical significance for multi-year O_3 trends was determined using the Spearman rank order correlation coefficient (p-value < 0.05). Plots showing a characterization of the entire O_3 distribution (not just discrete cut points of 5th, 25th, 50th, 75th, and 95th percentiles) are provided in Appendix 8-C.

These plots show consistent trends over the past 13 years for O_3, with high O_3 values decreasing fairly uniformly across different regions and areas of different degrees of urbanization. Conversely, mean and median trends appear quite different in high, medium, and

low population density areas. Mid-range O_3 concentrations at low population density locations within the case study areas (so still relatively close to a major city) have generally decreased over a period of substantial NOx emissions reductions[10]. This decrease is most pronounced in the summer months and in the Eastern half of the U.S. (low population density monitors in 3 out of the 5 Western[11] case-study areas and in only 4 out of the 10 Eastern[12] case study areas do not have significant decreases in summertime median O_3 concentrations). Mid-range O_3 concentrations in many, but not all, high population density areas have significantly increased in winter months. Wintertime increases were significant in 11 of the 15 areas (only Atlanta, Boston, Houston and Sacramento did not increase significantly). Thirteen out of 15 summertime high population density area trends in median O_3 were not significant[13], but combining winter and summer measurements to determine annual trends showed that Denver, Los Angeles, New York and Philadelphia high population density sites had significantly increasing annual median O_3 while Boston, Chicago, Dallas and St. Louis had significantly increasing 25th percentile O_3 but no significant median trend. These results reflect increasing mid-range O_3 concentrations mainly confined to urban centers during periods of NOx reductions. One important point to note is that the design value monitor (the monitor with the highest average (over three years) of 4th highest daily maximum value) in most of the case-study locations is located outside of the high population density area (as defined here). Downward trends in medium and low population density areas are therefore generally representative of the behavior at the highest O_3 monitor in an area, whereas trends in urban centers may be important from an exposure perspective.

In summary, any increasing O_3 trends occur more in highly populated areas, during cool months, and at the lower end of the O_3 distribution. Conversely, any decreasing O_3 trends occur more during warm months, in lower population areas, and at the upper end of the O_3 distribution. One result of these two phenomena is a narrowing of the range of O_3 concentrations over this period of decreasing NOx emissions. For instance, there are many cases where the top and bottom of a single distribution exhibit different trends. For example, the low population density monitors of Dallas, Los Angeles, Philadelphia and Saint Louis and the high population density monitors for Baltimore, Dallas, and Philadelphia for all months show a significant increase in the 5th percentile and a simultaneous significant decrease in the 95th percentile. More common is a

[10] Denver is an outlier among the case study areas with consistently increasing mid-range ozone trends across seasons and urban classifications. Denver may be subject to increasing emissions from large wildfires and oil and gas development which are not typical of other urban case study areas. In addition, Denver is particularly susceptible to influences from stratospheric intrusions and international transport due to its high altitude

[11] Western case study areas for this purpose include: Dallas, Denver, Houston, Los Angeles, and Sacramento

[12] Eastern case study areas for this purpose include: Atlanta, Baltimore, Boston, Chicago, Cleveland, Detroit, New York, Philadelphia, St. Louis, and Washington D.C.

[13] Only Houston and Dallas had statistically significant trends in median summertime urban ozone

significant change in one end of the distribution, but no significant change in the other (e.g., the summer months at high population density monitors in all case study areas except Baltimore, Chicago, and Detroit). It is important to note that there are also cases where both ends of the distribution change in the same manner and there is therefore no narrowing of the range of O_3 concentrations in these areas.

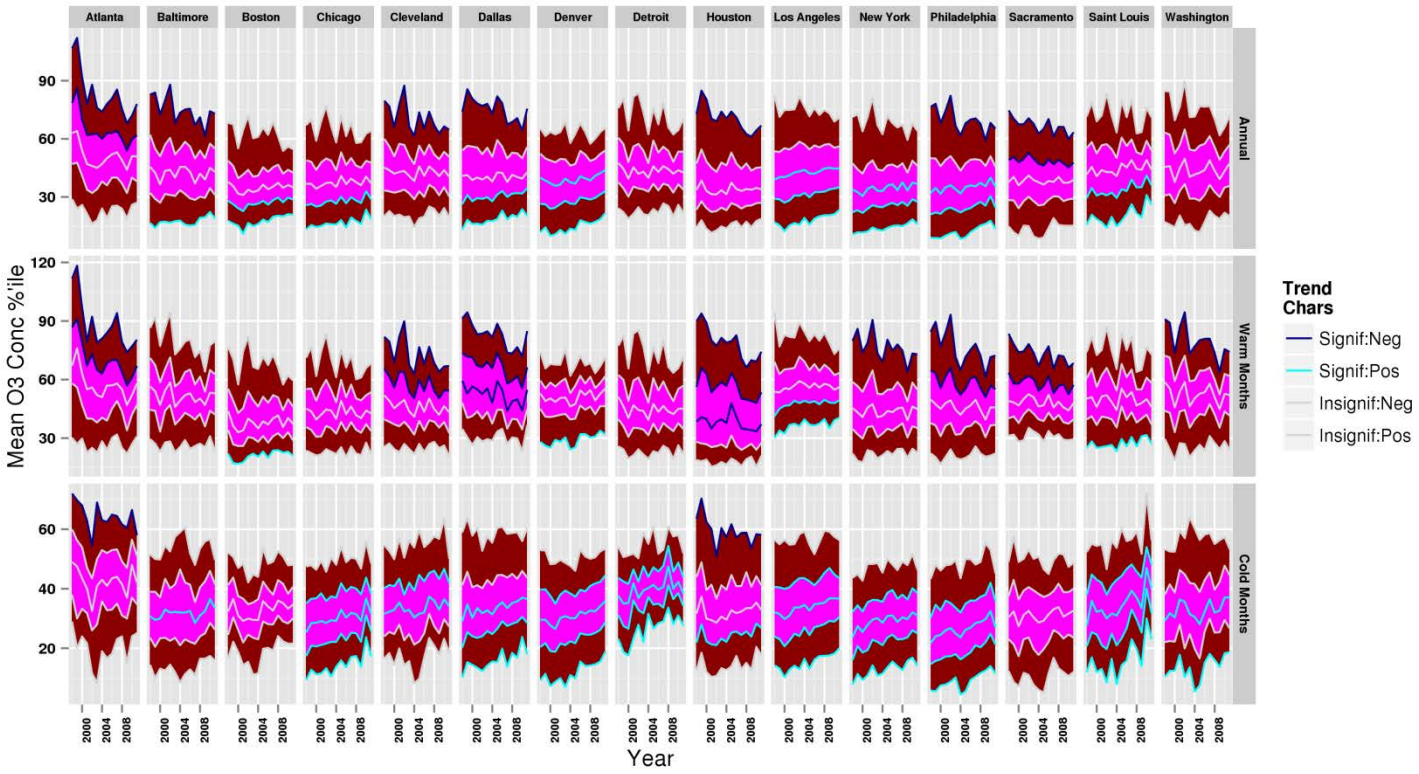


Figure 8.31 Distributions of O_3 concentrations for high population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011.

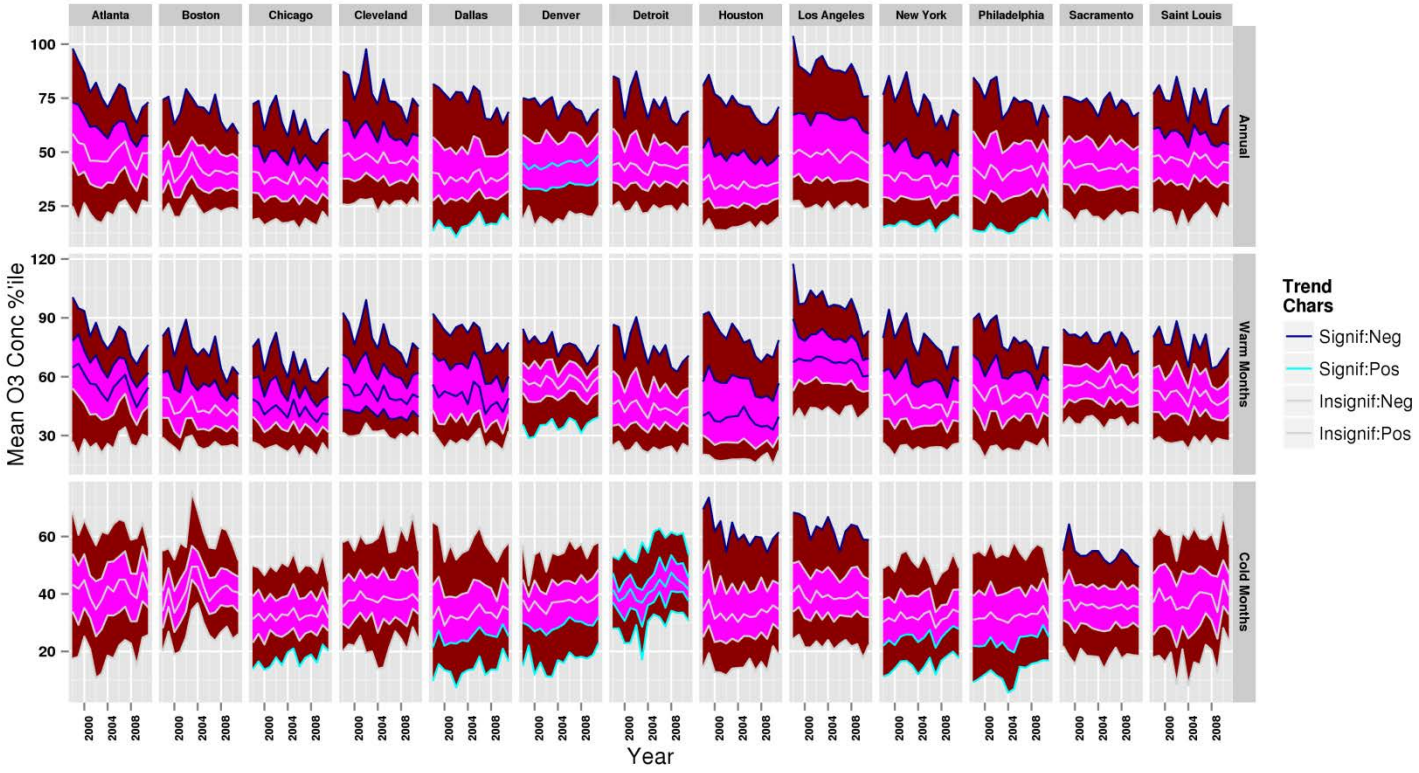


Figure 8.32 Distributions of O_3 concentrations for medium population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011.

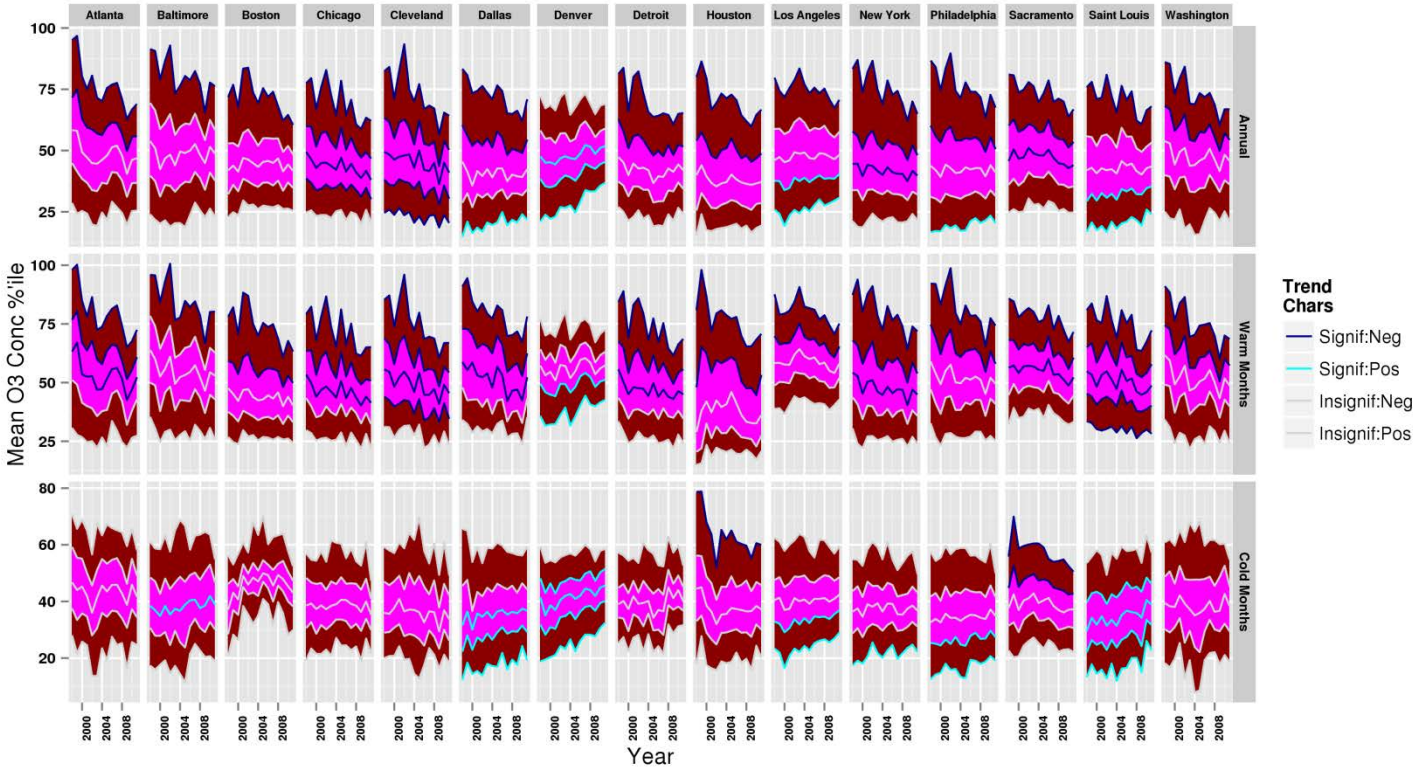


Figure 8.33 Distributions of O_3 concentrations for low population density monitors by different subsets of months over a 13-year period. From top to bottom in each ribbon plot, the blue and white lines indicate the spatial mean of the 95th, 75th, 50th, 25th, and 5th percentiles for each monitor for every year from 1998-2011.

Maps of ambient trends in both New York City and Chicago most clearly show these trends and further illustrate this behavior. Figures 8.34 and 8.35 show trends in daily maximum 8-hour O_3 values these two cities for May-September. Plots for other case-study areas are provided in Appendix 8-C. For both cities, the fourth highest 8-hr daily maximum O_3 value either has a downward trend or no trend at all monitors. In New York (Figure 8.34), mean and median O_3 values significantly decrease at downwind locations in New York and Connecticut. Conversely, median O_3 values significantly increase from 1998 to 2011 at two core urban sites (one at City College of NY in upper Manhattan and one near Queen's college) and at a nearby site on Long Island. Similarly, in Chicago (Figure 8.35), mean and median trends in O_3 are downward or insignificant in Indiana and in suburban Illinois locations and show increases near the highly populated urban core.

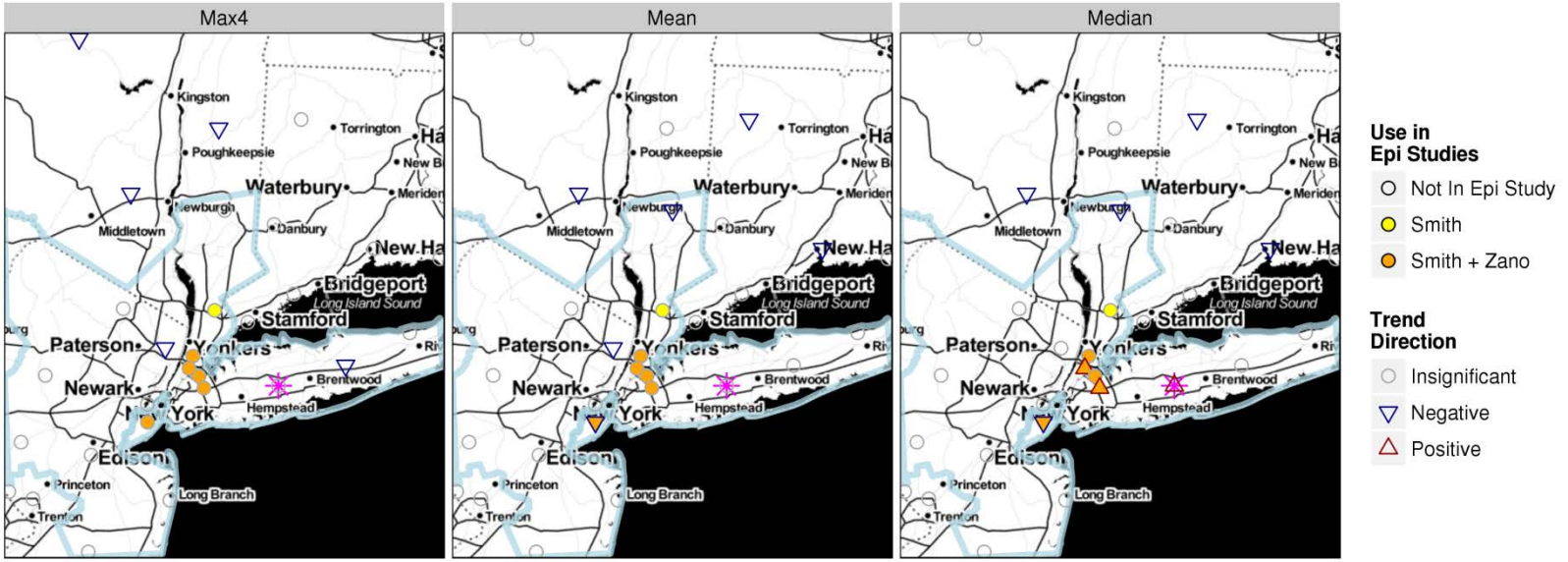


Figure 8.34 Map of O_3 trends at specific monitors in the New York area. All upward and downward facing triangles represent statistically significant trends from 1998-2011 ($p < 0.05$), circles represent locations with no significant trends. Sites used in Smith et al (2009) and the Zanobetti and Schwartz (2008) epidemiology studies are represented by colored dots. Only monitors with at least seven years of data are displayed. The pink star indicates the site with the higher design value in 2011. The MSA border as defined by the U.S. census bureau is delineated by the light blue line. Left panel shows trends in annual 4th highest 8-hr daily maximum O_3 values, center panel shows trends in annual mean 8-hr daily maximum O_3 values, and right panel shows trends in annual median 8-hr daily maximum O_3 values.

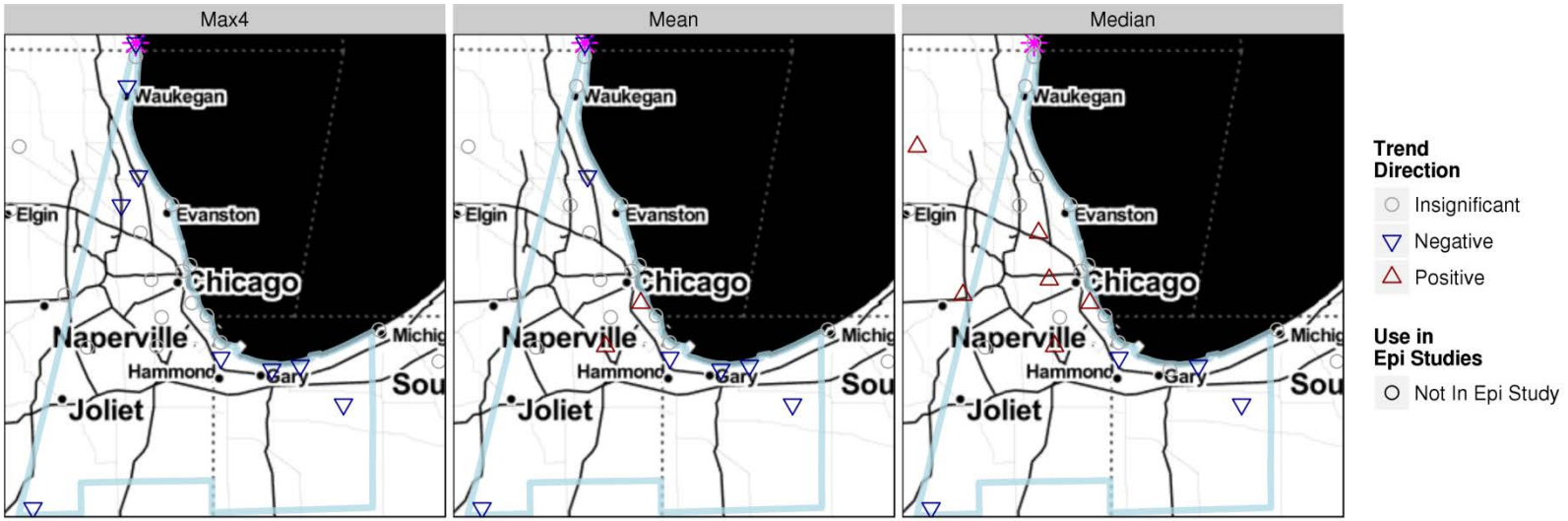


Figure 8.35 **Map of O_3 trends at specific monitors in the Chicago area. All upward and downward facing triangles represent statistically significant trends from 1998-2011 ($p < 0.05$), circles represent locations with no significant trends. Sites used in Smith et al (2009) and the Zanobetti and Schwartz (2008) epidemiology studies are represented by colored dots. Only monitors with at least seven years of data are displayed. The pink star indicates the site with the higher design value in 2011. The MSA border as defined by the U.S. census bureau is delineated by the light blue line. Left panel shows trends in annual 4th highest 8-hr daily maximum O_3 values, center panel shows trends in annual mean 8-hr daily maximum O_3 values, and right panel shows trends in annual median 8-hr daily maximum O_3 values.**

To demonstrate how changes in emissions of NOx and anthropogenic VOCs might be driving these trends, Table 8-7 shows trends of O_3 in high and low population areas and annual National Emissions Inventories (NEI) for 2002, 2005, 2008 and 2011[14] aggregated to the level of the NOAA Climate Regions[15]. There is moderate correspondence between the decreases in NOx emissions across the regions with the observed decreases in warm season O_3 concentrations in low population areas. VOCs show little correspondence to any of the O_3 trends, which is likely due to complications from 1) the mix of chemicals with a large range of reactivities; 2) complex non-linear chemistry; and 3) the potential impact of the much larger magnitude of biogenic vs. anthropogenic emissions on a regional scale. Details of these calculations can be found in Appendix 8-C.

Table - 8.7 Broad Regional Annual Trends of Concurrent O_3 Concentrations and Emissions of NOx and VOCs over the 2000-2011 Time Period

Trend	Central	East North Central	North East	South	South East	South West	West
High Pop Dens, May2Sep O_3	none	none	none	down	none	none	none
High Pop Dens, Oct2Apr O_3	up	up	up	low %'s up	none	up	up
Low Pop Dens, May2Sep O_3	down	down	down	down	down	low %'s up	top %'s down
Low Pop Dens, Oct2Apr O_3	none	none	top %'s up	none	none	up	none
NO_x Emission	down	down	down	down	down	down	down
VOC Emission	none	none	down	up	down	up	down

8.2.3.2 Modeled O_3 response to emissions reductions across the United States

This section presents an analysis of the CMAQ modeling of O_3 responses to "across-the-board" U.S. anthropogenic precursor emissions reductions described in Appendix 4b. In this analysis, we compare the modeled responses of O_3 concentrations in the case study areas to the modeled O_3 responses in the rest of the U.S. For this purpose, we used five CMAQ model simulations: 1) a base simulation which included 2007 emissions of all O_3 precursors, 2) a 50% NOx cut simulation in which U.S. anthropogenic NOx emissions were reduced by 50% from 2007 levels, 3) a 90% NOx cut simulation in which U.S. anthropogenic NOx emission were reduced by 90% from 2007 levels, 4) a 50% NOx/VOC cut simulation in which both U.S. anthropogenic NOx and VOC emission were reduced by 50% from 2007 levels, and 5) a 90% NOx/VOC cut simulation. These simulations are analyzed to characterize responses in O_3 to

[14] http://www.epa.gov/ttnchie1/trends/

[15] http://www.ncdc.noaa.gov/monitoring-references/maps/us-climate-regions.php

"across the board" emissions cuts at four distinct levels and do not represent the exact adjustment cases that were used to estimate O_3 concentrations consistent with individual case study areas just meeting various potential levels of the NAAQS standard. However, these four cases generally span the range of emissions perturbations that were applied in the HDDM adjustment methodology described in Chapter 4 and in Appendix 4d.

In this analysis we focus on seasonal mean O_3 and population as proxies for epidemiology based risk estimates in Chapter 7. Since the epidemiology studies used in Chapter 7 show relatively linear response of health outcomes to O_3 concentrations throughout the entire range of measured O_3 values, examining seasonal mean values should provide some understanding of locations where O_3 health effects are expected to increase and decrease as a result of precursor emission reductions. By combining population information with these spatial distributions of seasonal O_3 responses, we can better understand expected O_3 behavior in locations where people live. This is not a detailed risk assessment but can provide information on the representativeness of the case-study areas to the nation as a whole in terms of expected O_3-related health outcomes.

To begin, we examine maps displaying ratios of mean O_3 concentrations in the emissions cut simulations to mean O_3 concentrations in the 2007 base simulation. Figure 8.36 and Figure 8.37 show the ratio of seasonal (April-October) mean O_3 in the two NOx emissions reduction simulations to that in the base simulation for the entire model domain. Figure 8.38 and Figure 8.39 depict the ratio of January mean O_3 for the two NOx cut simulations. Figures showing the ratios based on the May-September seasonal average and figures for the NOx/VOC emissions reductions scenarios are provided in Appendix 8-C. The maps show widespread decreases (i.e., ratios less than 1) in seasonal mean O_3 across the country. These decreases are especially pronounced in the Eastern U.S. and in California. O_3 increases (i.e., ratios greater than 1) are confined to urban core areas except in January. The spatial extent of these O_3 increases are generally less for the 90% NOx cut simulation than for the 50% NOx cut simulation although the magnitude is greater over very limited areas in Chicago, Seattle, and San Francisco. The O_3 increases are most widespread in the cooler months (January, April, and October). For the April-October seasonal average O_3 concentrations, VOC in addition to NOx cuts did not substantially change the ratios of O_3 in the emissions reduction scenarios to O_3 in the base scenario. In the Northeast and Midwest, increases in seasonal mean O_3 concentrations were mainly confined to urban case study areas of New York, Detroit, Chicago, and St. Louis. In the Southeastern U.S., the urban areas which show up as having increased seasonal mean O_3 in the 50% NOx cut simulations include Miami, Orlando, Tampa, and New Orleans (only Miami has O_3 increase in the 90% NOx cut simulation). The only case-study area in the southeast, Atlanta, does not experience increases in seasonal mean O_3 in the model simulation (this is consistent with the

changes in risk estimated for Atlanta, which show no increases in total risk as alternative standards are simulated). In the central U.S., seasonal mean O_3 in the case study areas of Denver, Houston, and Dallas and non-case-study areas of San Antonio, Duluth, and Minneapolis increased with 50% NOx reductions. Seasonal mean O_3 increases were seen only in Houston, Minneapolis, and Duluth with 90% reductions in simulated NOx emissions. The Northwestern U.S. showed some of the most widespread increases in seasonal mean O_3 in the 50% and 90% NOx cut simulations covering the Seattle and Portland metro areas as well the San Francisco Bay area and in a single model grid cell for Sacramento (50% NOx reduction case only). Sacramento is the only city in the Northwest that was included as a case study area. Finally, Los Angeles (a case study area), San Diego, Phoenix, and Bakersfield were the areas for which CMAQ predicted seasonal mean O_3 increases with the 50% NOx cut simulation. These O_3 increases disappeared (or were largely diminished in the case of LA) in the 90% NOx cut case. Based on these maps, it appears that in the Northeast and the Central U.S., the case-study area selection likely oversampled these O_3 increases on a geographic basis since all locations outside of city centers experienced decreasing seasonal mean O_3 with the NOx reduction model simulations. However in two regions, the Southeast and the Northwest, the urban case study area did not experience increases in seasonal mean O_3 concentrations while other urban areas in the region did. In these two regions, the urban case-study selection likely under-sampled the locations which experienced increases in seasonal mean O_3.

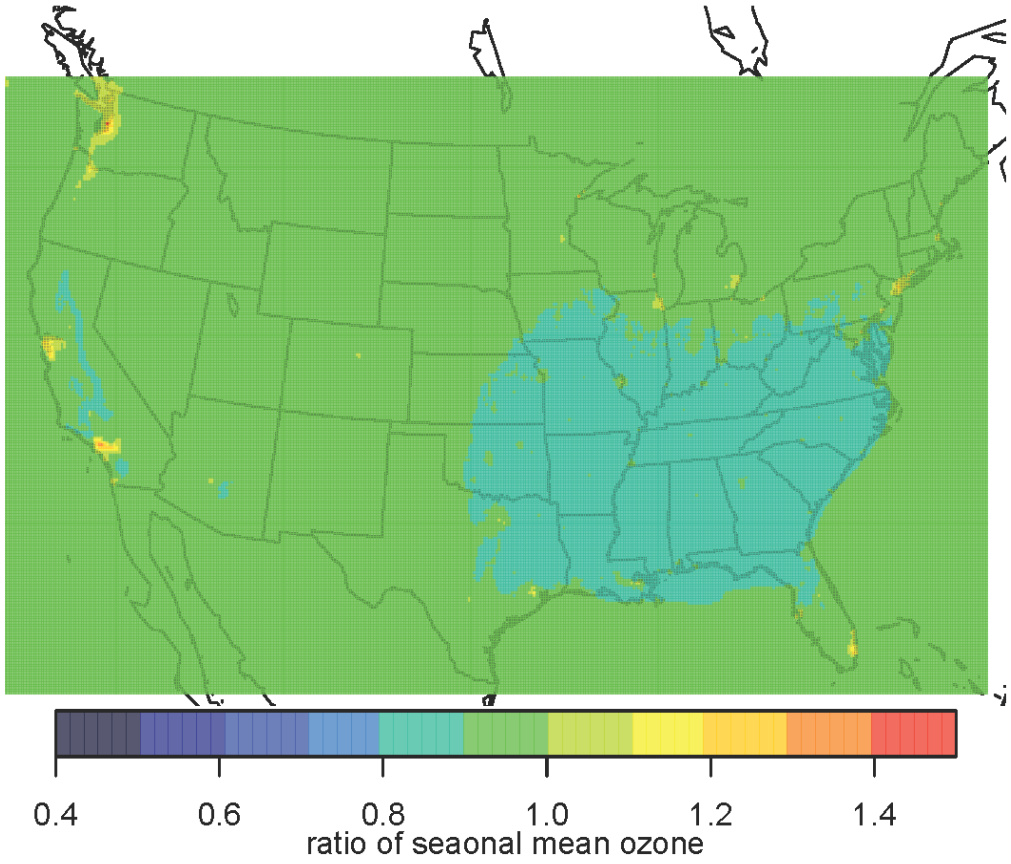


Figure 8-36 Ratio of April-October seasonal average O_3 concentrations in the brute force 50% NOx emissions reduction CMAQ simulations to April-October seasonal average O_3 concentrations in the 2007 base CMAQ simulation.

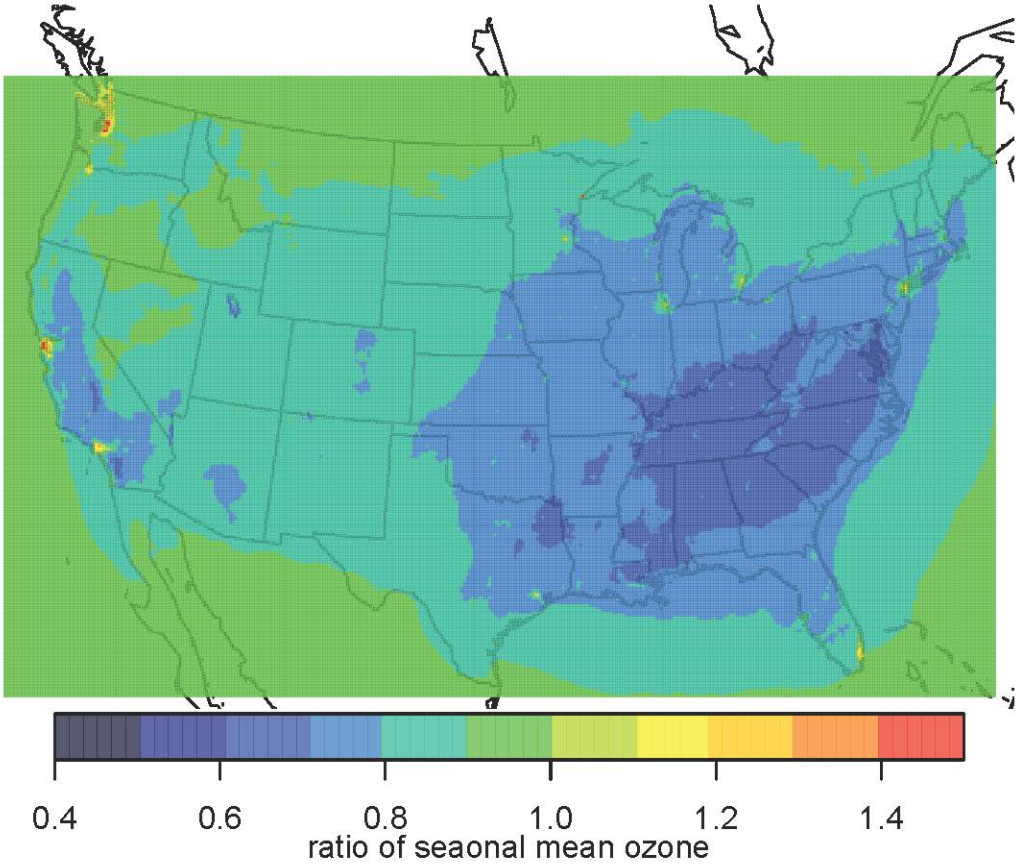


Figure 8.37 Ratio of April-October seasonal average O_3 concentrations in the brute force 90% NOx emissions reduction CMAQ simulations to April-October seasonal average O_3 concentrations in the 2007 base CMAQ simulation.

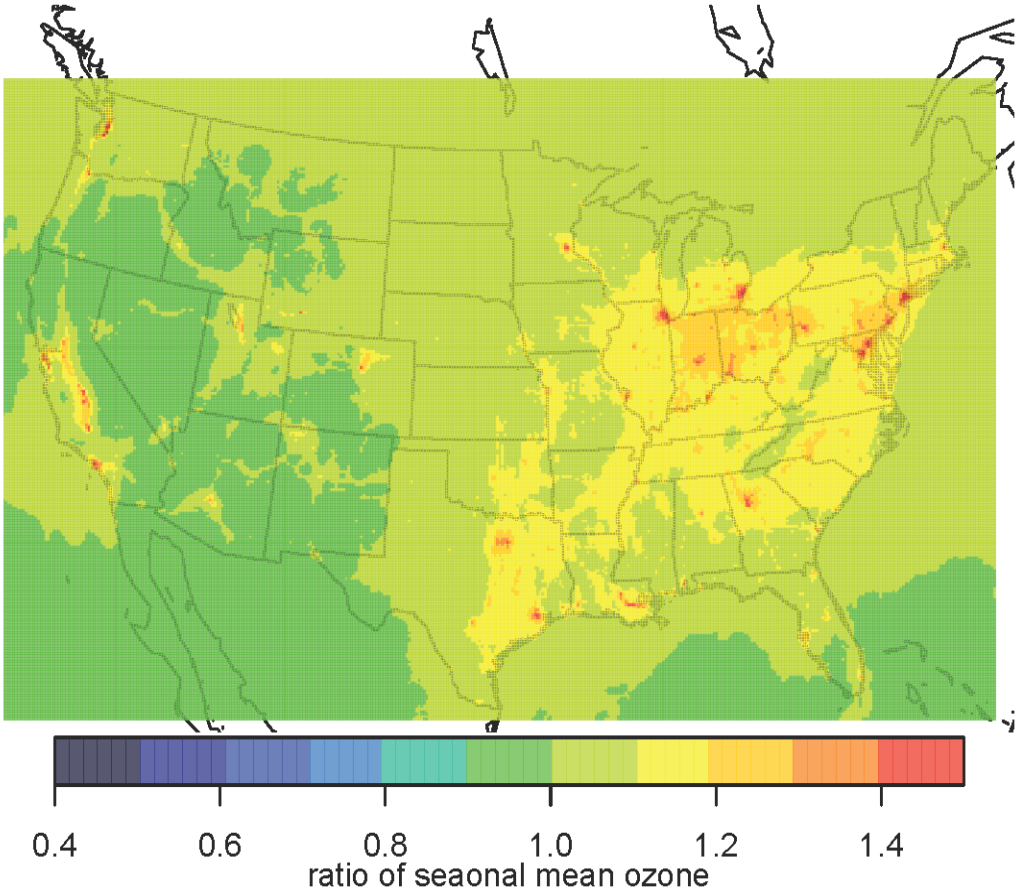


Figure 8.38 Ratio of January monthly average O_3 concentrations in brute force 50% NOx emissions reduction CMAQ simulations to January monthly average O_3 concentrations in the 2007 base CMAQ simulation.

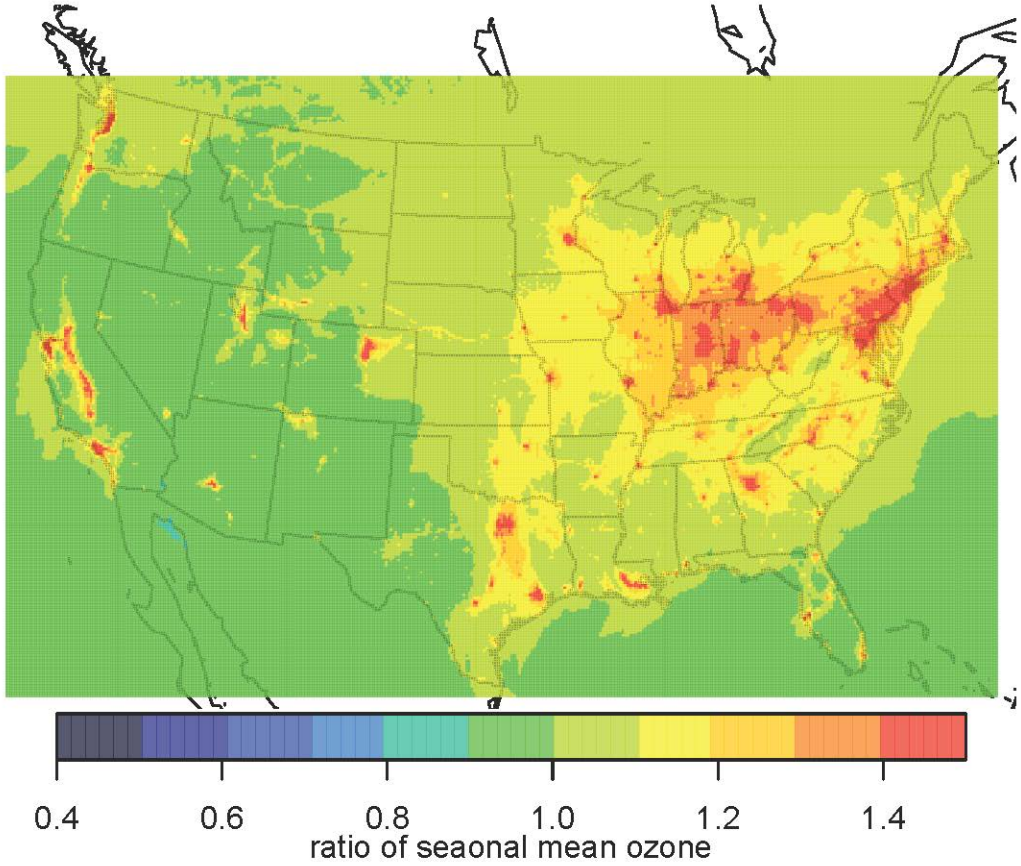


Figure 8.39 Ratio of January monthly average O_3 concentrations in brute force 90% NOx emissions reduction CMAQ simulations to January monthly average O_3 concentrations in the 2007 base CMAQ simulation.

In order to characterize the representativeness of case study areas in a more quantitative manner, paired O_3 concentrations and population data[16] were extracted from each model grid cell and categorized in various manners. Figure 8.40 and Figure 8.41 depict the percent of U.S. population living in areas with increases or decreases in monthly or seasonal mean O_3 under the emissions reductions scenarios (50% NOx cut and 90% NOx cut respectively) compared to O_3 in the base modeling scenario. The top panels show data for January monthly mean O_3, the center panels show data for seasonal mean O_3 (June-August), and the bottom panels show data for seasonal mean O_3 (April-October). Tabulated results and equivalent plots for the combined

[16] Block level population data from the 2010 Census was aggregated to the 12km CMAQ grid cell level. The 2007 population was then calculated using population growth factors developed by Woods and Poole Economics, Inc.

NOx/VOC cut simulations are provided in Appendix 8-C. Month by month break-outs for each case study area are also available in Appendix 8-C.

The vast majority of the U.S. population lives in areas where the CMAQ simulations predict mean O_3 decreases for the June-August and April-October time periods. The majority of population living in case-study areas also lives in locations with decreasing seasonal mean O_3 concentration under NOx reduction scenarios. As discussed previously, more locations have increasing mean O_3 in the cooler months as demonstrated by the fact that almost all of the U.S. population lives in locations where the model predicts increases in mean O_3 in January. The case study areas represent 29% of the total U.S. population. These areas account for 20-30% of the U.S. population that experience decreasing seasonal mean O_3 for April-October in the NOx cut simulations and 50-60% of the U.S. population that experience increasing seasonal O_3 for April-October. Consequently, the urban-case study areas over-sample populations living in locations with increasing seasonal mean O_3 in response to NOx cuts compared to populations living in locations with decreases in seasonal mean O_3. In all panels displayed in Figures 8.28 and 8.29, most of the population lives in locations where increases or decreases in mean O_3 were greater than 1 ppb.

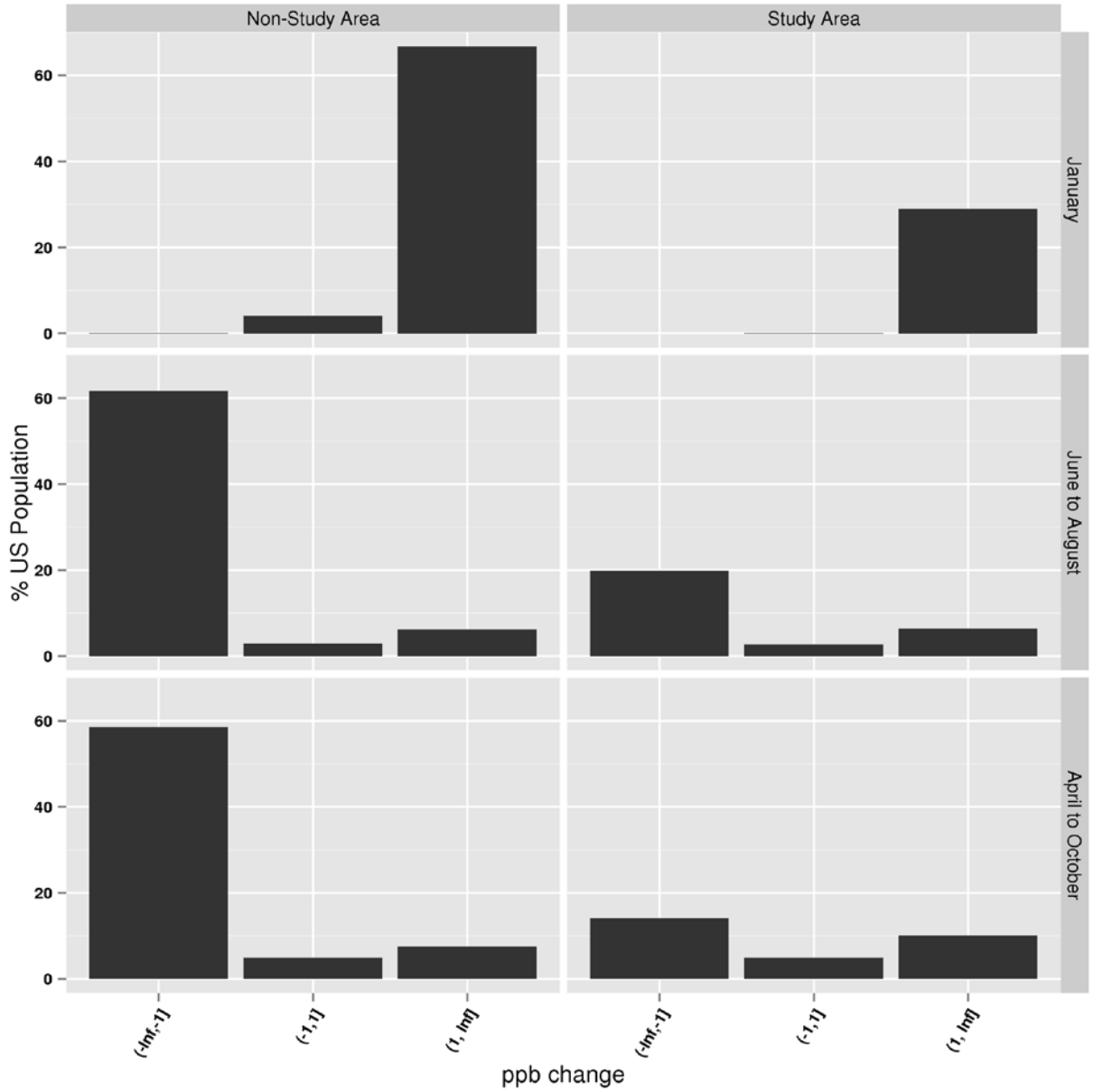


Figure 8.40 Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent change in mean O_3 (ppb) from the 2007 base CMAQ simulation to the 50% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the case study areas. Top plots show changes in January monthly mean O_3, middle plots show changes in seasonal mean June-August O_3, and bottom plots show changes in seasonal mean April-October O_3.

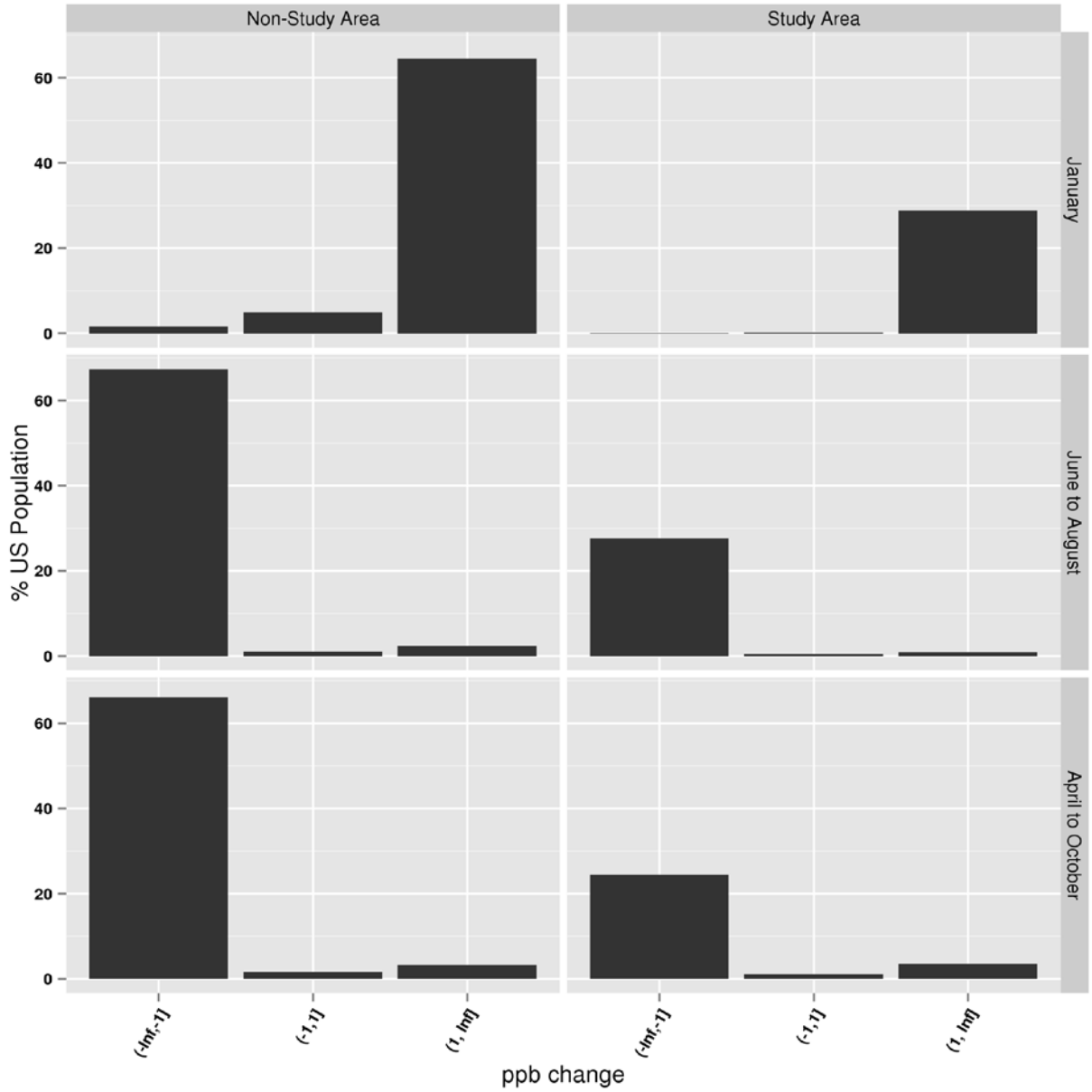


Figure 8.41 Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent change in mean O_3 (ppb) from the 2007 base CMAQ simulation to the 90% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the case study areas. Top plots show changes in January monthly mean O_3, middle plots show changes in seasonal mean June-August O_3, and bottom plots show changes in seasonal mean April-October O_3.

The proportion of the population living in locations of increasing seasonal mean O_3 in response to NOx emissions reductions varies considerably between case study areas. Figure 8.42 and Figure 8.43 show these proportions by city for the 50% and 90% NOx reduction scenarios. For the 50% NOx reduction scenario, the CMAQ results predict that four out of fifteen study ares (Chicago, Detroit, Los Angeles, and New York) have more than 50% of their populations living in locations with increasing mean O_3 for April-October. Most other urban case study areas have between 5% and 30% of their populations living in these areas with increasing mean O_3 levels. For the 90% NOx reduction scenario, the percent of population living in such locations decreases substantially for all cities, leaving four out of fifteen study areas (Detroit, Houston, Los Angeles, and New York) with more than 5% of their populations living in areas with increasing mean O_3 levels.

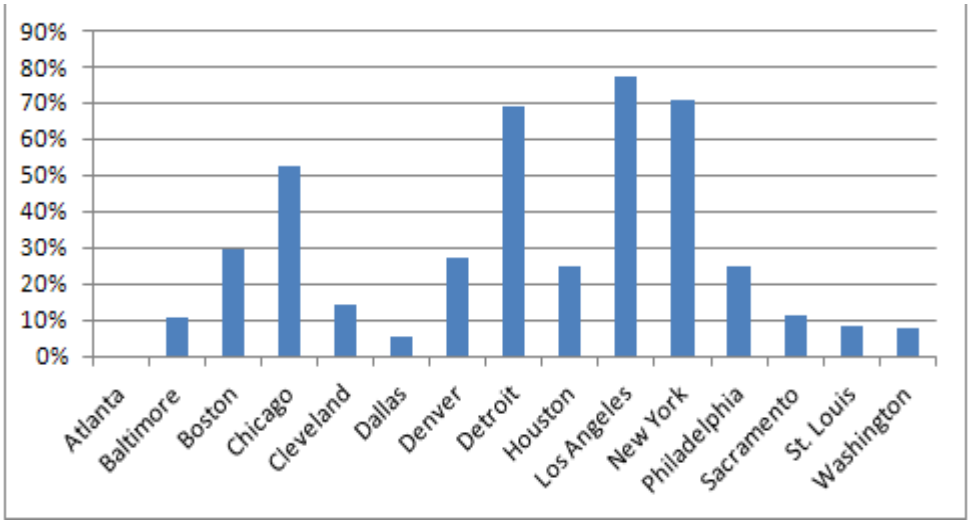


Figure 8.42 Population (as % of total case-study area population) living in locations of increasing April-October seasonal mean O_3 in the 50% NOx reduction CMAQ simulation.

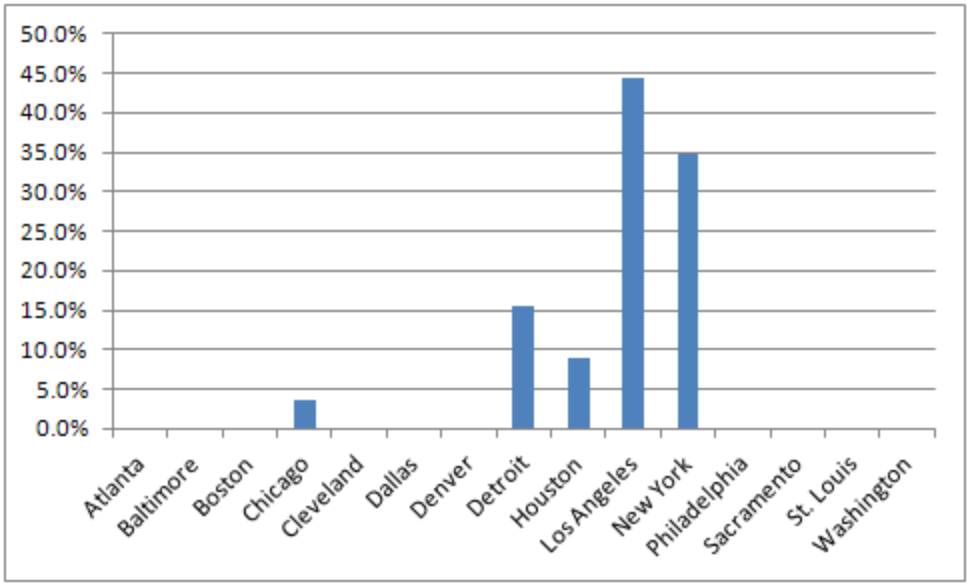


Figure 8.43 Population (as % of total case-study area population) living in locations of increasing April-October seasonal mean O_3 in the 90% NOx reduction CMAQ simulation.

We can further understand these results by looking at them in terms of population density in the case-study areas versus across the U.S. as a whole. As in Section 8.2.3.1, we define census tracts with population density greater than 1000 people/km^2 as high population density, but the low-mid population density classification used here is a combination of the low and medium classifications in that section. Figure 8.44 and Figure 8.45 split out the April-October results from Figure 8.40 and Figure 8.41 into high and low-mid sub-categories. Appendix 8-C provides similar breakouts for the other panels in Figure 8.40 and Figure 8.41. First, based on these definitions, we see that 57% of the population in case-study areas lives in high population density locations while only 27% of the U.S. population does. As discussed above, the high population areas are more likely to experience increases in mean O_3 as a result of NOx emission reductions compared to lower population areas. Therefore, the fact that the case-study areas used in the risk and exposure assessments are more densely populated than the country as a whole means that these analyses may estimate higher risks under emissions reduction scenarios than would be experienced, on average, across the country. Figure 8.44 and Figure 8.45 show generally similar shapes for the high population density histograms in the study-area and non-study area locations. In the 50% NOx cut simulation, 69% of the population living in high density case-study areas would experience increases in mean seasonal O_3 compared to 63% of the population the population living in high density areas of the country as a whole. Similarly in the 90% NOx cut simulation, 28% of the population in high density locations both within the study areas and across the U.S. as a whole lives in locations of increasing seasonal mean O_3. This

suggests that the selected study areas adequately represent population-weighted changes in mean O_3 for people living in high density areas. Similarly, less densely populated locations within the case-study areas show O_3 increases equivalent to those seen in less densely populated areas in the U.S. as a whole. In the 50% NOx cut simulation, 7% of people in low-mid density study area locations live where mean seasonal O_3 is increasing, while 5% of people in all low-mid density U.S. locations live where mean seasonal O_3 is increasing. Similarly, in the 90% NOx cut simulation, the numbers are 2% for both low-mid density study area populations and for low-mid density populations in the U.S. as a whole. Thus the oversampling of populations living in locations of increasing mean seasonal O_3 in response to NOx cuts, as shown in Figures 8.28 and 8.30, appears to be entirely due to the fact that the study areas oversample populations living in high density areas compared to the U.S. population as a whole.

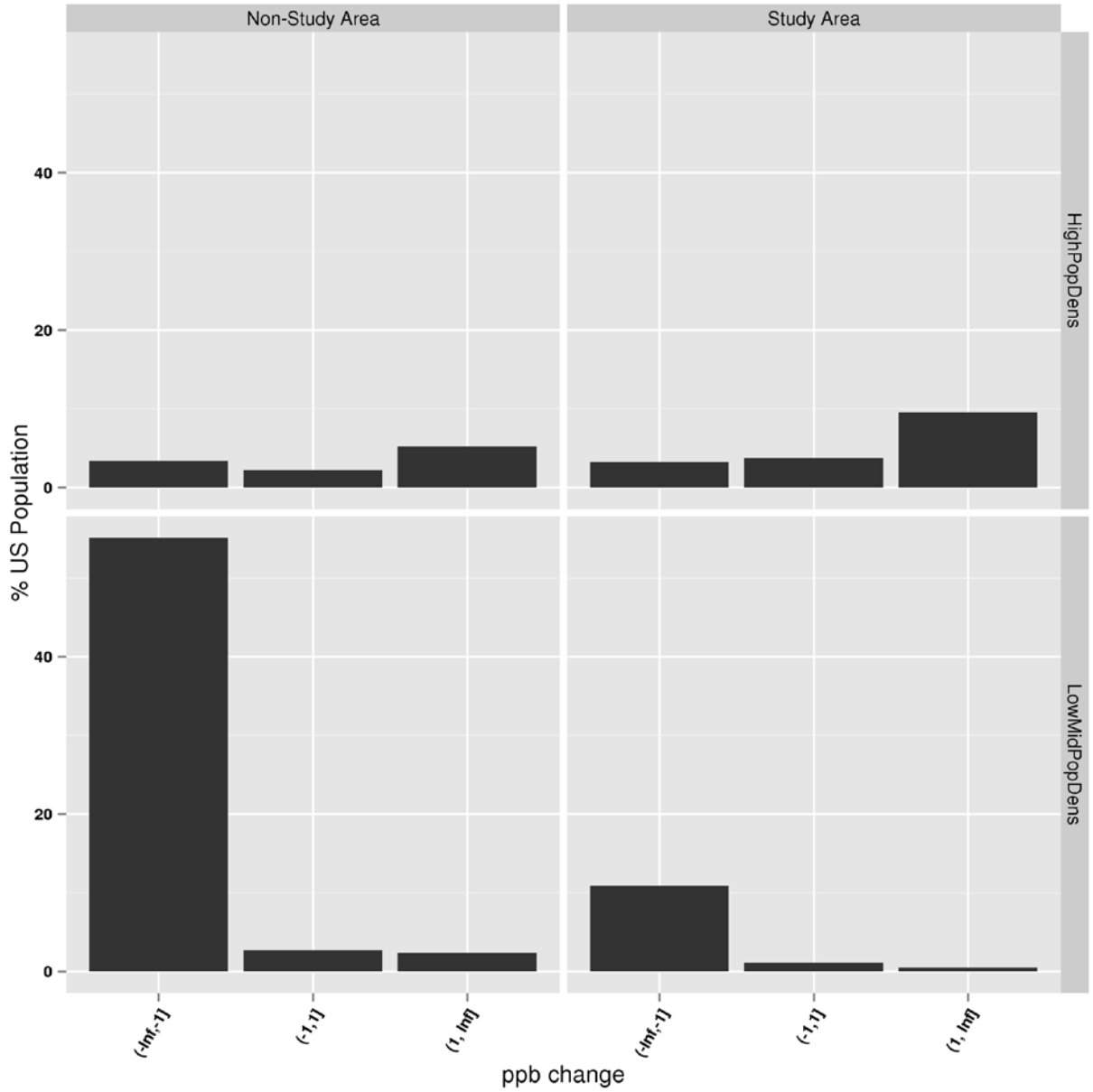


Figure 8.44 Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent the change in seasonal mean (April-October) O_3 from the 2007 base CMAQ simulation to the 50% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the urban case study areas. Bottom plots show histograms for low-mid population density areas while top plots show histograms for high population density areas.

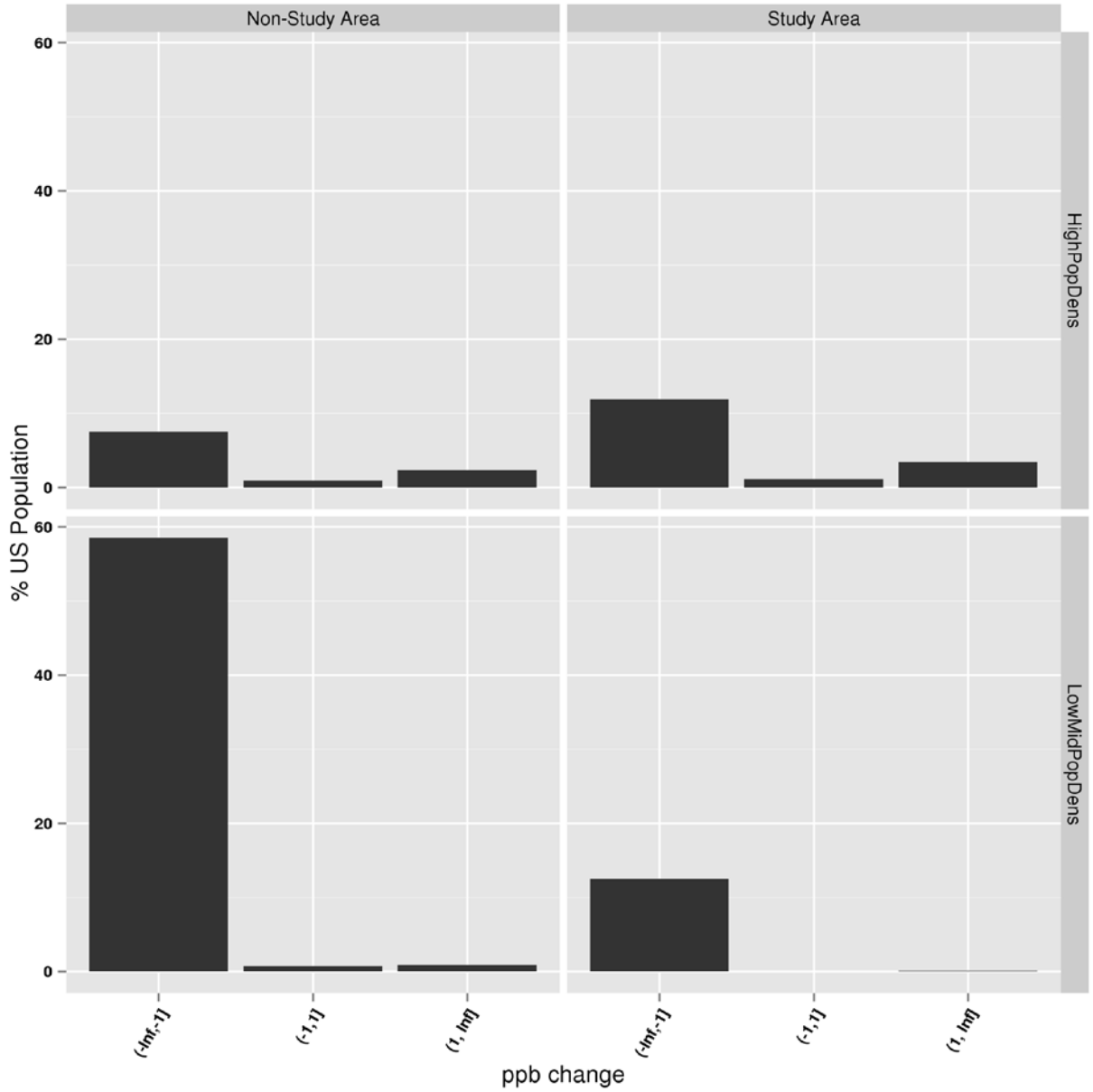


Figure 8.45 Histograms of U.S. population living in locations with increasing and decreasing mean O_3. Values on the x-axis represent the change in seasonal mean (April-October) O_3 from the 2007 base CMAQ simulation to the 90% NOx cut CMAQ simulation. The percentages of the U.S. population living in areas that have changes less than -1 ppb, from -1 to +1 ppb, and greater than 1 ppb are shown on the y-axis. Left plots show population numbers in locations not included in one of the cases study areas while right plots show population numbers in locations included in one of the urban case study areas. Bottom plots show histograms for low-mid population density areas while top plots show histograms for high population density areas.

8.2.5 Discussion

We evaluated two different questions, 1) to what degree are the 15 cities evaluated in the exposure and risk analyses representative of the overall U.S. population with regards to total O_3 risk?, and 2) to what degree are they representative of the overall U.S. population with regards to the degree of risk reduction that might be observed in response to just meeting the existing and alternative standards.

Regarding the first question, we observe that the 12 urban study areas considered in the urban-scale risk assessment presented in Section 7.2 capture urban areas that are among the most populated in the U.S., have relatively high O_3 levels, and represent the range of city-specific effect estimates found by Smith et al. (2009) and Zanobetti and Schwartz (2008). These three factors suggest that the urban study areas capture overall risk for the nation well, with a potential for better characterization of the high end of the risk distribution. We find that the urban study areas are not capturing areas with the highest baseline mortality rates, those with the oldest populations, and those with the lowest air conditioning prevalence. These areas tend to have relatively low O_3 concentrations and low total population, suggesting that the urban study areas are not missing high risk populations that have high O_3 concentrations in addition to greater susceptibility per unit O_3. We also find that the 12 urban study areas represent the full range of county-level O_3-related risk across the entire U.S. We conclude from these analyses that the 12 urban study areas adequately represent O_3-related risk across the U.S.

Concerning the second question, we observe that the 15 urban areas considered in the exposure and risk assessment case study areas over-sample populations living in locations with increasing seasonal mean O_3 in response to NOx cuts. This suggests that the selected study areas adequately represent population-weighted changes in mean O_3 concentrations for urban populations, but may be under-representing decreasing median O_3 concentrations in suburban and rural areas. As a result, the risk estimates for populations in the selected urban study areas may understate the risk reductions that might be achieved across the broader U.S. population.

8.3 REFERENCES

Abt Associates, Inc. 2010. *Model Attainment Test Software (Version 2)*, prepared for EPA. Research Triangle Park, NC: EPA Office of Air and Radiation, Office of Air Quality Planning and Standards. Research Triangle Park, NC. Available at: <http://www.epa.gov/scram001/modelingapps.mats.htm>.

Abt Associates, Inc. 2010. *Environmental Benefits and Mapping Program (Version 4.0)*, prepared for EPA. Research Triangle Park, NC: EPA Office of Air Quality Planning and Standards. Available at <http://www.epa.gov/air/benmap>.

- Anenberg, S.C.; J.J. West; L.W. Horowitz and D.Q. Tong. 2010. "An Estimate of the Global Burden of Anthropogenic Ozone and Fine Particulate Matter on Premature Human Mortality Using Atmospheric Modeling." *Environmental Health Perspective*, 118:1189-1195.

Anenberg, S.C.; J.J. West; L.W. Horowitz and D.Q. Tong. 2011. "The Global Burden of Air Pollution Mortality: Anenberg et al. respond." *Environmental Health Perspective*, 119:A158-A425.

Bell, M.L.; A. McDermott; S.L. Zeger; J.M. Samet and F. Dominici. 2004. "Ozone and Short-term Mortality in 95 U.S. Urban Communities, 1987-2000." *Journal of the American Medical Association*, 292:2372-2378.

Byun, D. and K. L. Schere. 2006. "Review of Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multi-scale Air Quality (CMAQ) Modeling System." *Applied Mechanics Reviews*, 59:51-77.

Centers for Disease Control: Wide-ranging Online Data for Epidemiological Research (CDC-Wonder) (data from years 2004-2006), Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services. Available on the Internet at <http://wonder.cdc.gov>.

Cooper, O.R.; R. Gao; D. Tarasick; T. Leblanc and C. Sweeney. 2012. Long-term Ozone Trends at Rural Ozone Monitoring Sites Across the United States, 1990-2010." *Journal of Geophysical Research*, 117, D22, doi:10.1029/2012JD018261.

Fann N.; A. D. Lamson; S. C. Anenberg; K. Wesson; D. Risley; B. J. Hubbell. 2012. "Estimating the National Public Health Burden Associated with Exposure to Ambient $PM_{2.5}$ and Ozone." *Risk Analysis*, 32:81-95.

George, B.J. and T. McCurdy. 2009. "Investigating the American Time Use Survey from an exposure modeling perspective." *Journal of Exposure Science and Environmental Epidemiology*, 21:92-105.

Graham, S. and T. McCurdy. 2004. "Developing Meaningful Cohorts for Human Exposure Models." *Journal of Exposure Analysis and Environmental Epidemiology*, 14:23-43.

Hollman, F.W.; T. J. Mulder and J. E. Kallan. 2000. "Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100." Population Division Working Paper No. 38, Population Projections Branch, Population Division, U.S. Census Bureau, Department of Commerce.

Jerrett, M.; R.T. Burnett; C. A. Pope, III; K. Ito; G. Thurston; D. Krewski; Y. Shi; E. Calle; M. Thun. 2009. "Long-term O_3 Exposure and Mortality." *New England Journal of Medicine*, 360:1085-1095.

Neidell, M. 2009. Information, avoidance behavior and health. J Human Res. 44:450-478.

Neidell, M. 2010. Air quality warnings and outdoor activities: evidence from Southern California using a regression discontinuity approach design. J Epidemiol Community Health. 64:921-926.

Samet, J. M.; S. L. Zeger; F. Dominici; F. Curriero; I. Coursac; D.W. Dockery; J. Schwartz and A. Zanobetti. (2000). "The National Morbidity, Mortality, and Air Pollution Study Part II: Morbidity and Mortality from Air Pollution in the United States." Health Effects Institute, Boston, MA, Number 94, Part II.

Smith R.L.; B. Xu and P. Switzer. 2009. "Reassessing the Relationship Between Ozone and Short-term Mortality in U.S. Urban Communities." *Inhalation Toxicology*, 21:37-61.

Timin, B.; K. Wesson and J. Thurman. "Application of Model and Ambient Data Fusion Techniques to Predict Current and Future Year $PM_{2.5}$ Concentrations in Unmonitored Areas. (2010)," pp. 175-179 in Steyn DG, Rao St (eds). *Air Pollution Modeling and Its Application XX*. Netherlands: Springer.

U.S. Environmental Protection Agency. 2013. *Integrated Science Assessment for Ozone and Related Photochemical Oxidants*. Research Triangle Park, NC: EPA. (EPA document number EPA 600/R-10/076F).

U.S. EPA. 2010. *Quantitative Health Risk Assessment for Particulate Matter*. Research Triangle Park, NC: EPA. (EPA document number EPA-452/R-10-005).

U.S. EPA. 2009. *Risk and Exposure Assessment to Support the Review of the SO2 Primary National Ambient Air Quality Standards: Final Report*. Research Triangle Park, NC: U. S. EPA. (EPA document number EPA-452/R-09-007).

U.S. EPA. 2007. *Ozone Population Exposure Analysis for Selected Urban Areas*. Research Triangle Park, NC: EPA. (EPA document number EPA-452/R-07-010). Available at: <http://www.epa.gov/ttn/naaqs/standards/ozone/data/2007_07_ozone_exposure_tsd.pdf>.

Woods and Poole Economics Inc. 2012. "Complete Demographic Database." Woods and Poole Economics, Inc. Washington, DC. Available at: <http://www.woodsandpoole.com/index.php.>.

Zanobetti, A. and J. Schwartz. 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*, 177:184-189.

Zanobetti, A. and J. Schwartz. 2011. "Ozone and Survival in Four Cohorts with Potentially Predisposing Diseases." *American Journal of Respiratory and Critical Care Medicine*, 194:836-841.

Zhang, L.; D. J. Jacob; N.V. Smith-Downey; D. A. Wood; D. Blewitt; C. C. Carouge; A. van Donkelaar; D. B. A. Jones; L.T. Murray and Y. Wang. 2011. "Improved Estimate of the Policy-relevant Background Ozone in the United States Using the GEOS-Chem Global Model with 1/2°x2/3° Horizontal Resolution Over North America." *Atmospheric Environment*, 45:6769-6776.

9 SYNTHESIS

9.1 INTRODUCTION

This assessment estimates exposures to O_3 and resulting mortality and morbidity health risks based on the findings of the O_3 ISA (U.S. EPA, 2013) that short-term O_3 exposures are causally related to respiratory effects, and likely causally related to cardiovascular effects, and that long term O_3 exposures are likely causally related to respiratory effects. The assessment evaluates total exposures and risks associated with the full range of observed O_3 concentrations, as well as the incremental changes in exposures and risks between just meeting the existing standard of 75 ppb and just meeting alternative standard levels of 70, 65, and 60 ppb using the form and averaging time of the existing standard: the annual 4th highest daily maximum 8-hour O_3 concentration, averaged over three consecutive years. We evaluated alternative standard levels of 70, 65, and 60 consistent with recommendations from CASAC to consider alternative standard levels between 60 and 70 ppb (Frey and Samet, 2012).

Following the conceptual framework described in Chapter 2, the assessment evaluates exposures and lung function risk in 15 urban case study areas, and mortality and morbidity risks based on concentration-response functions derived from epidemiology studies in 12 of these urban case study areas[1]. The results from these assessments will help inform consideration of the adequacy of the existing primary O_3 standards, and potential risk reductions associated with several alternative levels of the standard (for the current form and averaging time). In addition, to place the urban case study area analyses in a broader context, Chapter 8 of this assessment estimates the national burden of mortality associated with recent O_3 levels, and evaluates the representativeness of the 15 urban case study areas in characterizing O_3 exposures and risks across the U.S. This synthesis focuses on the urban case study area assessments of exposure and risk for the scenarios of just meeting the existing and alternative standards. For this synthesis, we discuss the results of the national-scale assessment as they relate to understanding the breadth of O_3 risks across the U.S. and to the national representativeness of the urban case study area risk results.

To facilitate interpretation of the results of the exposure and risk assessment, this chapter provides a synthesis of the various results, focusing on comparing and contrasting those results to identify common patterns, or important differences. These comparisons will focus on patterns

[1] Three additional urban case study areas were evaluated for the human exposure assessment and lung function risk assessment to provide greater geographic representation. There was insufficient information available to conduct the epidemiology-based risk assessment in these 3 additional areas. Also, we originally planned to include Seattle, WA as a 16th urban case study area, but due to limitations in the available air quality monitoring data, we determined that it would not be appropriate to model exposure and risks for Seattle (see appendix 4-E).

across urban case study areas, across years of analysis, and across alternative standards. In addition, factors related to each specific type of analysis that may influence comparisons between the analyses are identified and discussed. The degree to which the integrated results are representative of national patterns of exposure and risk is evaluated. Overall confidence in the results, as well as relative confidence between the different analyses is also assessed. The chapter concludes with an overall integrated characterization of exposure and risk in the context of key policy-relevant questions raised in Chapter 2.

9.2 SUMMARY OF KEY RESULTS

9.2.1 Air Quality Considerations (Chapter 4)

Table 9-1 below gives information on the monitoring network, population, and observed peak O_3 concentrations for the 15 case study areas, for the years included in the exposure, lung function risk, and epidemiology based risk assessments. The number of counties, number of O_3 monitors, population, and design values (DV) are based on the area definitions used in the exposure modeling and clinical-based lung function risk assessments, while the 2007 and 2009 annual 4th highest values are based on the Core Based Statistical Areas (CBSAs) used in the epidemiology-based risk assessment. The "N/A" values in the 2007 and 2009 4th high columns are for the three urban areas not included in the epidemiology-based risk assessment. The data show a trend of lower peak O_3 concentrations (i.e., the 2008-2010 design values and 2009 4th high values are generally lower than the 2006-2008 design values and 2007 4th high values, respectively).

Table 9-1 Area and Monitoring Information for the 15 Case Study Areas

Area Name	# of Counties	# of O_3 Monitors	Population (2010)	2006-2008 DV (ppb)	2007 4th high (ppb)	2008-2010 DV (ppb)	2009 4th high (ppb)
Atlanta	33	13	5,618,431	95	102	80	77
Baltimore	7	7	2,710,489	91	92	89	83
Boston	10	14	5,723,468	83	89	77	75
Chicago	16	26	9,686,021	78	N/A	74	N/A
Cleveland	8	13	2,881,937	82	83	77	72
Dallas	11	20	6,366,542	89	N/A	86	N/A
Denver	13	26	3,390,504	86	97	77	79
Detroit	9	12	5,218,852	81	93	75	73
Houston	10	22	5,946,800	91	90	85	91
Los Angeles	5	54	17,877,006	119	105	112	108
New York	27	31	21,056,173	90	94	84	81
Philadelphia	15	19	7,070,622	92	102	83	74
Sacramento	7	26	2,755,972	102	93	102	96
St. Louis	17	17	2,837,592	85	94	77	74
Washington	26	22	5,838,518	87	N/A	81	N/A

In this analysis, we employed a photochemical model-based adjustment methodology Simon et al. (2012) using the Higher-Order Decoupled Direct Method (HDDM) capabilities in the Community Multi-scale Air Quality Model (CMAQ) (hereafter referred to as HDDM air quality adjustment). The HDDM air quality adjustment methodology replaced the quadratic rollback technique used in the first draft REA to estimate O_3 concentrations consistent with just meeting existing and alternative O_3 standards. The HDDM air quality adjustment procedure estimates the change in observed hourly O_3 concentrations at a given set of monitoring sites resulting from national across-the-board reductions in U.S. anthropogenic NOx and/or VOC emissions. In this analysis, we adjusted O_3 concentrations to just meet the existing standard of 75 ppb[2] and potential alternative standards of 70, 65, and 60 ppb at ambient monitoring sites in the 15 case study areas for the 2006-2008 and 2008-2010 periods. In most locations, only NOx reductions were used to adjust the distribution of O_3 concentrations, because of the ineffectiveness of VOC reductions in reducing peak O_3 concentrations needed to meet the

[2] Attainment with the existing standard level of 75ppb is determined by the 4th highest maximum 8-hour O_3 concentration, averaged over 3 years (hereafter referred to as the existing standard).

existing and alternative standard levels. Sensitivity analyses were also conducted in some locations to evaluate the impact of decreasing both NOx and VOC emissions.

The HDDM air quality adjustment methodology represents a substantial improvement over the quadratic rollback method used to adjust O_3 concentrations in previous reviews. First, quadratic rollback was a purely mathematical technique which attempted to reproduce the distribution of observed O_3 concentrations just meeting various standards, while the new methodology uses photochemical modeling to simulate the response in O_3 concentrations due to changes in precursor emissions based on current understanding of atmospheric chemistry and transport. Second, quadratic rollback used the same mathematical formula to adjust concentrations at all monitors within each case study area for all hours, while HDDM allows the adjustments to vary both spatially across each case study area and temporally across hours of the day and across seasons. Finally, quadratic rollback was designed to only allow decreases in O_3 concentrations, while the HDDM air quality adjustment allows both increases and decreases in O_3 concentrations in response to reductions in NOx or VOC emissions. For example, in response to reductions in NOx emissions, the HDDM methodology is able to capture increases in O_3 concentrations that can occur in urban cores characterized by titration of O_3 by fresh NO emissions and decreases in O_3 concentrations downwind.

Following HDDM adjustment of O_3 concentrations, several general trends are evident in the changes in O_3 patterns across the case study areas and across the alternative standard levels. In all 15 case study areas, peak O_3 concentrations tended to decrease while the O_3 concentrations in the lower part of the distribution of O_3 tended to increase as the concentrations were adjusted to meet the existing and alternative standards. In addition, O_3 concentrations in the high and mid-range portions of the O_3 distribution generally decreased in the outer, more rural and suburban portions of the urban case study areas, while the O_3 response to NOx reductions was more varied within the urban cores. In particular, while the peak (annual 4th highest daily maximum 8-hour) concentrations upon which the existing and alternative standards are defined generally decreased in the urban core of the case study areas in response to modeled reductions in primarily NOx emissions, the O_3 responses near the center of the O_3 distribution at these locations followed one of three patterns when focusing on the mean of the daily maximum 8-hour O_3 concentrations from May to September, as shown in Table 9-2.

Table 9-2 General Patterns in Seasonal (May-Sept) Mean of Daily Maximum 8-hour O_3 Concentrations after Adjusting to Meet Existing and Alternative Standards*

After Adjusting to Meet Existing Standard	After Further Adjusting to Meet Lower Alternative Standards	Case Study Areas Showing Pattern
Decreased	Continued to decrease	Atlanta Sacramento Washington, D.C
Increased	Decreased	Baltimore Cleveland, Dallas Detroit Los Angeles New York Philadelphia St. Louis
Increased	Continued to increase or remained constant	Boston Chicago Denver Houston

* These patterns refer to O_3 responses in the urban core of each urban case study area based on analysis of the interpolated monitor values used as inputs to the exposure and lung function risk analyses.

The air quality inputs to the exposure modeling and clinical-based lung function risk assessments were estimated hourly O_3 concentrations at each census tract in the 15 case study areas. These values were interpolated from the observed and HDDM-adjusted monitoring data using the Voronoi Neighbor Averaging (VNA) technique. This technique was shown to be an improvement upon the nearest neighbor technique used in the first draft REA and previous O_3 NAAQS reviews (see Appendix 4-A for details). Consequently, the spatial variability of observed and HDDM-adjusted O_3 is better accounted for in these analyses compared to those in the first draft REA.

The air quality inputs to the epidemiology based risk assessment were "composite monitor" values, a time-series of the spatially averaged monitoring data, in 12 of the 15 case study areas. Consequently, in cases of urban case study areas within which O_3 was predicted to increase in some locations and decrease in others, the air quality inputs to this analysis represent a "net" effect for each case study area. The spatial extent of the case study areas used in the composite monitor averages were CBSAs. These CBSA areas are larger than the Zanobetti & Schwartz, 2008 (Z & S) study areas (used in the first draft REA) which include only a subset of the CBSA focused on urban cores. Figure 9-1 (reproduced from Figure 4-7) shows box plots of the composite monitor values for 2006-2008 based on the observed data (black), data adjusted to meet the existing standard using quadratic rollback (blue), and the HDDM adjustment procedure

(red) for the Z & S areas. The two adjustment methods were generally comparable in terms of the changes in the upper quartile of the distribution. However, by design, quadratic rollback always estimated decreases in the 75th percentile, median, and 25th percentile of the composite monitor values, while HDDM estimated decreases in these values in some urban case study areas, and increases in other areas consistent with atmospheric chemistry. HDDM-based adjustments always produced increases in the lower tail of the distribution, while the lower tail values generally remained unchanged with quadratic rollback. The differences between the two adjustment procedures were the most pronounced in Los Angeles and New York, where the largest reductions in NOx were required in order to meet the existing and alternative standards. These large reductions in NOx caused a relatively large increase in lower O_3 concentrations because of the reduction in NOx titration of O_3. As was noted in Chapter 4, the HDDM-based O_3 estimates become more uncertain for larger changes in NOx and VOC emissions, and thus there was less overall confidence in those results. Even in these cases, the HDDM approach is still preferable because it captures better the overall shift in the distribution of O_3 concentrations.

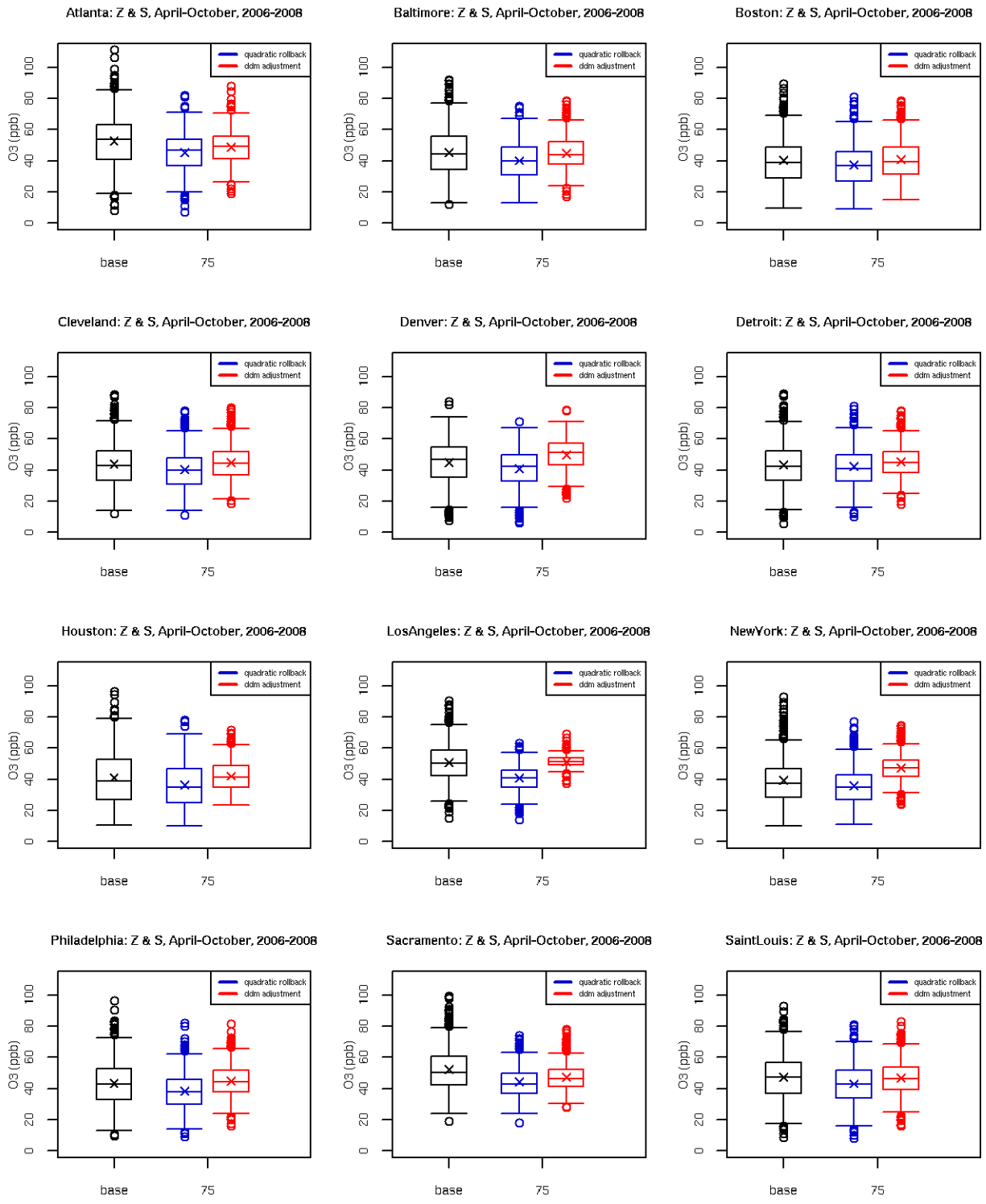


Figure 9-1 Distributions of composite monitor 8-hour daily maximum O_3 concentrations from ambient measurements (black), quadratic rollback (blue), and the HDDM adjustment methodology (red) for meeting the existing standard.

9.2.2 Human Exposure Modeling (Chapter 5)

The population exposure assessment evaluates exposures to O_3 using the Air Pollution Exposure (APEX) model for the general population, all school-aged children (ages 5-18), asthmatic school-aged children (ages 5-18), asthmatic adults (ages > 18), and older persons (ages 65 and older), with a focus on populations engaged in moderate or greater exertion (e.g. children engaged in outdoor recreational activities). The strong emphasis on children, asthmatics, and older adults reflected the findings of the last O_3 NAAQS review (U.S. EPA, 2007) and the ISA (U.S. EPA, 2013, Chapter 8) that these are important at-risk groups.

We assessed exposure in 15 urban case study areas – Atlanta, Baltimore, Boston, Chicago, Cleveland, Dallas, Denver, Detroit, Houston, Los Angeles, New York, Philadelphia, Sacramento, St. Louis, and Washington, D.C. – for recent O_3 concentrations (2006-2010) and for O_3 concentrations adjusted to just meet existing and alternative standards for two time periods (2006-2008 and 2008-2010)[3]. The analysis provided estimates of the percent of several populations of interest exposed to concentrations above three health-relevant 8-hour average O_3 exposure benchmarks: 60, 70, and 80 ppb. The ISA includes studies showing statistically significant effects at each of these benchmark levels (U.S. EPA, 2013). These benchmarks were selected to provide some perspective on the public health impacts from exposures to various concentrations that have been associated with O_3-related health effects (e.g., lung inflammation and increased airway responsiveness) in controlled human exposure and toxicological studies, but cannot currently be evaluated in quantitative risk assessments. In addition, the exposure assessment also identified the specific microenvironments and activities that contribute most to exposure and evaluated at what times and how long individuals were in key microenvironments and were engaged in key activities. This assessment focused on persons experiencing the highest daily maximum 8-hour exposure within each study area. The assessment found that:

- Childhood is an important lifestage where higher exposures and risks can occur, due to the higher time spent outdoors by children, the higher exposure concentration experienced by children while outdoors (i.e. when they are dismissed from school in the afternoon and during the summer, when they may be at an outdoor camp all day) and engagement in moderate or high exertion level activities.

[3] Attainment with the O_3 standard is based on the 4th highest maximum 8-hour O_3 concentration, averaged over 3 years. We evaluated two different 3-year periods in determining how air quality in each of the analytical years would respond to just meeting the existing and alterative levels of the standard. This was done to evaluate the effect of variability in meteorology and emissions on exposures and risks associated with just meeting the existing and alternative standards. For the exposure and lung function risk analyses, which provide estimates for each of the five analytical years, this results in two estimates for 2008, because 2008 is included in each of the 3-year averaging periods and there are separate analytical results for 2008 for the adjusted air quality resulting from simulating attainment in each of the two 3-year periods.

- Persons spending a large portion of their time outdoors during afternoon hours experienced the highest 8-hour O_3 exposure concentrations given that O_3 concentrations in other microenvironments were simulated to be lower than ambient concentrations.
- Highly exposed children spend half of their outdoor time (on average) engaged in moderate or greater exertion levels, such as in sporting activities. Highly exposed adults also spent their outdoor time engaged in moderate or greater exertion levels though on average, not as frequently as children.

Across the 15 urban case study areas, we find that children are of greatest concern for O_3 exposures compared to other lifestages due to the greater amount of time they spend outdoors engaged in moderate or higher exertion activities. The exposure analysis estimates that children have the highest percent of exposures of concern of any of the at-risk populations or lifestages. As a result, we focus on the results for children (ages 5-18) in the remainder of this discussion. Figure 9-2 (reproduced from Figure 5-11) shows the results of the exposure assessment for all 15 urban case study areas, showing trends across the analytical years for the percent of children with at least one 8-hour exposure greater than the 60, 70, and 80 ppb benchmarks.

The limited availability of longitudinal activity diary data and the general population modeling approach used may underestimate the correlation in activity patterns for certain susceptible populations (e.g., outdoor workers), and underestimate how often there are repeated exposures to O_3 concentrations above the exposure benchmarks. As a result, although we are able to report the percent of the population with at least one exposure greater than the alternative exposure benchmarks, we are less confident in the estimated percent of the population experiencing more than one exposure. Individuals with repeated exposures may be at greater risk of significant health effects (U.S. EPA, 2013, Section 6.2.1.1). In addition, the limited data on responses to air quality alerts (e.g., averting behavior) indicates that a small percentage of the population may engage in averting behavior in response to air pollution, which may overstate actual exposures if individuals reduce their exposure during periods of high O_3.

The benchmark exposures of concern are not equivalent to ambient standard levels, as exposures reflect the full pattern of O_3 concentrations throughout a season, coupled with time spent outdoors and indoors engaged in different activities. Thus, just meeting the existing standard will result in shifts in the entire distribution of O_3 over a three year period, and will change the percent of populations experiencing each of the exposure benchmarks of concern. Figure 9-2 shows that the percent of children above the 60 ppb benchmark declines consistently across the 15 urban case study areas when just meeting potential alternative standards of 70, 65, and 60. For most urban case study areas and years, the percent of children above the 60 ppb benchmark is reduced by over half when O_3 is adjusted to meet the 65 ppb alternative standard

relative to the 75 ppb standard. In many urban case study areas and years, just meeting the 65 ppb alternative standard results in close to zero percent of children above the 60 ppb benchmark. For the 70 and 80 ppb benchmarks, meeting an alternative standard of 70 ppb results in a small percentage of children exceeding the benchmarks.

Year-to-year variability is relatively pronounced for exceedances of the 60 ppb benchmark. In addition, we observe a geographic pattern to the years with the maximum percent of exceedances of the exposure benchmarks reflecting the regional O_3 patterns across years. In general, northeastern urban case study areas saw the highest percentage of exceedances during 2007, while southern and western urban case study areas saw a higher percentage of exceedances during 2006. However, these patterns are somewhat dependent on the 3-year averaging period used to determine whether the standards are met. In general variability in the percent of children exceeding the 60 ppb exposure benchmark across urban case study areas is similar to the variability across years.

The percent of children with multiple exposures above the exposure benchmarks is generally much lower compared to the percent of children with single exposures above the benchmarks. However, as noted above, we have lower confidence in these estimates. Even for the lowest benchmark level of 60 ppb, most locations and years have less than 10 percent of children experiencing 2 or more exposures when just meeting the existing standard of 75 ppb, less than 5 percent when just meeting an alternative standard of 70 ppb, and less than 1 percent when just meeting an alternative standard of 65 ppb. For most urban case study areas and years, less than 1 percent of children experience 2 or more exposures above the 70 ppb exposure benchmark when just meeting the existing standard.

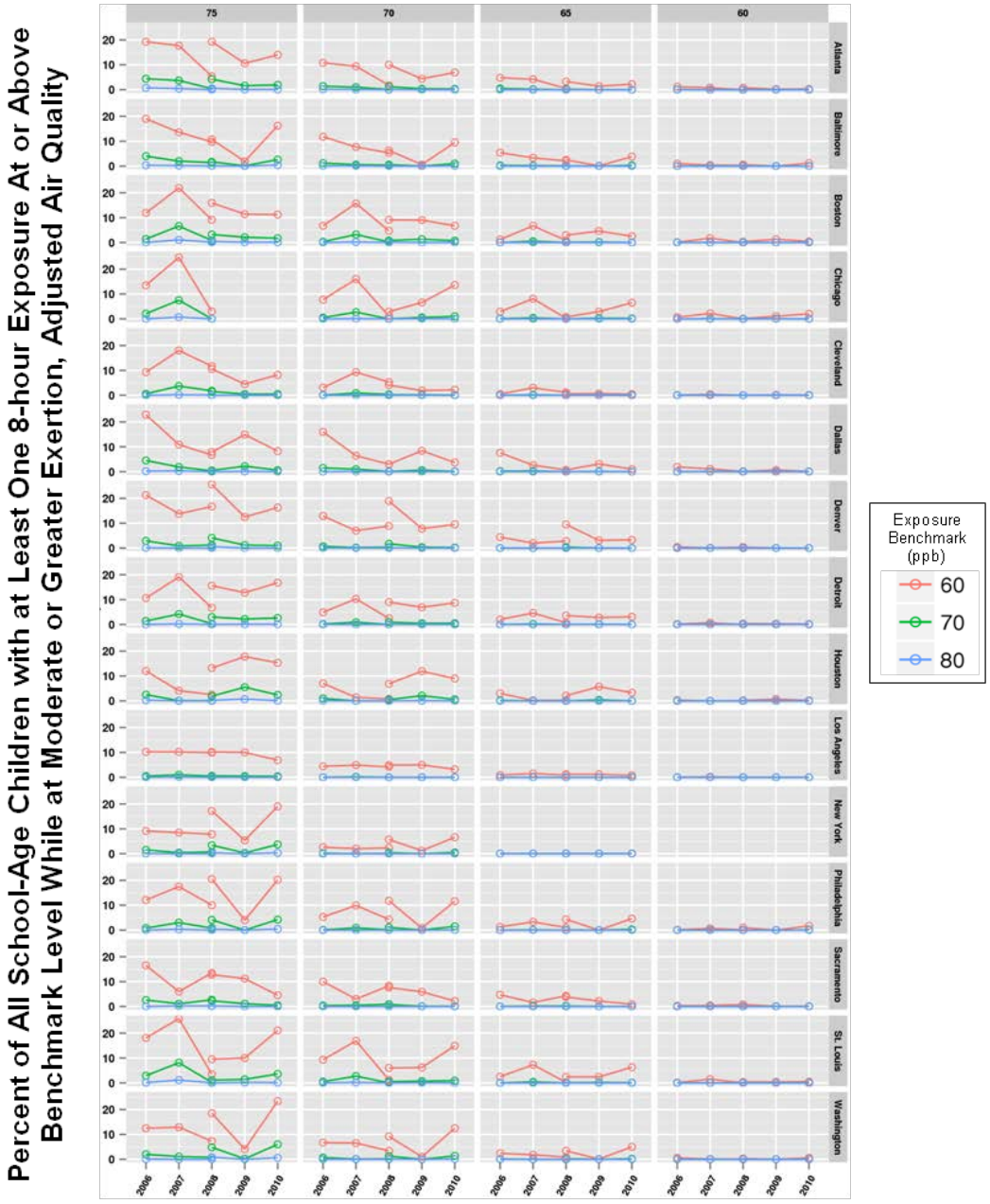


Figure 9-2 Effects of just meeting existing (column 1) and alternative (columns 2 through 4) standards on percent of children (ages 5-18) with at least one O_3 exposure at or above 60, 70, and 80 ppb-8hr while at moderate or greater exertion, years 2006-2010.[4]

[4] We were not able to adjust air quality to just meet the 60 ppb alternative standard in the New York City by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

Table 5-6 summarized the percent of the population of children (ages 5-18) with at least one daily 8-hour exposure above the 60, 70, and 80 ppb benchmarks, providing both the mean and maximum percentage across the five analytical years for each urban case study area. For O_3 adjusted to just meet the existing standard of 75 ppb, the highest maximum percentage of children exceeding the 60 ppb benchmark across years, 26 percent, occurs in Denver, which also has the highest mean percentage across years. After just meeting the existing standard, Los Angeles has the lowest maximum (10 percent) and mean (9.5 percent) percentage of children exceeding the 60 ppb benchmark across years, likely reflecting the highly skewed nature of O_3 concentrations in that urban case study area. For example, just meeting the existing standards in Los Angeles moves the majority of O_3 concentrations (sites and days) well below 60 ppb (See Appendix 4-D). Patterns across urban case study areas are generally similar after just meeting alternative standards of 70, 65, and 60 ppb, with the exception that the lowest maximum and mean percentage of children for the alternative standard level of 65 ppb occurs in the New York City urban case study area, which had very large (greater than 90 percent) reductions in NOx emissions that were used to adjust air quality to just meet the 65 ppb standard level in that urban case study area. This resulted in the distribution of O_3 concentrations covering most days of the year and most monitoring sites shifting dramatically downward, with most concentrations well below 60 ppb across the New York City urban case study area. The level of confidence in the results for the New York City and Los Angeles study areas for just meeting the alternative standards is lower than that for some of the other urban case study areas due to the HDDM-based O_3 estimates becoming more uncertain for very large changes in precursor emissions.

Figure 9-3 (reproduced from Figure 5-19) shows the results of the exposure assessment for all 15 urban case study areas, showing the effect on the percent of children with one or more exposures above the 60 ppb benchmark of just meeting the existing and alternative standards. For each alternative standard, Figure 9-2 shows the maximum percent of children exceeding the benchmark across the modeled years 2006-2010. Patterns of results are similar for the 70 ppb and 80 ppb benchmarks, however, the maximum percents of children exceeding those higher benchmarks are much smaller for all alternative standards. The percent of children exceeding the 80 ppb benchmark is close to zero once the existing standard is met. The percent of children with two or more exposures exceeding the 60 ppb benchmark level is substantially lower when just meeting the existing standard, and is close to zero for the 70 ppb and 80 ppb benchmarks. This percentage drops substantially when meeting the 70 ppb standard, and is close to zero in most urban case study areas when meeting the 65 ppb and 60 ppb alternative standards. Patterns for asthmatic children are very similar to patterns for all children.

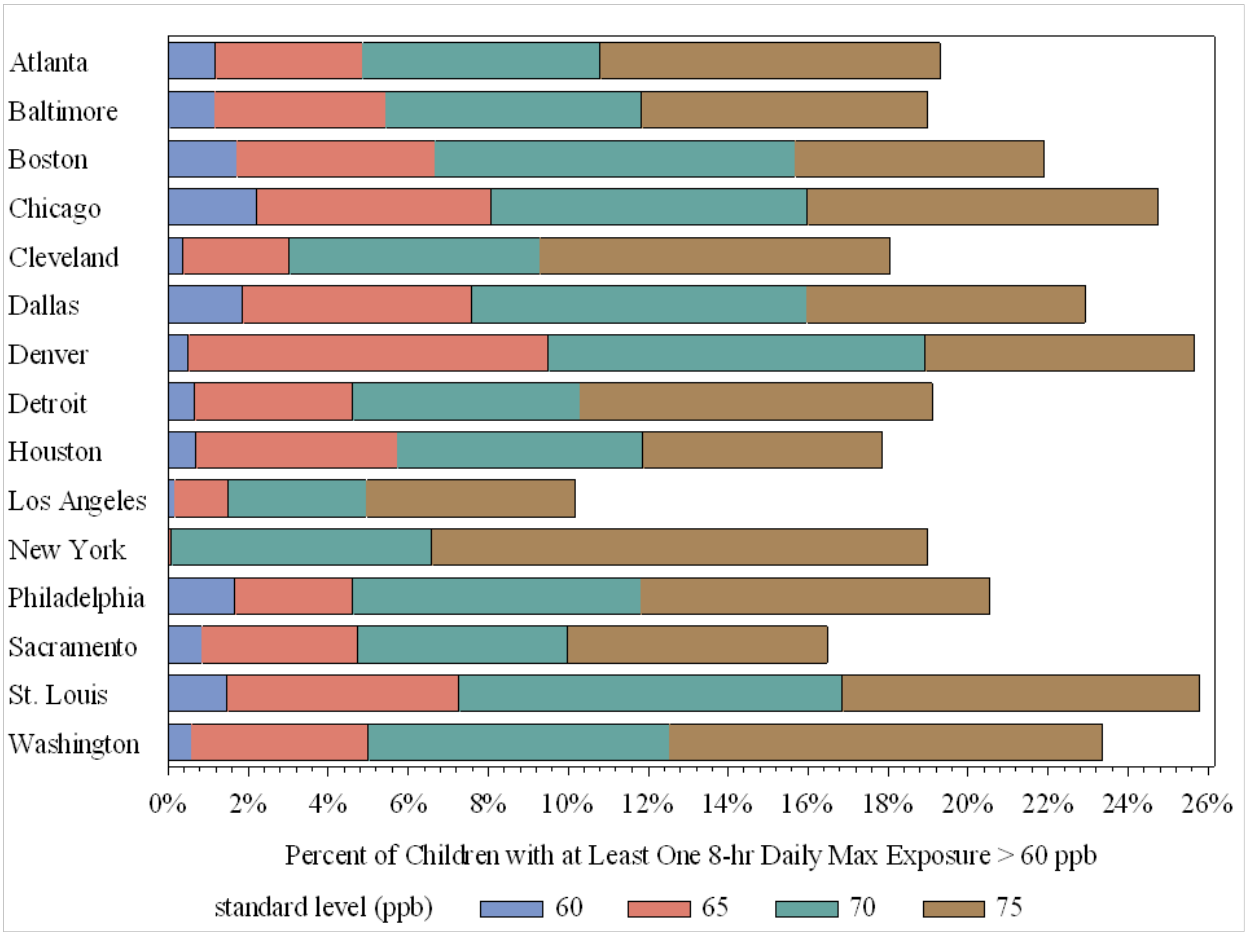


Figure 9-3 Effects of just meeting existing (75 ppb) and alternative standards on percent of children (ages 5-18) exceeding 60 ppb exposure benchmark, highest value across years for each urban case study area, 2006-2010.[5]

9.2.3 Health Risks Based on Controlled Human Exposure Studies (Chapter 6)

Using the estimates of exposure from APEX combined with results from controlled human exposure studies, we estimated the number and percent of at-risk populations or lifestages (all children aged 5-18, children with asthma aged 5-18, adults aged 18-35, adults aged 36-55, and outdoor workers) experiencing selected decrements in lung function. The analysis focuses on estimates of the percent of each at-risk population or lifestage experiencing a reduction in lung function (mostly for durations of one to five hours) for three different levels of impact, 10, 15,

[5] We were not able to adjust air quality to just meet the 60 ppb alternative standard in New York City urban case study area by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

and 20 percent decrements in FEV1. These levels of impact were selected based on the literature discussing the adversity associated with increasing lung function decrements (US EPA, 2013, Section 6.2.1.1). Consistent with the exposure assessment, we focus this summary on lung function decrements in children as they are the lifestage likely to have the greatest percentage at-risk due to higher levels of exposure and exertion. Within the overall population of children, asthmatic children may have less reserve lung capacity to draw upon when faced with decrements, and therefore a ≥10% decrement in lung function may be a more adverse event in an asthmatic child than a healthy child.

Lung function risks (based on experiencing an estimated 10, 15, or 20 percent decrement in lung function) were estimated for each of the 15 urban case study areas in which human exposures were modeled. Two models were used to estimate lung function risks: one based on application of a population level exposure-response (E-R) function consistent with the approach used in the previous O_3 NAAQS review, and one based on application of an individual level E-R function (the McDonnell-Stewart-Smith (MSS) model), introduced in this review, which incorporates individual differences in physiology, age, and activity patterns (McDonnell et al., 2012). Because the individual level E-R function approach allows for a more complete estimate of risk (incorporating risk responses at varying activity levels, not just moderate or greater exertion), we focus on the results of that approach for this discussion.

The MSS model as implemented in APEX has a term that adjusts the lung function response according to an individual's age. The MSS model was fit using data from subjects who ranged in age from 18 to 35. Thus, the MSS model is not able to account for differences in lung function at different age groups between the ages of 5 and 18. However, age does have a pronounced effect on lung function response in the APEX model. APEX models differences in physiological parameters due to age, and these result in age-dependent predictions of ventilation rates, which are used in the MSS model. Ventilation rates also depend on the activities being performed, which are also age-dependent. As a result of differences in physiology and activities, the lung function responses vary by age (see Appendix 6-E).

Figure 9-4 (reproduced from Figure 6-6) shows the results of the lung function risk assessment for all 15 urban case study areas, showing trends across the analytical years for the percent of children with predicted lung function decrements greater than or equal to 10 percent[6].

[6] We have introduced a new method (relative to the O_3 NAAQS review completed in 2008) for calculating the percent of the at-risk populations (all children and asthmatic children) experiencing lung function decrements, based on modeling of individual level responses to O_3 exposures. This model yields significantly higher estimates of the percent of children experiencing lung function decrements greater than 10, 15, and 20 percent. This may be partly due to the specific data inputs from clinical studies used to derive the function, but is also to be expected because the MSS model can reflect greater sensitivity of children to O_3 exposures because it allows for age variability in the relationship between O_3 and FEV1 decrements, and younger populations are more responsive to O_3 exposures than older populations.

Specifically, Figure 9-4 shows that the percent of children (age 5-18) with greater than or equal to 10 percent lung function decrement declines consistently across the 15 urban case study areas when just meeting the existing 75 ppb standard, as well as the alternative standards of 70, 65, and 60. The percent of children at-risk at the 10 percent decrement level remains at or above 10 percent in many locations after just meeting the 60 ppb alternative standard. The percentage of children with greater than or equal to a 15 or 20 percent lung function decrement is much lower for all alternative standards, with close to zero percent of children at-risk when just meeting the alternative standard of 60 ppb. In general variability in percent of children at-risk across urban case study areas is similar to variability across years.

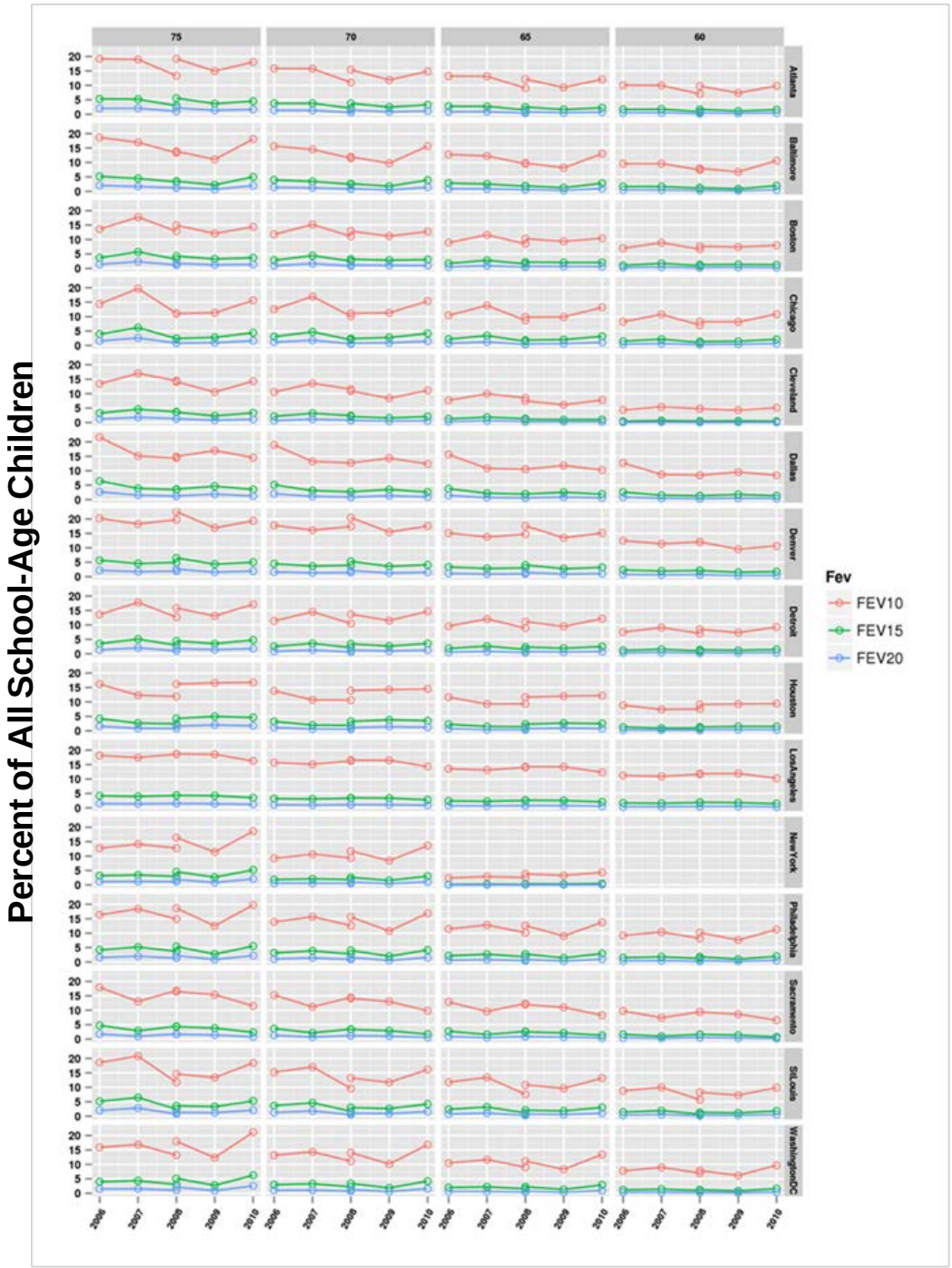


Figure 9-4 Effects of just meeting existing (column 1) and alternative (columns 2-4) standards on percent of children (ages 5-18) with FEV_1 decrement > 10, 15, and 20%, years 2006-2010.[7]

[7] We were not able to adjust air quality to just meet the 60 ppb alternative standard in New York City urban case study area by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

Figure 9-5 (reproduced from Figure 6-11) shows the results of the lung function risk assessment for all 15 urban case study areas, showing the effect on the risk of a 10 percent or greater lung function decrement in children (ages 5-18) of just meeting the existing and alternative O_3 standards. For each alternative standard, Figure 9-5 shows the maximum percent risk over all of the modeled years 2006-2010.

There is no consistent pattern in the percent of children with 10 percent or greater lung function decrement across urban case study areas just meeting the existing standard of 75 ppb. The 5-year maximum estimated percent of children at-risk ranges from 17 to 22 percent across urban case study areas. The percent reduction in 5-year maximum risk when just meeting the 70 ppb alternative standard is more consistent across urban case study areas, ranging from 8 to 23 percent (excluding New York City, which had a reduction of 29 percent). Reductions in risk when just meeting the 65 ppb alternative standard are also generally consistent across urban case study areas, with the exception of New York City. Incremental reductions in risk when just meeting the alternative 65 ppb standard compared with just meeting the 70 ppb alternative standard range from 17 to 31 percent excluding New York City, which has a reduction in risk of more than twice as much as the next largest reduction. Incremental reductions in risk from just meeting the alternative 60 ppb standard compared with just meeting the 65 ppb standard are generally consistent, ranging from 16 to 46 percent, with somewhat larger reductions in risk occurring in Cleveland and Denver. Overall, the 5-year maximum percent of children at-risk for lung function decrements of 10 percent or more exceeds 13 percent, 10 percent, and 5 percent in all urban case study areas except New York after just meeting alternative standards of 70, 65, and 60, respectively. Patterns of risk reductions are also similar for the alternative lung function decrement levels of 15 percent and 20 percent. However, the initial percent of the population experiencing these decrements when just meeting the existing standard are substantially lower.

Patterns of risk responses using the population level exposure-response model are similar to the MSS individual risk model. However, the starting values for the percent of the population at risk are lower, reflecting the limits of the model in reflecting individual level responses, and the limited coverage of the model for exposures at lower exertion levels. For children, the MSS model gives results typically a factor of three higher than the population level E-R model used in the previous O_3 NAAQS review.

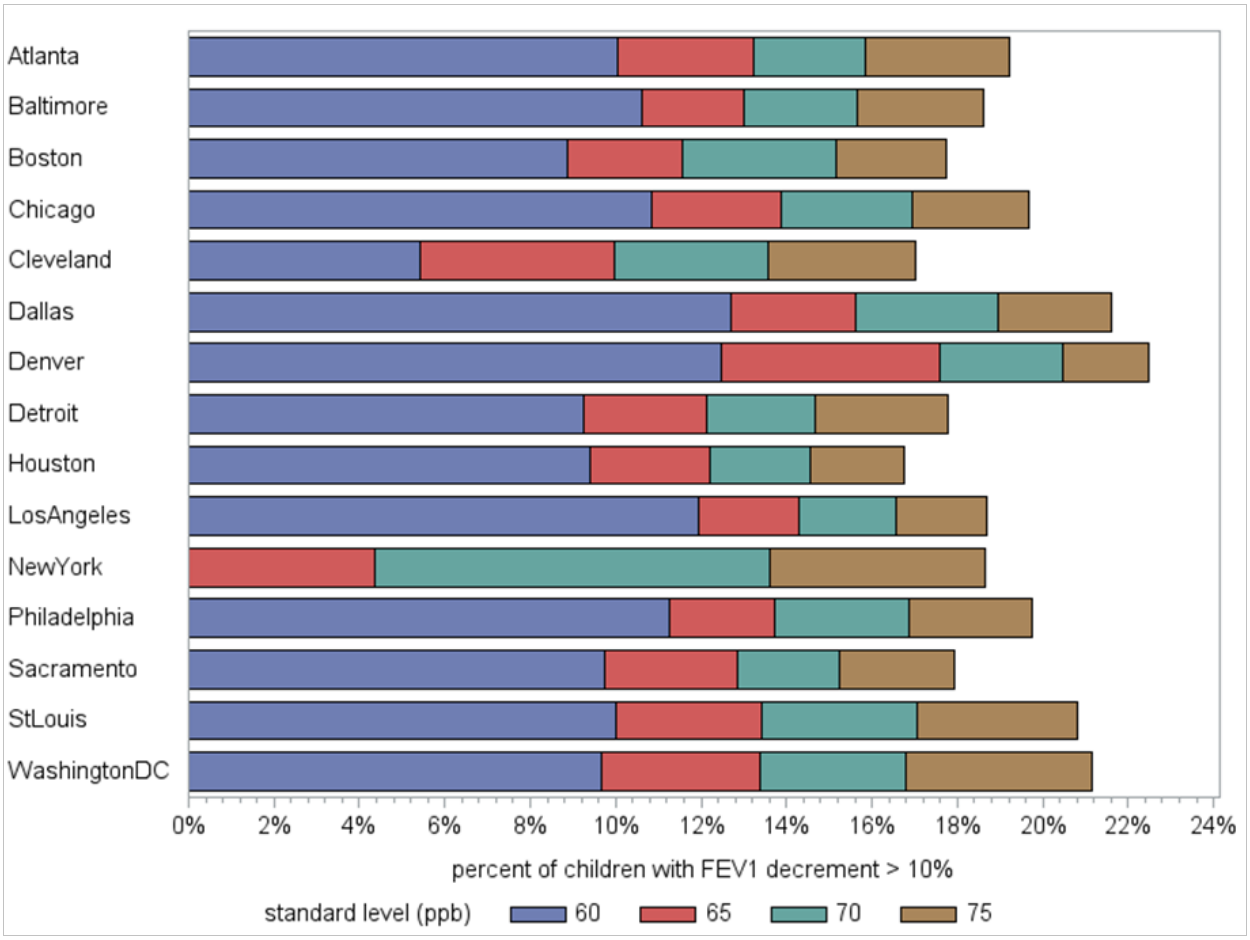


Figure 9-5 Impact of just meeting existing (75 ppb) and alternative standards on percent of children (ages 5-18) with FEV_1 decrement > 10%, highest value for each urban case study area, 2006-2010.[8]

9.2.4 Health Risks Based on Epidemiological Studies (Chapters 7 and 8)

The epidemiology-based risk assessment evaluated mortality and morbidity risks from short-term O_3 exposures and mortality risks from long-term exposures to O_3 by applying concentration-response (C-R) functions derived from selected epidemiology studies. The analysis included both a set of urban case study area case studies and a national-scale assessment. The urban case study analyses evaluated mortality and emergency department (ED) visits, hospitalizations, and respiratory symptoms associated with recent O_3 concentrations (2006-2010) and with O_3 concentrations adjusted to just meet the existing and alternative O_3

[8] We were not able to adjust air quality to just meet the 60 ppb alternative standard in New York City by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

standards (see section 9.2.1 and Chapter 4). Mortality and hospital admissions (HA) were evaluated in 12 urban case study areas, while ED visits and respiratory symptoms were evaluated in a subset of areas with supporting epidemiology studies. The 12 urban case study areas were: Atlanta, GA; Baltimore, MD; Boston, MA; Cleveland, OH; Denver, CO; Detroit, MI; Houston, TX; Los Angeles, CA; New York, NY; Philadelphia, PA; Sacramento, CA; and St. Louis, MO. The urban case study analyses focus on risk estimates for the middle year of each three-year design value period (2006-2008 and 2008-2010) in order to provide estimates of risk for a year with generally higher O_3 concentrations (2007) and a year with generally lower O_3 concentrations (2009).

Most of the endpoints evaluated in epidemiology studies cover the entire study population including children and adults. Because most mortality and hospitalizations occur in older persons, these epidemiology-based risk estimates are better indicators of effects in adults than in children. This is an important distinction from the human exposure and lung function risk assessments, which focus on children. The only endpoints specific to children are asthma and all respiratory hospital admissions using the New York specific epidemiology study, respiratory ER visits in Atlanta, and respiratory symptoms in asthmatic children in Boston.

Both the urban case study area and national-scale assessments provide the absolute incidence and percent of incidence attributable to O_3. In addition, risks are presented in terms of incidence per 100,000 population to control for the differences in the sizes of the populations across urban case study areas, and to allow for comparison of risks using different definitions of urban extent. In previous reviews, O_3 risks have only been estimated for the portion of total O_3 attributable to North American anthropogenic sources (above what was referred to in previous reviews as "policy-relevant background O_3"). In contrast, this assessment estimates risk for O_3 concentrations down to zero, reflecting the lack of evidence for a detectable threshold in the C-R functions (U.S. EPA, 2013, Chapter 2), and the understanding that U.S. populations may experience health risks associated with O_3 resulting from emissions from all sources, both natural and anthropogenic, within and outside the U.S. In order to better reflect how O_3 distributions are likely to respond to just meeting existing and potential alternative standard levels, we adjusted O_3 concentrations to just meet existing and potential alternative standard levels using reductions in only U.S. anthropogenic emissions of O_3 precursors. Thus, the estimated changes in risk between just meeting the existing standards and just meeting potential alternative standard levels only reflect reductions in U.S. anthropogenic emissions.

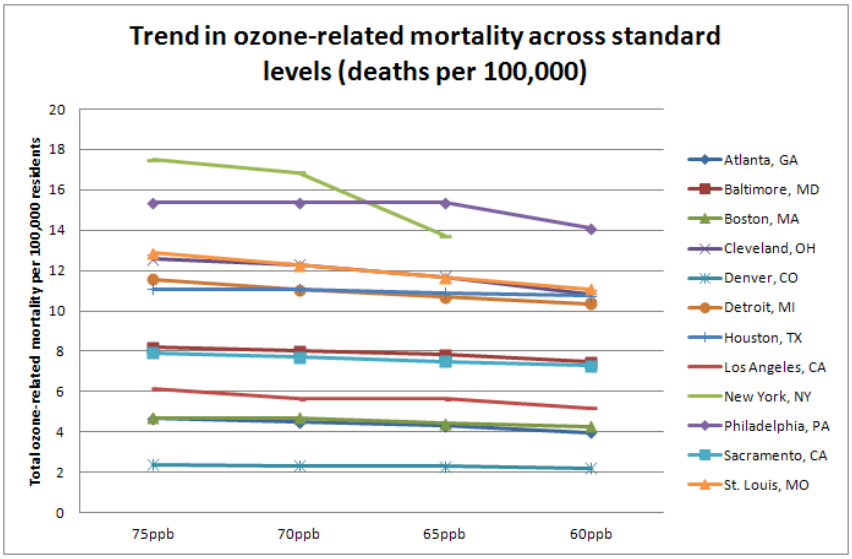
However, consistent with the conclusions in the O_3 ISA (U.S. EPA, 2013), we have relatively lower certainty about the shape of the C-R function towards the lower end of the distribution of O_3 concentrations used in fitting the function due to the reduction in the number of O_3 measurements in this portion of the distribution. We discuss this source of uncertainty

below. In addition, we provide the distribution of mortality incidence across the range of O_3 concentrations in Chapter 7 to inform discussions of uncertainty in the results.

9.2.4.1 Urban Case Study Results

Figures 9-6 and 9-7 (reproduced from Figures 7-4 and 7-5) show the results of the mortality and adult (ages 65 and older) respiratory hospital admissions risk assessments for all 12 urban case study areas, showing the effect on the incidence per 100,000 population just meeting the existing 75 ppb standard and alternative O_3 standards of 70, 65, and 60 ppb in 2007 and 2009.

2007 Simulation year



2009 Simulation year

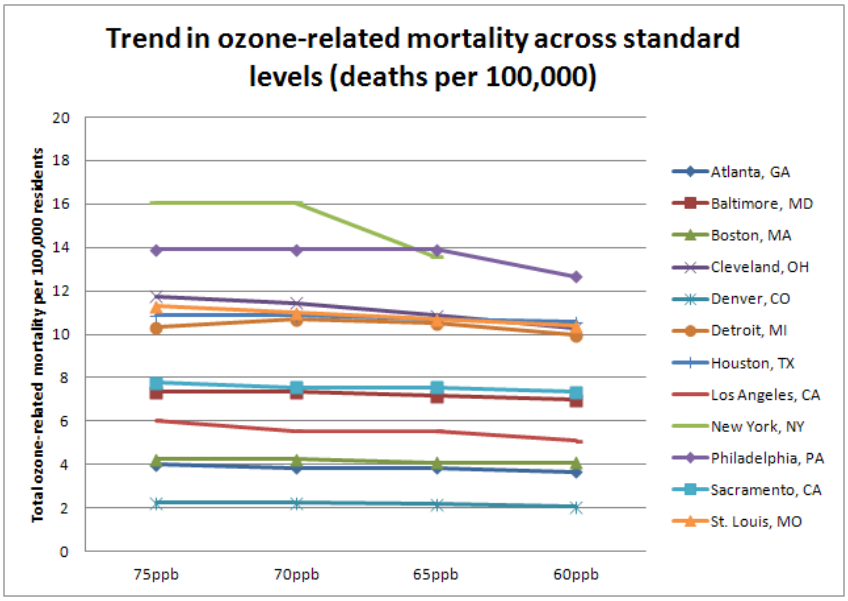
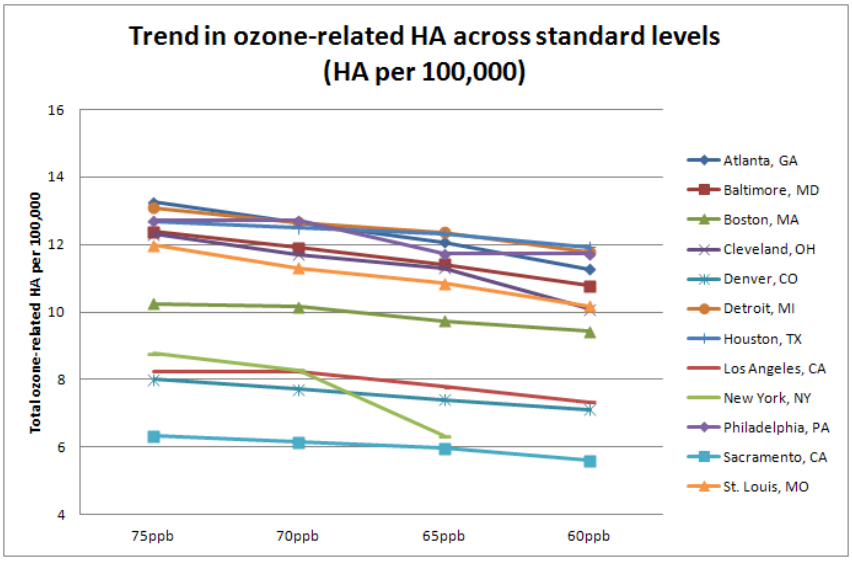


Figure 9-6 Impacts of just meeting existing (75 ppb) and alternative standard levels on mortality risk per 100,000 population for 2007 and 2009.[9]

[9] As noted earlier, we were not able to adjust air quality to just meet the 60 ppb alternative standard in New York City by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

2007 Simulation year



2009 Simulation year

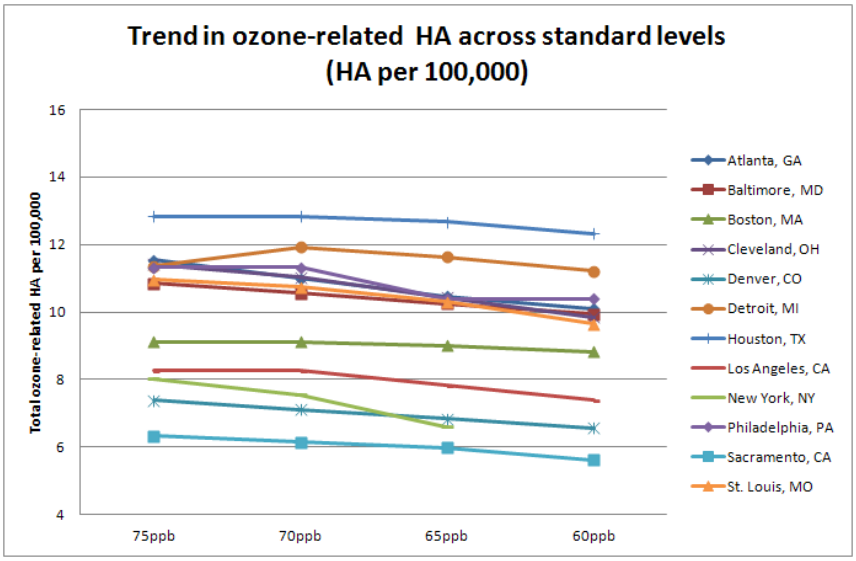


Figure 9-7 Impacts of just meeting existing and alternative standard levels on adult (ages 65 and older) respiratory hospital admissions risk per 100,000 population for 2007 and 2009.[10]

[10] We were not able to adjust air quality to just meet the 60 ppb alternative standard in New York City urban case study area by reducing U.S. NOx and VOC emissions (see chapter3 and appendix 4-D for details). Detroit was already meeting the existing standard for 2008-2010.

In some urban case study areas which have large NOx emissions (e.g. from heavy downtown traffic), O_3 levels are artificially low because the NOx emissions remove O_3 through a chemical reaction (see section 9.2.1 and Chapter 4). In these places, when NOx emissions are decreased to reduce peak O_3 concentrations across the entire CBSA, which often includes locations outside of the urban core areas, lower concentrations of O_3 can go up. This can also happen in other areas on the lowest O_3 days. This phenomenon occurs in some locations when meeting lower alternative standards as well.

The overall trend across urban case study areas is small decreases in mortality and morbidity risk as O_3 concentrations are adjusted to just meet incrementally lower alternative standard levels. In New York, there are somewhat greater decreases in these risks, reflecting the relatively large emission reductions used to adjust air quality to just meet the 65 ppb alternative standard, and the substantial change in the distribution of O_3 concentrations that resulted. We were not able to adjust O_3 concentrations to just meet the 60 ppb alternative standard in the New York City urban case study area. Risks vary substantially across urban case study areas; however, the general pattern of reductions across the alternative standards is similar between urban case study areas. Because of the generally lower baseline O_3 concentrations in 2009, risks are generally slightly lower in 2009 relative to 2007; however, the patterns of reductions in risk are very similar between the two years.

Mortality and morbidity risks generally do not show large responses to meeting existing or alternative levels of the standard for several reasons. First, these risks are based on C-R functions that are approximately linear along the full range of concentrations, and therefore reflect the impact of changes in O_3 along the complete range of 8-hour average O_3 concentrations. This includes days with low baseline[11] O_3 concentrations that are predicted to have increases in O_3 concentrations, as well as days with higher starting O_3 concentrations that are predicted to have decreases in O_3 concentrations as a result of just meeting existing and alternative standards. Second, these risks reflect changes in the urban-area wide monitor average, which will not be as responsive to air quality adjustments as the design value monitor, and which includes monitors with both decreases and increases in 8-hour concentrations. Third, the days and locations with predicted increases in O_3 concentrations (generally those with low to midrange starting O_3 concentrations) resulting from just meeting the existing or alternative standard levels generally are frequent enough to offset days and locations with predicted decreases in O_3. The heat maps presented in Figures 7-2 and 7-3 demonstrate that just meeting progressively lower alternative standard levels narrows the distribution of risk across the range

[11] By low baseline concentrations, we mean area-wide average O_3 concentrations between approximately 10 and 40 ppb prior to adjustments to just meet the existing and alternative standards.

of O_3 concentrations. In addition, the distribution of risk tends to be more centered on area-wide average concentrations in the range of 25 to 55 ppb after just meeting an alternative standard of 60 ppb. The focus of the epidemiological studies on urban case study area-wide average O_3 concentrations, and the lack of thresholds coupled with the linear nature of the C-R functions mean that in this analysis, the impact of a peak-based standard (which seeks to reduce peak concentrations regardless of effects on low or mean concentrations) on estimates of mortality and morbidity risks based on results of those studies is relatively small. For example, for mortality and hospital admissions, we find a less than 10 percent reduction in risk for most urban case study areas when just meeting the 70 ppb and 65 ppb alternative standards compared to just meeting the existing standard, and a less than 25 percent reduction in risk for all urban case study areas when just meeting the 60 ppb standard compared to just meeting the existing standard. The general pattern for other morbidity risks is similar to hospital admissions. However, we are not able to draw strong conclusions about the results across urban case study areas, because of the limited number of urban case study areas represented for most of the endpoints.

We have applied city-specific mortality effect estimates to each urban case study area based on the largest multi-city epidemiological study. However, for many of the urban case study areas, the risk estimates have wide confidence intervals that can include zero, due to the lower statistical power of some of the city-specific effect estimates relative to the national combined effect estimate across cities. Furthermore, there is significant variability in these effect estimates across the 12 urban case study areas, with some urban case study areas having effect estimates from 5 to 7 times greater than other cites (see Chapter 7, section 7.4.1).[12] The variability in effect estimates, along with differences in O_3 concentrations, is a driver for the overall variability in the risk results across cities. Smith et al (2009) reports an overall significant national mortality effect estimate with confidence intervals that do not include zero, reflecting the much greater statistical power available when pooling information across urban case study areas.

We also evaluated mortality risks in the 12 urban case study areas associated with long-term O_3 exposures (based on the seasonal average (April to September) of the peak daily one-hour maximum concentrations). Risks from long-term exposures after just meeting the existing standard are substantially greater than risks from short-term exposures, ranging from 16 to 20 percent of respiratory mortality across urban case study areas. However, the percent reductions in long-term mortality risks are similar to those for mortality from short-term exposures. For example, we find a less than 10 percent reduction in risk relative to just meeting the existing

[12] This substantial heterogeneity in effect estimates can reflect a number of factors including differences in population susceptibility and behavior related to O_3 exposure and risk (e.g., proximity to roadways, use of air conditioning, commuting patterns, time spent outdoors) and differences in the degree to which the O_3 monitoring network used in the epidemiological study reflects patterns of population exposure.

standard in most areas when just meeting the 70 ppb and 65 ppb alternative standards, and a less than 20 percent reduction when just meeting the 60 ppb alternative standard level. Risk reductions for the New York City urban case study area are much greater when just meeting the 65 ppb alternative standard compared to just meeting the existing standard, with a 24 percent reduction in risk in 2007.

New York and Los Angeles have characteristics that make epidemiological risk estimates particularly uncertain. In the case of New York, the expansion of the urban case study area definition to the CBSA adds uncertainty due to the large and diverse nature of the CBSA. The New York CBSA includes two urban case study areas which have separate effect estimates available from the Smith et al. (2009) study. These separate effect estimates (for Newark, NJ and Jersey City, NJ) are smaller than the effect estimate for New York, however, they are also based on much smaller populations, and have relatively wider confidence bounds, reflecting low statistical power. For consistency with other urban case study areas and to allow for comparison between the CBSA-based risk estimates and the smaller study area based estimates (see the sensitivity analyses in Chapter 7), we elected to apply the New York city effect estimate, which is based on a very large population and has high precision, to all of the counties in the New York CBSA. While this adds substantial uncertainty to the absolute incidence of mortality for the New York CBSA, it does not affect the pattern of risk reductions when just meeting alternative standards. In addition, as noted earlier, the O_3 adjustments to meet existing and alternative standards in New York and Los Angeles also have additional uncertainties relative to the other 10 urban case study areas.

We conducted a number of sensitivity analyses based on a population normalized mortality risk metric, e.g. mortality risk per 100,000 population. Maintaining the general linear, no-threshold functional form, mortality risks per 100,000 population are generally robust to alternative specifications of the C-R functions, although in several urban case study areas, using effect estimates from Smith et al. (2009) which were derived using regional priors rather than national priors results in higher risk estimates[13]. Using the effect estimates from Zanobetti and Schwartz (2008) has no consistent effect on risk results across the urban case study areas. Using effect estimates based on a copollutant model with PM_{10}, mortality risks are higher in some locations and lower in others. However, in all locations the confidence intervals are substantially

[13] In Bayesian modeling, effect estimates are "updated" from an assumed prior value using observational data. In the Smith et al (2009) approach, the prior values are either a regional or national mean of the individual effect estimates obtained for each individual city. The Bayesian adjusted city specific effect estimates are then calculated by updating the selected prior value based on the relative precision of each city-specific estimate and the variation observed across all city-specific individual effect estimates. City-specific estimates are pulled towards the prior value if they have low precision and/or there is low overall variation across estimates. City-specific estimates are given less adjustment if they are precisely estimated and/or there is greater overall variation across estimates.

wider using the copollutant model with PM_{10} (due to fewer days with both pollutants measured), which makes it difficult to determine whether the increases and decreases in estimates relative to the core estimates are real or the result of statistical error.

We selected the CBSA as the spatial definition for the urban case study areas. We made this selection to address a downward bias that we identified resulting from a mismatch between the smaller urban core areas used in the epidemiology studies and the larger areas where O_3 concentrations are expected to change as a result of meeting the existing and alternative standard levels (see Chapter 7). We included a sensitivity analysis evaluating the result of using a smaller geographic area including only the counties used in the epidemiology study. As expected, using a smaller geographic extent for the urban case study areas results in smaller, and in some cases negative risk reductions when compared to using the CBSA definitions. This reflects the fact that the controlling[14] monitor in many of the 12 urban case study areas is located outside of the small set of counties included in the Smith et al. (2009) urban case study area definitions, and some of the monitors that are within that more limited spatial extent are more prone to O_3 titration due to local NOx emission sources. As a result, those monitors are more likely to see increases in O_3 which will, if other monitors with higher concentrations in the broader regions are not included, lead to estimated increases in risk due to the application of a linear, no threshold C-R function. This bias can be substantial, especially in St. Louis and several urban case study areas in the Northeast, including Boston, New York, and Philadelphia, where the highest concentration monitors are outside the Smith et al. (2009) urban case study area definitions.

Sensitivity analyses were conducted for scenarios of just meeting existing and alternative standards using combinations of NOx and VOC emissions reductions (as compared to NOx reductions alone). The addition of VOC emissions reductions had little impact with the exception of New York and Los Angeles, where risk was decreased relative to the NOx-only reduction scenario.

9.2.4.2 National-scale Assessment Results

The national-scale assessment evaluated only mortality associated with recent O_3 concentrations across the entire U.S for 2006-2008. The national-scale assessment is a complement to the urban scale analysis, providing both a broader geographic assessment of O_3-related health risks across the U.S., as well as an evaluation of how well the 12 urban study areas represented the full distribution of O_3-related health risks in the U.S. The national-scale assessment demonstrates that there are O_3 risks across the U.S, not just in urban case study areas, even though the O_3 concentrations in many areas were lower than the existing standard level.

[14] The controlling monitor is the monitor with the highest design value within a defined non-attainment area.

While we did not assess the changes in risk at a national level associated with just meeting existing and alternative standards, just meeting existing and alternative standards would likely reduce O_3 concentrations both in areas that are not meeting those standards and in locations surrounding those areas, leading to risk reductions that are not included in the urban-scale analysis.

9.3 COMPARISON OF RESULTS ACROSS EXPOSURE, LUNG FUNCTION RISK, AND EPIDEMIOLOGY-BASED MORTALITY AND MORBIDITY RISK ANALYSES

In considering the overall results across the human exposure, lung function risk, and epidemiology-based risk assessments, we focus on the key policy-relevant metrics and levels for each type of assessment. For the human exposure assessment, we selected exposures above the 60 ppb exposure benchmark for all children (ages 5-18). We select this exposure metric because children represent a key at-risk lifestage, and the 60 ppb exposure benchmark is the lowest exposure level associated with significant findings in controlled human exposure studies. For the lung function risk assessment, we selected the results for lung function decrements greater than or equal to 10 percent for all children (ages 5-18). We select this lung function risk metric because children represent a key at-risk lifestage, and a 10 percent lung function decrement represents a potentially more adverse event in asthmatic children. For the epidemiology-based risk assessment we selected the core short-term exposure mortality results and the respiratory hospital admission results, because these endpoints were estimated for all of the 12 urban case study areas. Generally speaking, these metrics provide the most differentiation between the alternative standards, helping to inform policy-relevant questions regarding adequacy of the existing standard, and public health impacts of meeting alternative standards. The other metrics analyzed in this REA (e.g. other exposure benchmarks and other lung function decrements) show less response to just meeting the existing standard or potential alternative standard levels.

As discussed in Chapter 2, we designed the exposure and risk assessment to help inform two fundamental questions related to the adequacy of the existing standard in protecting public health and the degree of exposure and risk reductions associated with alternative standards compared with the existing standard. The following discussion evaluates the three types of analyses we conducted in terms of the consistency of the information provided to inform these questions.

9.3.1 Evaluation of Exposures and Risks After Just Meeting the Existing Standard

To compare the results of the three assessments in urban case study areas, we plot the key metrics from each analysis across urban case study areas for the two common years of analysis (i.e., 2007 and 2009). For three urban case study areas (i.e., Chicago, Dallas, and Washington

D.C.) we have only the exposure and lung function risk assessments, as these urban case study areas did not have sufficient information to estimate epidemiology-based risks. The epidemiology-based metrics are the percent of baseline short-term exposure mortality, based on the core estimates using the C-R functions from Smith et al. (2009), and respiratory hospital admissions based on the core estimates using the C-R functions from Medina-Ramon (2006), attributable to O_3. Figure 9-8 presents the exposures and risks after just meeting the existing standard of 75 ppb. Each row represents one of the key analytical results; each column gives the results for 2007 and 2009 for each urban case study area. The scale of each analytical metric for each analysis differs, and thus the comparisons across analyses should focus on overall patterns rather than on direct comparisons of numeric estimates.

All of the metrics show substantial variability among urban case study areas, although there appears to be less variability in lung function risk and hospital admission risk compared with the exposure metric and mortality risk. The differences between estimates for 2007 and 2009 are much higher for some urban case study areas (e.g. Baltimore and Philadelphia) for the exposure metric than any of the risk metrics. This may reflect the explicit threshold nature of the exposure metric, which focused on exposures above a benchmark level of 60 ppb. Differences between years in exposures above the 60 ppb benchmark after just meeting the existing standard are dependent on the number of days during each year with decreases in higher O_3 concentrations, as well as the magnitude of the decreases in O_3 on those higher O_3 concentration days. These in turn are sensitive to the shape of the O_3 distribution in the analytical year prior to just meeting the existing standard (which determines the starting number of days above 60 ppb) and the response to emissions reductions applied in meeting the existing standard for 2007 or 2009. There is some consistency between metrics in the urban case study areas with highest values for the exposure and lung function risk metrics. However, there were still differences, especially for Los Angeles, which had one of the higher values for lung function risk in 2009, but had one of the lower percentages of children exposed above the 60 ppb benchmark. This again points to the importance of the threshold nature of the exposure metric, combined with the tendency for more substantial decreases in peak O_3 concentrations relative to mid-range and low concentrations when just meeting the existing standard.

There is little consistency within urban case study areas between the epidemiology risk metrics and the exposure and lung function risk metrics, and there is also little consistency

between the mortality and hospital admission risks. Houston has the lowest metric values in 2007 (except for mortality risk), but in 2009 has some of the higher risk metrics (except for hospital admission risk). New York has the highest mortality risk in 2007 and 2009 but has among the lowest hospital admission risks in both years.

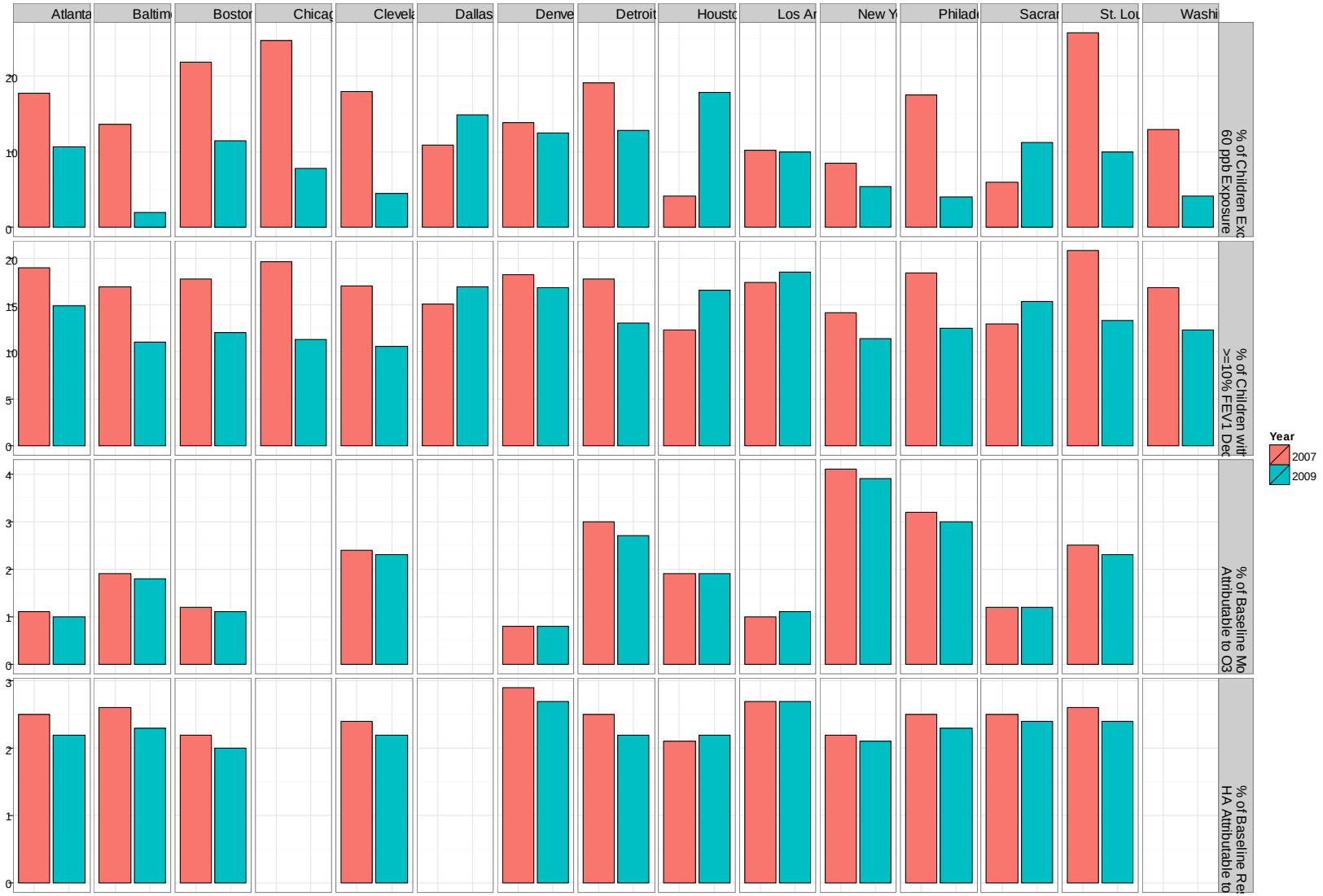


Figure 9-8 Comparison of Exposure (Row 1) Lung Function Risk (Row 2) and Epidemiology-Based Risk (Rows 3 and 4) Metrics after Just Meeting the Existing 75 ppb Standard.

9.3.2 Reductions in Exposure and Risk Metrics after Just Meeting Alternative Standards

To compare the results of the three assessments for urban case study areas after just meeting alternative standards relative to the existing standard, we express each result as a percent of the metric value when just meeting the existing standard. Figure 9-9 presents the percent reduction in exposures and risks after just meeting alternative standards relative to just meeting the existing standard of 75 ppb. In this plot, each row represents one of the key analytical results and each column gives the results for 2007 and 2009 for each urban case study area. The scales are the same between analyses, and as such, it is informative to examine both the overall patterns of change between alternative standards, and also the absolute value of the percent reductions in risk metrics between analyses. In interpreting this chart, higher values mean greater reductions in risk or exposure relative to just meeting the existing standards. Because these are percent reductions, the maximum value is one hundred percent, which if reached would indicate that risks or benchmark exposures are completely eliminated when the alternative standard is met in the urban case study area as was seen for the 60 ppb exposure benchmark.

Many of the differences in results across the metrics are driven by how each metric is affected by the O_3 data input to the analysis. In general, the impact of the HDDM adjustments to O_3 vary based on three main considerations: 1) the degree to which the exposure or risk metric is sensitive to changes across the various ranges of O_3 concentrations (e.g. high, mid-range, low); 2) whether the exposure or risk metric uses individual census tract concentrations or area-wide average concentrations; and 3) changes in the distribution of O_3 concentrations in the year of analysis between recent O_3 concentrations and adjusted (meeting the existing or alternative standards) O_3 scenarios. With respect to 1), the exposure benchmark metric, which focuses only on exposures above 60 ppb, will not be sensitive at all to changes in O_3 concentrations in the range below 60 ppb. The lung function risk metric, which depends on the dose rate and individuals' characteristics, does not have a concentration threshold. However, because of the logistic form of the response function, it is less sensitive to lower O_3 concentrations and has very few FEV1 responses greater than 10 percent when exposure concentrations are below 20 ppb and very few FEV1 responses greater than 15 percent when exposure concentrations are below 40 ppb. On the other hand, the mortality and hospital admission risk metrics are based on non-threshold, approximately linear C-R functions, and therefore will be sensitive to changes in O_3 along the full range of O_3. As discussed in Chapter 4, because O_3 at lower concentrations may increase following HDDM adjustment in some locations and on some days to just meet alternative standards[15], this can lead to increases in risk on some days, which can lead to a net

[15] The frequency and magnitude of increases in spatially averaged mean concentrations in an urban case study area occur during a season when adjusting air quality to just meet a standard vary considerably between the existing

increase or decrease in risk over the entire year, depending on whether the days with increased risk exceed days with decreased risk (generally due to a preponderance of days with lower O_3 concentrations). With respect to 2), the exposure and lung-function risk metrics are based on concentrations at individual census tracts since they depend on O_3 exposure modeled by moving each individual through their environment. Because of this, the exposure and lung-function risk metrics are most affected by the spatial and temporal variability of O_3 concentrations across the urban case study area. The mortality and hospital admission risk metrics are calculated applying C-R functions to area-wide, daily maximum 8-hr average O_3 concentrations. As a result, the spatial variability in O_3 concentrations between the monitors will only influence the epidemiology-based risk estimates in how they influence the area-wide average. With respect to 3), all three metrics are influenced by how the distribution of O_3 concentrations changes between recent O_3 conditions and after adjustment to just meet existing and alternative O_3 standard levels.

and alternative standards. The highest frequency of occurrence of days with increasing O_3 happens when adjusting air quality to just meet the existing standards, and decreases as air quality is further adjusted to just meet lower alternative standard levels.

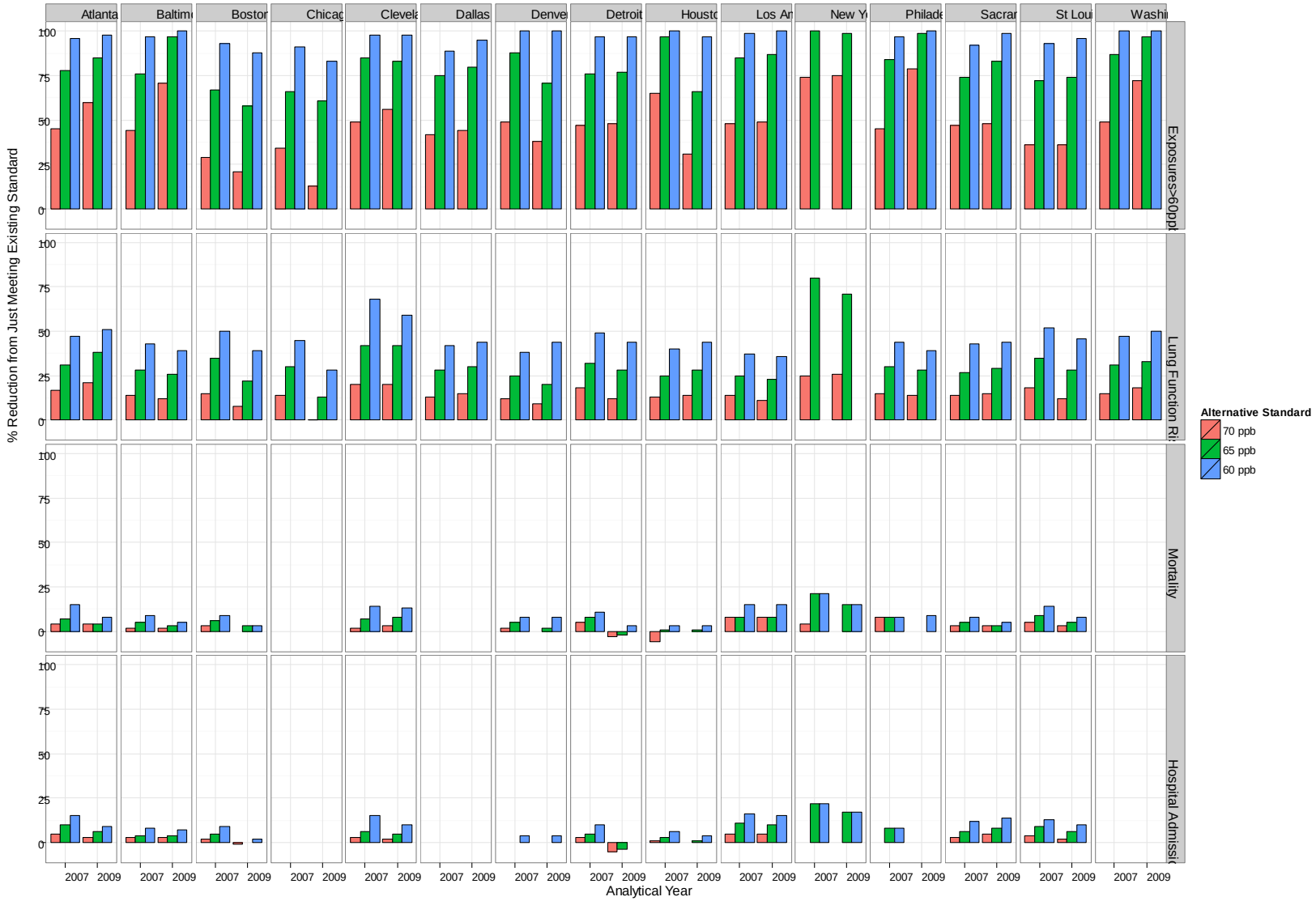


Figure 9-9 Comparison of Percent Reduction in Key Risk Metrics for Alternative Standard Levels Relative to Just Meeting the Existing 75 ppb Standard.

The exposure and lung function risk metrics are most affected by the reductions in the individual monitors' peak O_3 concentrations, including the magnitude of these reductions and the number of days that experience these reductions. In contrast, the mortality and hospital admission risk metrics are affected by changes in the mean of the seasonal, area-wide average O_3 concentrations, where the mean is determined by the frequency and magnitude of increases versus decreases in area-wide, maximum daily 8-hr O_3 concentrations[16]. In addition to O_3 concentrations, there are other factors that affect the variability across urban case study areas for these three metrics, such as activity data and exposure factors for the exposure and lung function risk metrics and the study-specific C-R functions for the mortality and hospital admission risk metrics.

One clear observation is that the percent reductions in risk from meeting alternative standard levels relative to meeting the existing standard for the two epidemiology-based endpoints are much smaller than for the exposure benchmark and lung function risk endpoints. The maximum percent reduction in the mortality and hospital admissions risk relative to just meeting the existing standard across years, locations, and alternative standards is less than 25 percent, and for many years/locations, the reductions in these risks when just meeting the lowest alternative standard, 60 ppb, are less than 10 percent. The exposure benchmark results show the most reductions when comparing just meeting the existing standard to just meeting alternative standards. Just meeting the 65 ppb standard results in reductions in the percent of children exceeding the 60 ppb exposure benchmark by over 50 percent in all urban case study areas, and by over 75 percent in 12 of the 15 urban case study areas evaluated. For most locations and years, just meeting the 60 ppb alternative standard reduced the percent of children exceeding the 60 ppb exposure benchmark by over 90 percent compared to just meeting the existing standard. Reductions in lung function risk were also much higher than reductions in mortality and hospital admissions risk. Just meeting the 65 ppb standard results in reductions in lung function risk by over 25 percent in most locations and years, and just meeting the 60 ppb standard results in reductions by over 40 percent in most locations and years.

There is general consistency in the city-to-city patterns of reductions in the exposure and lung function risk metrics, although the decreases in lung function risk are less than half as large as the reductions in the percent of children exceeding the 60 ppb exposure benchmark (with the clear exception of New York city, which we will discuss further below). The patterns of reductions in mortality and hospital admission risk are generally consistent with the patterns for

[16] As noted previously, changes in the spatial extent of the urban case study areas over which monitors are averaged can change the magnitude and sign of the change in the spatial average O_3 concentration for an urban case study area. For example, we found that we bias the risk estimates low when using urban case study area definitions that include only urban core counties and not the counties with monitors experiencing the most reductions in O_3.

exposure and lung function risk for 2007, with the exception of Houston and Philadelphia. However, for 2009, the patterns for mortality and hospital admission risk are quite different, both from the 2007 results, and from the exposure and lung function risk results. This is due to the generally lower O_3 concentrations in 2009, which results in a greater number of days with predicted increases in O_3 concentrations at low concentrations, fewer days with very high concentrations where predicted reductions in O_3 occur, and a smaller predicted decrease in O_3 concentrations on those high days. This affects the mortality and hospital admissions risk more than the exposure and lung function risk metrics because those metrics incorporate thresholds, and therefore are not responsive to changes in O_3 concentrations below those thresholds.

Additional considerations are important in interpreting the reduction in exposure and risk between the existing standard and alternative standards. The REA analyses focus on reducing peak O_3 concentrations, in particular the 4th high O_3 concentration averaged over 3-years so as to simulate meeting the existing standard or various alternative standards. In addition, the air quality adjustments are based on applying reductions in U.S. anthropogenic emissions. In this way, the adjusted air quality reflects day-to-day O_3 concentrations that could occur when focusing on reducing high O_3 concentrations rather than on reducing mean O_3 concentrations. In addition, because the analyses do not include reductions of O_3 precursor emissions from sources other than U.S. anthropogenic emissions (e.g. international emissions, biogenics, etc), the O_3 concentrations in the adjusted air quality account for O_3 created from natural and international sources, even if 100 percent emissions reductions are applied to U.S. anthropogenic sources in adjusting air quality scenarios.

Finally, with respect to the epidemiology based analyses, we note that 2007, which had generally higher O_3 concentrations than 2009, had more days where O_3 concentrations decreased as a result of adjusting peak O_3 concentrations to just meet alternative standards. Thus just meeting alternative standards resulted in net decreases in risk in all locations, with the exception of Houston for just meeting the 70 ppb alternative standard. In contrast, 2009, which had generally lower concentrations than 2007, had more days in the range where O_3 concentrations were increased as a result of adjusting peak O_3 concentrations to just meet alternative standards, and thus the patterns reflect some locations where mortality and hospital admissions risk increases. However, for 2009, in all locations, when just meeting the lowest alternative standard of 60 ppb, mortality and hospital admission risks are decreased relative to just meeting the existing standard.

9.4 OVERALL ASSESSMENT OF REPRESENTATIVENESS OF EXPOSURE AND RISK RESULTS

9.4.1 Representativeness of Selected Urban Case Study Areas in Reflecting Areas Across the Nation with Elevated Risk

We selected urban case study areas for the exposure and risk analyses based on several criteria (e.g. recent elevated O_3 concentrations and presence of at-risk populations and lifestages) we identified as likely indicators of areas and populations likely to experience high O_3 exposures and risks (see Section 7.3.1). We then conducted several analyses to determine the extent to which our selected urban case study areas actually represent the highest mortality and morbidity risk areas. We compared the distributions of risk characteristics[17] and mortality risk (based on recent O_3 concentrations) for the 12 urban case study areas used in the epidemiology-based risk assessment with the corresponding national distributions. We also evaluated the degree to which our selected urban case study areas represent the patterns of O_3 concentration changes experienced by the overall U.S. population.

Based on the comparisons of distributions of risk characteristics, the selected urban case study areas represent urban case study areas that are among the most populated in the U.S., have relatively high peak O_3 concentrations, and capture well the range of city-specific mortality risk effect estimates. These three factors alone would suggest that the case study urban case study areas should capture well the overall risk for other heavily populated urban case study areas in the nation, with a potential for better characterization of the high end of the risk distribution. The selected urban case study areas do not include those with the highest numbers of some at-risk populations or lifestages, specifically older people with high baseline mortality rates. However, most locations in the U.S. (except Florida) with high percentages of older people have low overall populations, less than 50,000 people in a county, or low O_3 concentrations. This suggests that while the risk per exposed person per ppb of O_3 may be higher in these locations, the overall risk to the population is likely to be within the range of risks represented by the urban case study locations.

Based on the comparisons of distributions of short-term O_3 exposure mortality risk (using the percent of mortality metric) for recent O_3 concentrations, the 12 selected urban case study areas are representative of the full distribution of U.S. O_3-related mortality risk in urban case study areas. Two of the selected areas, New York and Philadelphia are representative of the highest end of the distribution of short-term O_3 mortality risk. Overall, O_3 mortality risk for short-term O_3 exposures in the 12 urban study areas are representative of the full distribution of

[17] In this context, risk characteristics are the elements of populations, air quality, and inputs to the C-R functions that are expected to be correlated with estimated mortality risks (see Chapter 8).

U.S. urban O_3-related mortality, representing both high end and low end risk counties. For the long-term O_3 exposure mortality risk metric (again using the percent of mortality), the 12 urban study areas are representative of the central portion of the distribution of risks across all U.S. counties; however, the selected 12 urban case study areas do not capture the very highest (greater than 98th percentile) or lowest (less than 25th percentile) ends of the national distribution of long-term exposure-related O_3-related risk.

9.4.2 Representativeness of Selected Urban Case Study Areas in Reflecting Responsiveness of Risk to Just Meeting Existing and Alternative O_3 Standards

While we selected urban case study areas to represent those populations likely to experience elevated risks from O_3 exposure, we did not include among the selection criteria the responsiveness of O_3 in the urban case study area to decreases in O_3 precursor emissions that would be needed to just meet existing or alternative standards.

In our preliminary evaluations of risk modeling results, we observed a consistent presence of days with low to midrange starting O_3 concentrations for which O_3 concentrations (using the 8-hour maximum metric) increased after adjustments to just meet the existing and alternative standards across the selected urban case study areas. As noted above, this led to estimates of increased risk on those days, and in some cases, estimates of increased risk over the course of the O_3 season, reflecting both the magnitude and frequency of the predicted increases relative to the predicted decreases in O_3 concentrations. As explained above, this pattern was more pronounced when using a more spatially limited definition of the urban case study areas, but even when using the CBSA definitions, there were still days when the area-wide average O_3 increased, primarily due to predicted increases in O_3 in the core counties of the urban case study areas.

In order to better understand how prevalent this type of air quality response was across the U.S., we conducted several additional analyses of O_3 concentrations. These included evaluations of trends at O_3 monitors during a period of time with significant O_3 precursor emission reductions, and evaluations of temporal and spatial patterns of O_3 changes across the U.S., based on air quality modeling results, to simulate how O_3 would change across the U.S. in response to NOx (and VOC) emissions reductions (relative to recent 2007 levels) similar to those used in the HDDM adjustments (see section 9.2.1 above). The latter analysis includes an assessment of the association of different types of O_3 responses with population counts to help characterize the degree to which populations in the U.S. experience O_3 conditions like those in the selected 15 urban case study areas (see Chapter 8).

Overall, both types of analyses showed that decreases in O_3 precursor emissions lead to decreases in O_3 concentrations in areas with higher starting O_3 concentrations, which tend to be rural or suburban case study areas, and on days with higher O_3 concentrations. The analyses also

indicate that in urban core areas (those with high levels of fresh NOx emissions), decreases in NOx emissions can lead to increases in O_3, primarily for days when initial O_3 concentrations are suppressed due to NOx titration. The observed widespread decreases of median O_3 in suburban and rural locations when NOx emissions are decreased suggest the efficacy of large NOx emissions reductions on reducing O_3 over large regions of the country.

These results suggest that many of the urban case study areas may show O_3 responses that are typical of other large urban case study areas in the U.S., but may not represent the response of O_3 in other populated areas of the U.S., including suburban case study areas, smaller urban case study areas, and rural areas. These smaller urban case study areas would be more likely than our urban case study areas to experience area-wide average decreases in mean O_3 concentrations as O_3 standards are met. Even though large urban case study areas throughout the U.S. have high population density, 73 percent of the U.S. population lives outside of these high population density areas[18], and thus, a large proportion of the population is likely to experience greater mortality and morbidity risk reductions in response to reductions in 8-hour O_3 concentrations than are predicted by our modeling in the selected urban case study areas. The analyses presented in Section 8.2.3.2 show that populations in the case study areas we selected are approximately twice as likely to experience increasing mean O_3 concentrations as populations in the U.S. as a whole. Because our selection strategy for risk modeling was focused on identifying areas with high risk, we tended to select large urban population centers. As discussed in the previous section, this strategy was largely successful in including those urban case study areas in the upper end of the O_3 risk distribution. However, this also has led to an overrepresentation of the populations living in locations where we estimate increasing mean seasonal O_3 in response to adjusting air quality to just meet the existing and alternative standards using NOx emissions reductions. The implication of this is that our estimates of mortality and morbidity risk reductions for the selected urban case study areas are likely to understate the average risk reduction that would be experienced across the population and should not be seen as representative of potential risk reductions for most of the U.S. population.

9.5 OVERALL ASSESSMENT OF CONFIDENCE IN EXPOSURE AND RISK RESULTS

As with any complex analysis using estimated parameters and inputs from numerous data sources and models, there are many sources of uncertainty that may affect our exposure and risk estimates. These sources of uncertainty are discussed in each of the chapters related to air quality, exposure, lung function risk, and epidemiology based mortality and morbidity risk. The

[18] High population density areas are defined here as locations with population densities greater than 1000 people/km^2

overall effect of the combined set of uncertainties on confidence in the interpretation of the results of the analyses is difficult to quantify. However, we provide our judgment of our overall confidence here, with an understanding that alternative judgments may also be supported.

The degree to which each analysis was able to incorporate quantitative assessments of uncertainty differed, due to differences in available information on uncertain parameters and complexities in propagating uncertainties through the models. In general, we followed the World Health Organization tiered approach to uncertainty characterization (WHO, 2008), which includes both quantitative and qualitative assessments. Each chapter includes a table identifying and characterizing the potential impact of key uncertainties on risk estimates, including the degree to which we were able to quantitatively address those uncertainties.

In considering our overall confidence in the results, there are several key considerations discussed below related to sources of uncertainty which we were not able to fully quantify, but which may have a large impact on both overall confidence and confidence in individual analyses.

9.5.1 Uncertainties in Modeling O_3 Responses to Meeting Standards

There is inherent uncertainty in all deterministic air quality models, such as CMAQ, the photochemical grid model used to develop the model-based O_3 adjustment methodology. Evaluations of air quality models against observed pollutant concentrations build confidence that the model performs with reasonable accuracy despite both structural and parametric uncertainties. A comprehensive model performance evaluation provided in Appendix 4-B shows generally acceptable model performance that is equivalent to or better than typical state-of-the science regional modeling simulations described in Simon et al. (2012). Two additional sources of uncertainties in the HDDM adjustment methodology are the applicability of HDDM sensitivities over large emissions perturbations and the variability in data used to create regressions which allowed the application of these sensitivities to ambient data. Both sources of uncertainty are shown to be reasonably small in chapter 4 with the first having a mean error of less than 1ppb for 50% NOx cuts and less than 4 ppb for 90% NOx cuts. The uncertainty introduced from the application of regressions to determine sensitivities were quantified by propagating uncertainties in the sensitivities through to uncertainties in the final predicted O_3 concentrations which had standard errors less than 1.4 ppb for all adjustment scenarios. New York and Los Angeles had the largest uncertainties in these two areas due to the fact that they required the largest reductions in NOx emissions. Uncertainties stemming from the application of 8-months of model data to 5-years of ambient data and the across-the-board emissions cut assumptions are further discussed in chapter 4 but are not expected to substantially degrade confidence in the air quality results.

9.5.2 Uncertainties in Modeling Exposure and Lung-function Risk

With regard to the exposure and lung-function risk estimates, the modeling explicitly incorporates population variability in many of the modeling inputs. We did not attempt to probabilistically incorporate the many sources of uncertainty in model parameters or input data due to limitations in the ability to specify distributions characterizing our confidence in those variables. To explore the impacts of some of the more important sources of uncertainty, we conducted a limited set of sensitivity analyses. For the exposure assessment, the estimate of repeated exposures above exposure benchmarks is based on the limited set of diaries of activity data available in the Consolidated Human Activity Database (CHAD) database (see Chapter 5). The method for constructing activity patterns over the course of an O_3 season may not fully capture the behavior of children who have systematically high outdoor activity levels. As a result, while we are able to report the percent of children with two or more exposures, modeling of the distribution of multiple exposures is limited, and the ability to identify the percent of the population with unusually high numbers of multiple exposures is not possible.

For the lung function risk assessment, sensitivity analyses indicate that the MSS model parameter related to the impact of the ventilation rate was most influential in determining the estimated number of children with FEV_1 decrements greater than 10 percent. Estimates of lung function decrements are also influenced by how much variability in individual response is assumed in the MSS model. Sensitivity analyses indicate that when a greater amount of variability is allowed in the MSS model, the percent of children ages 5-18 with FEV_1 decrements greater than 10 percent can increase substantially. In addition, we performed analyses to understand the age-related factors in APEX that could influence the estimated FEV_1 decrements. It was found that the four most influential factors influencing the relationship between the predicted FEV_1 decrement and age are the decreasing level of exertion, the decreasing equivalent ventilation rate (with increasing age), the higher time spent outdoors by children, and the higher exposure concentration experienced by children while outdoors. These all lead to children having higher FEV_1 decrements than adults, and are more influential than the MSS model age term.

9.5.3 Uncertainties in Modeling Epidemiological-based Risk

A major issue in using the results of the epidemiology studies in estimating risk is the narrow geographic definition used for urban case study areas in the epidemiology studies. In many of the urban case study areas, we observe two distinct patterns of O_3 response to the reductions in precursor emissions we evaluated to just meet the existing and alternative standard levels. The first pattern generally occurs in areas outside the urban core (e.g. suburban and rural areas), and on days when O_3 concentrations are on the higher end of the distribution of O_3 concentrations, and is characterized by predicted decreases in 8-hour O_3 concentrations. These

tend to be the locations where the highest 8-hour design values occur. The second pattern generally occurs in the urban core, and on days when O_3 concentrations are on the lower end of the distribution of O_3 concentrations, and is characterized by predicted increases in 8-hour O_3 concentrations. The narrow definitions of urban case study areas used in the epidemiological studies generally included the urban core areas, but did not include all of the suburban or rural areas. The narrow geographic definitions led to a clear downward bias in the estimates of risk changes that would be associated with just meeting the standards in the urban case study areas, because the risk changes would reflect the locations with a tendency towards increases in 8-hour O_3, but would not include locations outside the urban core with decreases in O_3. In many cases, the narrowly defined geographic definitions used in the epidemiology studies did not even include the location with the monitor that was violating the standard. We addressed this bias by expanding the urban case study area to the CBSA. However, this adds additional uncertainty to the risk estimates, and reduces our confidence that we have a good match between the basis of the C-R function (just urban core locations) and the risk analysis context (including both urban core counties and other counties in the CBSA). A clear implication of this decision is that the absolute incidence estimates will be larger than if the analysis was limited to a smaller number of counties. For this reason, we have placed more emphasis on risk metrics that have been normalized for population size (e.g. risks per 100,000 population and percent risk), so as to facilitate comparisons between cities of different population sizes and to reduce the influence of population size on the risk metrics.

The epidemiology studies used as the source for C-R functions for short-term exposure mortality and morbidity endpoints all use time-series approaches to estimate the effect of daily variations in O_3 concentrations on daily mortality or morbidity incidence. The effect estimates developed in these epidemiology studies were based on air quality and health information observed over periods of time in the past (1987-2000). These effect estimates were based on day-to-day variations in area-wide O_3 concentrations estimated from observed concentrations at monitors that reflect a specific set of emissions and atmospheric conditions. In our REA analyses, we apply these effect estimates to adjusted air quality scenarios that are reflective of substantial changes in O_3 concentrations across an area due to, in some cases, large decreases in NOx and VOC emissions reductions. The resulting spatial and temporal patterns of O_3 may not be the same as the spatial and temporal patterns of O_3 that existed at the time of the epidemiology study. The potential for different spatial and temporal patterns in O_3 concentrations between the adjusted air quality scenarios and the air quality observed during the epidemiology study period potentially adds uncertainty to the estimates of risk, as it is not clear the degree to which the exposure surrogate used in the epidemiology study correlates with the exposure surrogate used in the risk analysis. The degree of this potential uncertainty increases

with the amount of emissions reductions applied in the adjusted air quality scenario. This is because as the amount of emissions reductions applied increases, the spatial and temporal patterns of O_3 concentrations become increasingly different from those patterns observed for recent O_3 concentrations (2006-2010) that are more similar, although not identical (due to reductions in NOx between 2000 and 2006), to the patterns for the time period covered by the epidemiology studies (1987-2000) . We are not able to quantify the effect or magnitude of this uncertainty, because we do not know the relationship between O_3 variability and the C-R functions. However, to the extent that the uncertainty is shown to be important, it seems reasonable to conclude that the larger the adjustment to the O_3 distributions, the more likely there could be a mismatch in the exposure surrogates.

Overall, these sources of uncertainty cause us to have reduced confidence in estimates of short-term risk based on modeling the larger (CBSA-based) study areas using the multi-city time series-based effect estimates. This reduces the utility of the risk assessment in directly informing the decision regarding the level of the standard since we have lower confidence in estimates of absolute risk associated with a given standard level. However, the risk assessment can still be useful in providing estimates of the general magnitude and direction of changes in risk associated with an alternative standard level.

9.6 OVERALL INTEGRATED CHARACTERIZATION OF RISK IN THE CONTEXT OF KEY POLICY RELEVANT QUESTIONS

Our analyses set out to inform two questions: 1) what are the magnitudes of exposures of concern and risks for O_3-related health effects that are estimated to occur with O_3 concentrations that just meet the existing O_3 standard?; and 2) to what extent do alternative standards reduce estimated exposures and risks of concern attributable to O_3, focusing on at-risk populations and lifestages? In evaluating risk, we did not limit the assessment to just the absolute risk that is attributable to U.S. or North American emissions, as this is not relevant to answer the two questions. Instead, we estimated total risk from all O_3 concentrations and the distribution of risk over the range of O_3 concentrations. Our estimates of changes in risk from meeting alternative O_3 standard levels relative to meeting the existing standard reflect only the impact of reductions in U.S. precursor emissions on O_3 distributions, recognizing that these emissions are most likely to be affected by implementation of the standards.

To inform these questions, we conducted air quality, exposure, and risk analyses for selected urban case study areas. We evaluated changes in the distribution of O_3 concentrations along the full range of O_3 concentrations down to zero. We have utilized a new method (compared to the O_3 NAAQS review completed in 2008) for estimating O_3 concentrations consistent with attaining existing and alternative standards, based on modeling the response of

O_3 concentrations to reductions in U.S. anthropogenic NOx and VOC emissions, using the HDDM capabilities in CMAQ. This modeling incorporates all known emissions, including emissions from non-anthropogenic sources and anthropogenic emissions from sources in and outside of the U.S. As a result, background O_3 concentrations are directly modeled and, therefore, do not need to be separately specified. Application of this approach also addresses the recommendation by the National Research Council of the National Academies (NRC, 2008) to explore how emissions reductions might affect temporal and spatial variations in O_3 concentrations, and to include information on how NOx versus VOC control strategies might affect exposure to O_3 and potential risks.

We estimated exposures and risks using several different metrics. Consistent with the available evidence, we estimated the percentages of different study populations and lifestages with exposures exceeding several health-based exposure benchmarks. We estimated lung function risks based on a model of individual risk of lung function decrements that incorporates a dose-equivalent threshold and individual exposures, activity levels, and physiology. We estimated mortality and morbidity risks based on non-threshold C-R functions derived from epidemiology studies. These three different analyses result in differing sensitivities of results to changes in the O_3 concentration distribution. Because the three metrics are affected differently in the analyses by changes in O_3 at low concentration levels, it is important to understand these changes in O_3 at low concentrations in interpreting differences in the results across metrics.

We also evaluated the degree to which exposures of concern and lung function risk were reduced in the portions of urban case study areas (urban core areas) that were more likely to experience an increase in low concentrations of O_3, and in some cases an overall net increase in epidemiology based mortality and morbidity risk (results for this assessment are presented in Appendix 9A). We compared these estimates of changes in exposures and lung function risk to estimates of changes in exposures and lung function risk in the areas outside of the urban core areas to judge whether for exposures of concern and lung function risk we see the same pattern of risk reduction between those areas.

Both exposures of concern and lung-function risk estimates in the core urban case study areas showed similar patterns compared with the areas outside the urban cores when just meeting the existing and potential alternative standards. Thus, we observe that in urban core areas which in some cases showed overall increases in epidemiology based mortality and morbidity risk when looking across these same air quality scenarios (see section 9.5.3), we generally see reductions in exposures of concern and lung function risk. These findings illustrate that populations within core urban case study areas are likely to experience risk reductions for health endpoints reflected in the exposure and lung-function analyses.

The mortality and morbidity risk assessment is the analysis that is most sensitive to the increases in O_3 in the lower part of the distribution of initial O_3 concentrations at some monitors and on some days after meeting the existing and alternative standards in some urban case study areas. As demonstrated in the heat maps (Figures 7-2 and 7-3), the increases in O_3 (and resulting estimated increases in risk) occur largely on days with initial O_3 concentrations in the range of 10 to 40 ppb. In addition, mean O_3 concentrations for the urban case study areas change little between air quality scenarios for meeting the existing and alternative standards, because mean concentrations reflect both the increases in O_3 at lower concentrations and the decreases in O_3 occurring on days with high O_3 concentrations. This leads to small net changes in mortality and morbidity risk estimates for many of the urban case study areas. For New York, we find there is a larger decrease relative to other urban case study areas (nearly five times as large as the next largest result for Los Angeles), in mortality and respiratory hospital admissions when just meeting the 65 ppb alternative standard compared to just meeting the existing standard, reflecting the large degree of air quality adjustment needed to meet the standard at all monitors in New York. Both the net change in risk and the distribution of risk across the range of O_3 concentrations in the urban case study areas may be relevant in considering the degree of additional protection provided by just meeting existing and alternative standards.

The dampened response of short-term mortality risk can be contrasted with lung function risk estimates based on application of results from controlled human exposure studies. The lung function risk estimates primarily reflect changes in the upper end of the O_3 distribution and reflect counts of exceedances of lung function decrement benchmarks, rather than summing risks across all days in the season. In addition, lung function risks are based on detailed micro-environmental exposure modeling which uses individual monitor values instead of composite monitor values, thereby resulting in less dampening of spatial variability in O_3 within a given urban study area.

The exposure benchmark analysis is the least sensitive to changes in O_3 in the lower part of the distribution of initial O_3 concentrations, because the lowest of the exposure benchmarks is at 60 ppb, well above the portion of the distribution of initial O_3 concentrations that increased. Since the modeled exposures will always be less than or equal to the monitor concentrations, a benchmark of exposure at 60 ppb is above the range of O_3 concentrations where the HDDM approach estimates increases in concentrations. Thus, this metric is most reflective of the decreases in O_3 at high concentrations that are expected to result from just meeting the existing and alternative standards.

The lung function risk analysis is less sensitive than the mortality and morbidity risk assessments to increases at very low concentrations of O_3, because the risk function is logistic and shows little response at lower O_3 dose rates that tend to occur when ambient concentrations

are lower (generally less than 20 ppb for the 10 percent FEV_1 decrement and generally less than 40 ppb for the 15 percent FEV_1 decrement). However, because there are still some increases in O_3 concentrations that occur in the 50 to 60 ppb range where the estimated risk is more responsive, there may be some reduction in the magnitude of the risk decrease (this is evident when comparing the lung function risk metric with the exposure benchmark metric in figure 9-8).

The exposure-based lung function risk assessment is based on controlled human exposure studies which studied responses in healthy adults. Although the lung function model based on this population shows less responsiveness at lower ambient concentrations, the applicability of this model to the responses of more sensitive populations and lifestages, including children and asthmatics, is uncertain. In addition, although the most complete information for generating an exposure-response function is available for FEV1 as a measure of lung function, there are other, potentially more public health relevant effects, such as lung inflammation, which have also been shown to respond to O_3. As such, the lung function risk analysis should be seen as providing useful but not complete information on risks of health responses to O_3.

Exposures above health benchmarks and risks remain after adjusting O_3 to just meet the existing standard. The percentage of children with at least one 8-hour O_3 exposure exceeding 60 ppb is greater than 10 percent in at least one of the five analytical years for all of the 15 urban case study areas. The percent of children with a predicted decrement in lung function greater than or equal to 10 percent is greater than 16 percent in at least one of the five analytical years for all of the 15 urban case study areas, and for a 15 percent decrement is less than 7 percent for all years and areas. O_3-attributable mortality is slightly less than one percent up to four percent of total mortality across the 12 urban case study areas, with little variation between 2007 and 2009. O_3-attributable respiratory hospital admissions are between 2 and 3 percent across the 12 urban case study areas, with little variation between 2007 and 2009. The percent attributable risk for other morbidity endpoints is somewhat higher than for respiratory hospital admissions, but we only estimated these endpoints for a more limited set of urban case study areas due to data limitations.

The degree of reduction in exposures and risks when adjusting O_3 from just meeting the existing standard to just meeting lower alternative standard levels varies considerably between metrics. The greatest degree of reduction occurs in exposures above the 60 ppb exposure benchmark, followed by reductions in lung function decrements greater than or equal to 10 percent, with the smallest changes in mortality and respiratory hospital admissions. Although the magnitude of reduction differs between the different exposure and risk metrics, there are generally the same patterns of reductions for the exposure benchmark and lung function risk metrics, showing consistent reductions across all 15 urban case study areas. Risk reductions also

occur in most of the urban case study areas for mortality and respiratory hospital admissions. However, these reductions are small, and reflect net changes in risk that include days with risk increases as well as risk decreases. For most urban case study areas, the greatest incremental reductions in exposures above the 60 ppb benchmark occurred when just meeting 70 ppb compared to just meeting the existing standard. Just meeting lower standards of 65 ppb and 60 ppb had incrementally smaller reductions in the percent of children exposed above 60 ppb. Incremental lung function risk reductions are more even between alternative standards, with similar or greater incremental reductions for the 65 ppb and 60 ppb alternatives compared with the incremental reductions for just meeting 70 ppb. Incremental reductions in mortality and respiratory hospital admissions risk are small between alternative standards, but more urban case study areas have somewhat larger risk reductions when comparing just meeting the 60 ppb alternative to just meeting the 65 ppb standard, than when comparing 65 ppb to 70 ppb or 70 ppb to 75 ppb. Long-term exposure mortality risk results show larger absolute estimates of mortality risk and more consistent reductions across urban case study areas. However, percent changes in long-term exposure mortality are similar to those for short-term exposure mortality.

In conclusion, we have estimated that exposures and risks remain after just meeting the existing standards and that in many cases, just meeting alternative standard levels results in reductions in those exposures and risks. Meeting alternative standards has larger impacts on metrics that are not sensitive to changes in lower O_3 concentrations. When meeting the 70, 65, and 60 ppb alternative standards, the percent of children experiencing exposures above the 60 ppb health benchmark falls to less than 20 percent, less than 10 percent, and less than 3 percent in the worst O_3 year for all 15 case study urban case study areas, respectively. Lung function risk also drops considerably as lower standards are met. When meeting the 70, 65, and 60 ppb alternative standards, the percent of children with lung function decrements greater than or equal to 10 percent falls to less than 21 percent, less than 18 percent, and less than 14 percent in the worst O_3 year for all 15 case study urban case study areas, respectively. Mortality from short- and long-term O_3 exposures and respiratory hospitalization risk is not greatly affected by meeting lower standards, reflecting the impact of increasing O_3 on low concentration days, and the non-threshold nature of the C-R function.

9.7 REFERENCES

Frey, C. and J. Samet. 2012. *CASAC Review of the EPA's Policy Assessment for the Review of the O_3 National Ambient Air Quality Standards (First External Review Draft – August 2012).* U.S. Environmental Protection Agency Science Advisory Board, (Document number EPA-CASAC-13-003).

Medina-Ramon, M.; A. Zanobetti and J. Schwartz. 2006. "The Effect of O_3 and PM_{10} on Hospital Admissions for Pneumonia and Chronic Obstructive Pulmonary Disease: A National Multicity Study."*American Journal of Epidemiology*, 163(6), 579-588.

McDonnell, W.F.; P. W. Stewart; M. V. Smith; C. S. Kim C.S. and E. S. Schelegle. 2012. "Prediction of lung function response for populations exposed to a wide range of O_3 conditions." *Inhalation Toxicology*, 24:619-633.

Simon, H.; K. R. Baker; F. Akhtar; S. L. Napelenok; N. Possiel; B. Wells and B. Timin. 2013. "A Direct Sensitivity Approach to Predict Hourly O_3 Resulting from Compliance with the National Ambient Air Quality Standard." *Environmental Science and Technology*, 47, 2304-2313.

Smith, RL; B. Xu and P. Switzer. 2009. "Reassessing the Relationship Between O_3 and Short-term Mortality in U.S. Urban Communities." *Inhalation Toxicology*, 21: 37-61.

U.S. Environmental Protection Agency. 2007. O_3 *Health Risk Assessment for Selected Urban case study areas*. Research Triangle Park, NC: Office of Air Quality Planning and Standards. (EPA document number EPA 452/R-07-009). Available at: <http://www.epa.gov/ttn/naaqs/standards/O_3/s_O_3_cr.html>.

U.S. EPA. 2013. *Integrated Science Assessment for O_3 and Related Photochemical Oxidants: Final.* Research Triangle Park, NC: U.S. Environmental Protection Agency. (EPA document number EPA/600/R-10/076F).

World Health Organization. 2008. Part 1: *Guidance Document on Characterizing and Communicating Uncertainty in Exposure Assessment, Harmonization Project Document No. 6*. Published under joint sponsorship of the World Health Organization, the International Labour Organization and the United Nations Environment Programme. WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland.

Zanobetti, A. and J. Schwartz. 2008. "Mortality Displacement in the Association of O_3 with Mortality: An analysis of 48 Cities in the United States." *American Journal of Respiratory and Critical Care Medicine*, 177:184-189.

United States
Environmental Protection
Agency

Office of Air Quality Planning and Standards
Air Quality Strategies and Standards Division
Research Triangle Park, NC

Publication No. EPA-452/P-14-004a
February 2014