Julie E. Goodman, PhD, DABT
Gradient Corporation, Cambridge, MA

appearing on behalf of the American Petroleum Institute

Public Advisory Committee Meeting and Teleconference of the CASAC Sulfur Oxides (SO_X) Review Panel

Durham, NC
July 30, 2008

Thank you for the opportunity to speak today about the SO_2 epidemiology data. I am Dr. Julie Goodman, an epidemiologist and board-certified toxicologist at Gradient Corporation, an environmental consulting firm in Cambridge, Massachusetts. I am going to address the Integrated Science Assessment for Sulfur Oxides – Health Criteria (ISA) conclusion that there is a causal relationship between short-term SO_2 exposure and respiratory morbidity, with "clear and convincing evidence of consistency, specificity, temporal and biologic gradients, biological plausibility, and coherence" (US EPA, 2008). The ISA also says, "In the epidemiological studies, the SO_2-related respiratory effects were consistently observed in areas where the maximum ambient 24-h avg SO_2 concentrations [were] below the current 24-h avg NAAQS level of 0.14 ppm," and "[t]he evidence is suggestive but not sufficient to infer a causal relationship between short-term exposure to SO_2 and mortality" (US EPA, 2008). As I will discus over the next few minutes, because of potential exposure misclassification, measured and unmeasured confounders, and a lack of biological plausibility based on lag times and on a lack of effects observed in clinical studies, the epidemiology data are insufficient to conclude whether short-term ambient-level SO_2 exposure is associated with respiratory morbidity or mortality.

1. Exposure misclassification could have biased results towards or away from the null.

Exposure misclassification could have resulted from issues with data quality or because levels measured at central monitors were not representative of personal exposures. As stated in the ISA:

> The strength of the associations between personal exposures and ambient concentrations could also be affected by the quality of the data collected during the exposure studies. There are at least six aspects associated with the quality of the data: method precision, method accuracy (compared with FRM), percent of data above method detection limits (based on field blanks), completeness of the data collection, sample size, and soundness of

> the quality assurance/quality control procedures. Unfortunately, not all studies reported the SIX aspects of the data quality issue. (US EPA, 2008, p. 2-51).

It is also stated that "[a]mbient concentrations of SO_2 have been declining since the 1980s and are now at or very near the limit of detection of the ambient monitors in the regulatory network" (US EPA, 2008). This could bias results in either direction, as it cannot be known how far below the detection limits each person's or community's exposure lies.

With respect to time-series studies, there are other issues besides those with measurement error at low concentrations. For example, local sources may cause SO_2 to be unevenly distributed; a monitoring site may represent a nearby source and not human exposures a small distance away; terrain features and weather can affect pollution patterns; and daily variations in SO_2 concentrations at a central monitoring site differ from variations experienced by individuals living in the geographic area from which health measurements are drawn. These factors may bias results in either direction. Also, people spend most of their time indoors. Although exposures inside are less than those outdoors, the difference between indoor and outdoor measurements varies greatly by region (see ISA Figure 2-22). In a single study, overestimating exposure (by basing it on outdoor measurements) will bias results towards the null. The degree of bias will vary across studies, and no attempt has been made in the ISA to determine how risk estimates vary across studies with different ratios of indoor to outdoor SO_2 concentrations.

Overall, SO_2 exposure misclassification is likely to be non-differential. Non-differential misclassification occurs when, regardless of health status, each exposed and non-exposed subject has the same probability of being misclassified. Some individuals have mistakenly interpreted non-differential misclassification to mean that an equal fraction of subjects are misclassified in the diseased and non-diseased groups. If this indeed were to occur, then risk measures would be biased towards the null. This is because, if some percent of the "exposed" study group was actually unexposed and vice versa, then the exposure levels in the two groups would overlap. Thus, actual differences between exposed and unexposed individuals, if they existed, would appear smaller (Wacholder *et al.*, 1995). In fact, in some cases of low sensitivity and specificity, bias beyond the null can occur (Wacholder *et al.*, 1995).

By definition, "bias refers to a systematic tendency and not to a particular result" (Wacholder *et al.*, 1995). Non-differential misclassification actually means that every subject, regardless of disease

status, has an equal chance of being misclassified. This is because whether a subject is misclassified is a matter of chance. Yet, the actual fraction of subjects in a particular study who are misclassified in the diseased group is likely to differ from the fraction of subjects misclassified in the non-diseased group. Thus, even if misclassification is non-differential on average, due to random variation, misclassification rates in a single study will most likely be differential (Jurek *et al.*, 2005; 2008), and may bias results in any direction. This was demonstrated in a study by Sorahan and Gilthorpe (1994), who presented relative risks from simulated cohort studies with various degrees of non-differential misclassification. This analysis showed that a considerable percentage of studies with non-differential misclassification present produced risk estimates that were larger than those from data sets classified correctly.

According to Wacholder *et al.* (1995):

> Several papers published since 1990 have shown that there are special circumstances where there is a bias towards exaggeration of effects. Dosemeci et al. identified a scenario where non-differential misclassification of exposure more often than not leads to an overestimate of the odds ratio in an intermediate exposure category when there are more than two exposure levels. Other papers that have appeared since the textbooks cited by Sorahan and Gilthorpe were published during the 1980s, have identified circumstances where an overestimate is more likely than an underestimate. These include particular forms of non-differential misclassification when an exposure is not binary, when grouping has occurred, or when the errors in a continuous exposure are correlated with their true value.

Even though certain exposure misclassification issues will necessarily bias results towards the null, others could bias results in either direction. Because of the high likelihood of exposure misclassification and the impossibility of knowing with certainty in which way this will bias results, the epidemiology data are insufficient to determine whether short-term SO_2 exposure leads to adverse respiratory health effects.

2. Observed associations could be attributable to confounders.

In the ISA, US EPA (2008) discusses co-pollutants and their effects on the risk estimates for SO_2, emphasizing PM, NO_2, ozone, CO, vanadium, nickel, selenium, and arsenic. Throughout the text, US EPA (2008) states that several of the risk estimates are "robust" to the adjustment for these confounders. Yet several studies have calculated risk estimates a number of ways and found statistically significant associations under certain circumstances, but not others. A robust risk estimate should not be highly

dependent on the model selection – those risk estimates that were only significant using certain models may not have been indicative of a true effect.

More specifically, US EPA (2008) claims that the associations between short-term SO_2 concentration and respiratory health effects remain significant after co-pollutants were accounted for in statistical analyses. An assessment of the data demonstrates this is not necessarily the case. For example, US EPA (2008) states that the strongest evidence for an association with asthma in children is from the National Cooperative Inner-City Asthma Study (NCICAS) (Mortimer *et al.*, 2002) and the Childhood Asthma Management Program (CAMP) study (Schildcrout *et al.*, 2006). As shown in ISA Figure 3-2, however, the asthma risk in the NCICAS becomes non-significant when it was adjusted for O_3, PM_{10}, and NO_2. US EPA (2008) states that this could be a result of small sample size, but it is also likely that at least some of these pollutants are true confounders of the SO_2/asthma association.

3. Even if risk estimates are adjusted for known confounders, exposure misclassification, residual confounding, and unmeasured or unknown confounders could affect observed association between short-term SO_2 exposure and respiratory effects.

The ISA states that associations between short-term SO_2 exposure and respiratory effects are "largely positive, with several of the more precise effect estimates (suggestive of greater study power) indicating statistical significance" (US EPA, 2008). The ISA also states that most risk estimates were "robust," but, in addition to a number of non-significant risk estimates in certain statistical models, a good number of them have very large confidence intervals (CIs), indicating they were unstable. Those with smaller CIs are generally very close to 1, indicating a weak association, if any (*e.g.*, see ISA Figure 3-8). If any misclassification or confounders led to biases away from the null, the true risk estimate and 95% CI likely included 1.

Although the ISA says that most risk estimates were independent of inclusion of co-pollutants, even when risk estimates were adjusted for them, it is possible that co-pollutants were not fully accounted for, leading to what is known as "residual confounding". This means that confounding was still present after adjustment (Glymour and Greenland, 2008). If the risk estimates were adjusted for the residual confounders, it is possible, and perhaps likely, the estimates would no longer be statistically significant.

There are also unmeasured or even unknown confounders that may account for observed associations between short-term SO_2 exposure and respiratory health effects. For example, temperature, humidity, several hazardous air pollutants (HAPs), and even day-to-day variation in activity and stress may affect health risks (Bukowski, 2007; 2008a; 2008b; Goldberg *et al*. 2008). As evidence of this, US EPA (2008) reports in the ISA that some studies found associations between SO_2 and emergency department visits and hospitalizations in summer but not winter months (when SO_2 levels are higher). This indicates temperature and/or humidity may play a major role in risk. In spite of this, most risk estimates were not adjusted for these factors or several other confounders.

4. The criteria that US EPA uses for determining causality are inconsistent across endpoints.

US EPA (2008) states that the data are *sufficient* for concluding that short-term SO_2 exposures are associated with certain morbidities, but they are *suggestive but not sufficient* for determining whether SO_2 exposure is associated with mortality. For both morbidity and mortality, risk coefficient and 95% CIs – both not adjusted and adjusted for co-pollutants – were of similar orders of magnitude (*e.g.*, see Figures 3-2, 3-3, and 3-11). Also, US EPA (2008) states that effect estimates for mortality are "generally reduced after adjusting for co-pollutants in the regression models" but they do not emphasize this point for many analyses of morbidities. This disparity indicates US EPA is not using the same criteria to determine causation throughout the ISA. The criteria used for the mortality data should be used throughout the ISA. If this were done, the ISA would conclude that the data are not sufficient to establish causality.

5. Risk estimates calculated from several statistical models are not consistent or biologically plausible.

Several studies calculated risk estimates a number of ways and found statistically significant associations under certain circumstances, but not others. This includes associations calculated using a number of different lag times. For example, Schwartz *et al.* (1994) found no association between SO_2 and cough with a 0-day lag, but an increased risk with a 1-4-day lag (Schwartz *et al.*, 1994). There is no biological explanation provided for why a cough would appear with a 1-4-day lag but not a 0-day lag.

In addition, the lag time for which statistically significant findings were found varied across studies. Were SO_2 truly causing adverse health effects, the timing of these effects should be consistent across studies, and this simply is not the case.

6. Clinical studies demonstrated no significant effects at higher SO_2 concentrations than those in epidemiology studies reporting associations between short-term SO_2 exposure and respiratory effects.

The ISA states that the strongest evidence for a causal association between SO_2 and respiratory effects comes from clinical studies (US EPA, 2008):

> Collectively, evidence from earlier studies considered in the previous review, along with a limited number of new human clinical studies, consistently indicates that with elevated ventilation rates, asthmatic individuals experience moderate or greater decrements in lung function, as well as increased respiratory symptoms, following peak exposures to SO_2 at concentrations as low as 0.4-0.6 ppm (Balmes et al., 1987; Gong et al., 1995; Horstman et al., 1986; Linn et al., 1987; Linn et al., 1983)…. Some sensitive asthmatics have been shown to experience moderate decrements in lung function at concentrations below 0.3 ppm (Balmes et al., 1987; Linn et al., 1987; Sheppard et al., 1981), although there is limited evidence of a significant increase in respiratory symptoms at these exposure concentrations.

These effects are transient. As stated by US EPA (2008), a transient decrement in lung function "is not automatically considered to be an adverse effect." It should also be noted that in all studies of adolescents, SO_2 was administered *via* a mouthpiece, which bypassed nasal and nasopharynx absorption and led to an increased presentation of SO_2 into the lungs.

The ISA reports that associations reported in epidemiology studies were consistently observed in areas where the maximum ambient 24-hour average SO_2 concentrations were below the current 24-hour average NAAQS level of 0.14 ppm. These exposures are generally far below those in the clinical studies at which no significant response was seen in sensitive asthmatics (0.3 ppm), yet the ISA does not discuss this inconsistency. Most of the effects observed in the clinical studies were resolved quickly, and this is not consistent with 1- or more-day lags observed with epidemiology studies. Also, the effects reported in the epidemiology studies were more severe than those noted in the clinical trials at higher doses, which is not biologically plausible. Were SO_2 truly causing effects, one would expect to see an exposure-response relationship, or increased effects (both in severity and the number of individuals affected) at higher exposures. This is clearly not the case.

Conclusions

There have been several statistically significant and non-significant associations reported between short-term SO_2 exposure and respiratory health effects. Issues with both the quality of exposure measurements and whether these measurements represent human exposure could bias results either towards or away from the null. Furthermore, co-pollutants and other confounders, known and unknown, likely also bias results away from the null. Although this is noted for certain endpoints (*e.g.*, mortality), it is not fully accounted for regarding other endpoints (*e.g.*, respiratory symptoms). Also, risk estimates based on several statistical models are not consistent across studies or biologically plausible. Finally, associations were noted in epidemiology studies at exposure concentrations much lower than those at which no effects were seen in clinical trials. In addition, those effects seen at higher doses in clinical trials were less severe than those noted in the epidemiology studies. Thus, there is no compelling new scientific evidence for a causal association between short-term SO_2 exposure and respiratory morbidity and mortality. The epidemiology data are not sufficient to determine whether any adverse health effects result from short-term SO_2 exposure at levels below the current NAAQS for SO_X of 0.14 ppm averaged over a 24-hour period.

References

Bukowski, JA. 2007. "Do time-series studies contain residual confounding by risk factors for acute health events?" *Epidemiol.* 18:839.

Bukowski, J. 2008a. "Do pollution time-series studies contain uncontrolled or residual confounding by risk factors for acute health events?" *Regul. Toxicol. Pharmacol.* 51:135-140.

Bukowski, JA. 2008b. "Response to counterpoint: Time-series studies of acute health events and environmental conditions are not confounded by personal risk factors by Mark S. Goldberg, Richard T. Burnett, Jeffrey R. Brook." *Regul. Toxicol. Pharmacol.* 51:148-150.

Goldberg, MS; Burnett, RT; Brook, JR. 2008. "Counterpoint: Time-series studies of acute health events and environmental conditions are not confounded by personal risk factors." *Regul. Toxicol. Pharmacol.* 51:141-147.

Glymour, MM; Greenland, S. 2008. "Causal Diagrams." In *Modern Epidemiology (Third Edition).* (Eds.: Rothman, KJ; Greenland, S; Lash, TL), Lippincott, Williams, & Wilkins, Philadelphia, PA.

Jurek, AM; Greenland, S; Maldonado, G. 2008. "How far from non-differential does exposure or disease misclassification have to be to bias measures of association away from the null?" *Int. J. Epidemiol.* 37(2):382-385.

Jurek, AM; Greenland, S; Maldonado, G; Church, TR. 2005. "Proper interpretation of non-differential misclassification effects: Expectations *vs.* observations." *Int. J. Epidemiol.* 34(3):680-687.

Mortimer, KM; Neas, LM; Dockery, DW; Redline, S; Tager, IB. 2002. "The effect of air pollution on inner-city children with asthma." *Eur. Respir. J.* 19:699-705.

Schildcrout, JS; Sheppard, L; Lumley, T; Slaughter, JC; Koenig, JQ; Shapiro, GG. 2006. "Ambient air pollution and asthma exacerbations in children: An eight-city analysis." *Am. J. Epidemiol.* 164:505-517.

Schwartz, J; Dockery, DW; Neas, LM; Wypij, D; Ware, JH; Spengler, JD; Koutrakis, P; Speizer, FE; Ferris, BG, Jr. 1994. "Acute effects of summer air pollution on respiratory symptom reporting in children." *Am. J. Respir. Crit. Care Med.* 150:1234-42.

Sorahan, T; Gilthorpe, MS. 1994. "Non-differential misclassification of exposure always leads to an underestimate of risk: An incorrect conclusion." *Occup. Environ. Med.* 51(12):839-840.

Spengler, JD; Ferris, BG; Dockery, DW; Speizer, FE. 1979. "Sulfur dioxide and nitrogen dioxide levels inside and outside homes and the implications on health effects research." *Environ. Sci. Technol.* 13(10):1276-1280.

US EPA. 2008. "Integrated Science Assessment for Sulfur Oxides - Health Criteria (Second External Review Draft)." Office of Research and Development. EPA/600/R-08/047. 288p., May.

Wacholder, S; Hartge, P; Lubin, JH; Dosemeci, M. 1995. "Non-differential misclassification and bias towards the null: A clarification [Letter to the editor]." *Occup. Environ. Med.* 52(8):557.

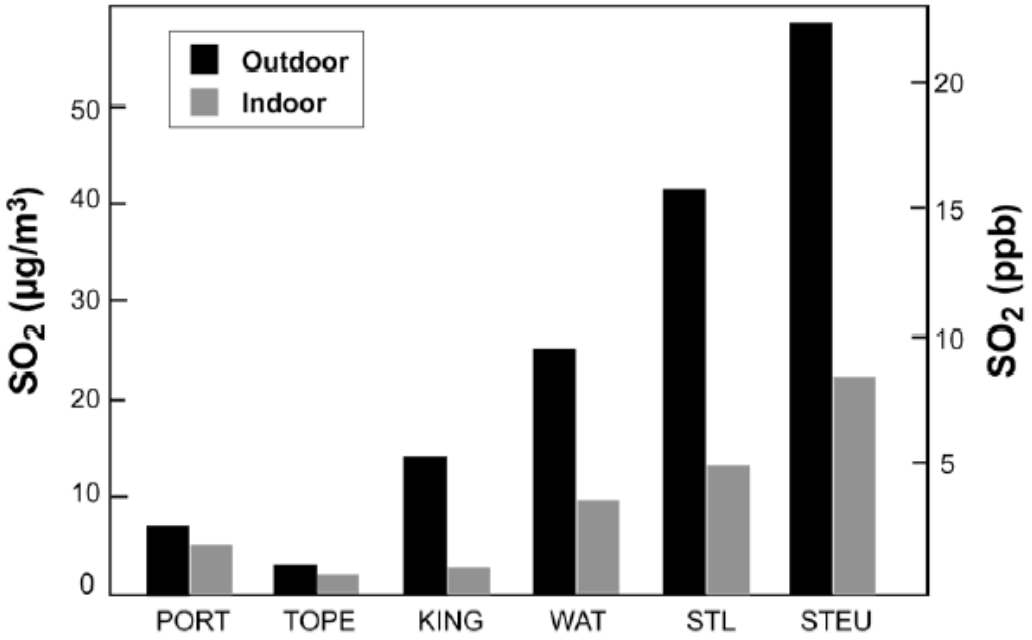

Figure 2-22. Average annual indoor and outdoor SO_2 concentrations for each of the six cities included in the analysis.

PORT = Portage, WI TOPE = Topeka, KS KING = Kingston, TN
WAT = Watertown, MA STL = St. Louis, MO STEU = Steubenville, OH.

Source: Adapted from Spengler et al. (1979).

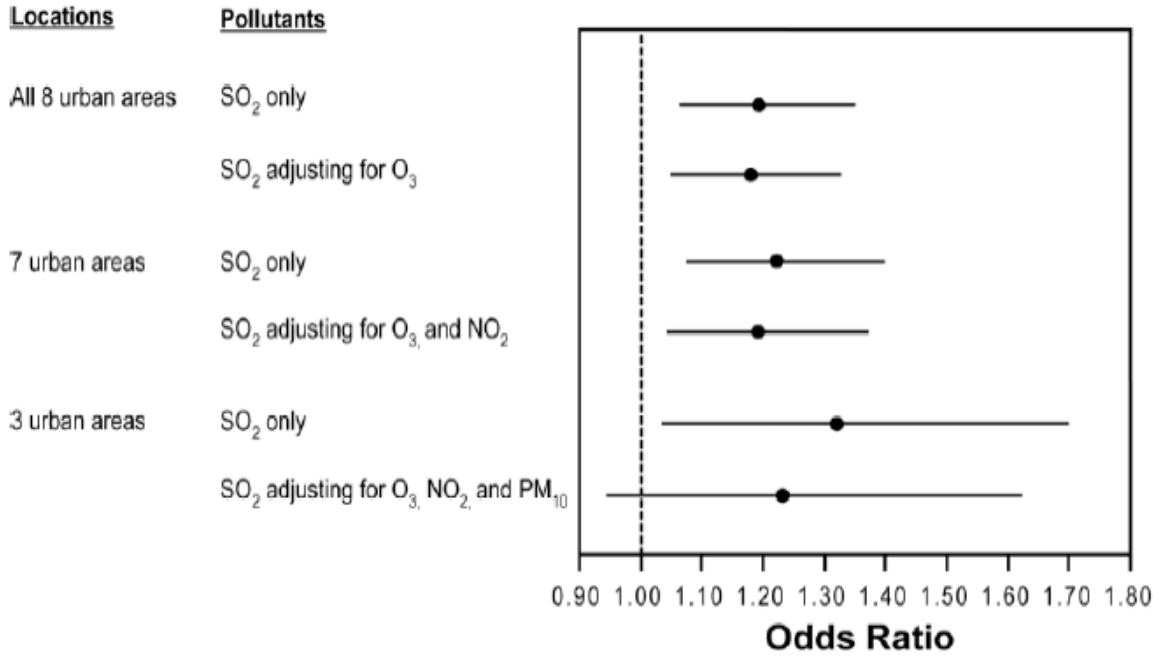

Figure 3-2. Odds ratios (95% CI) for incidence of morning asthma symptoms of 846 asthmatic children from the National Cooperative Inner-City Asthma Study. Effects associated with a 20 ppb increase in 3-h avg SO_2 with a lag of 1-2 day moving average are presented. SO_2 effect estimates from single- and multipollutant models are shown.

Source: Mortimer et al. (2002).

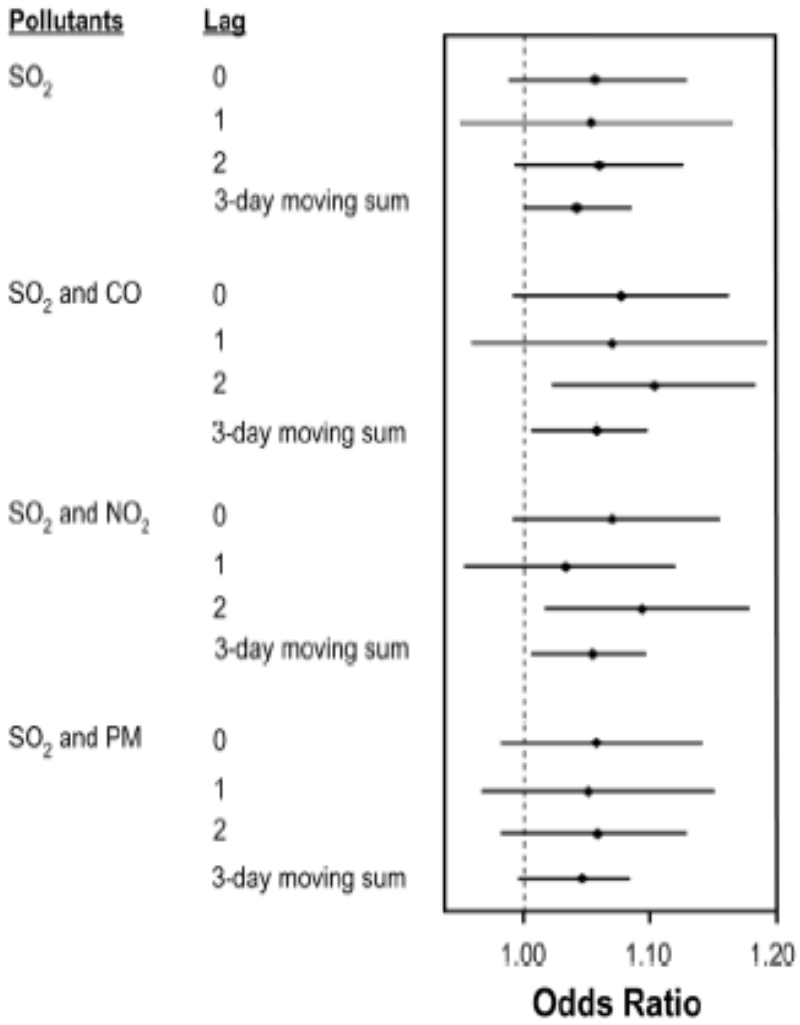

Figure 3-3. Odds ratios (95% CI) for daily asthma symptoms of 990 asthmatic children from the Childhood Asthma Management Program Study. Effects associated with a 10 ppb increase in within-subject concentrations of 24-h avg SO_2 are presented. Data collected from November 1993 to September 1995 were used. Results from single- and joint-pollutant models are shown.

Source: Schildcrout et al. (2006).

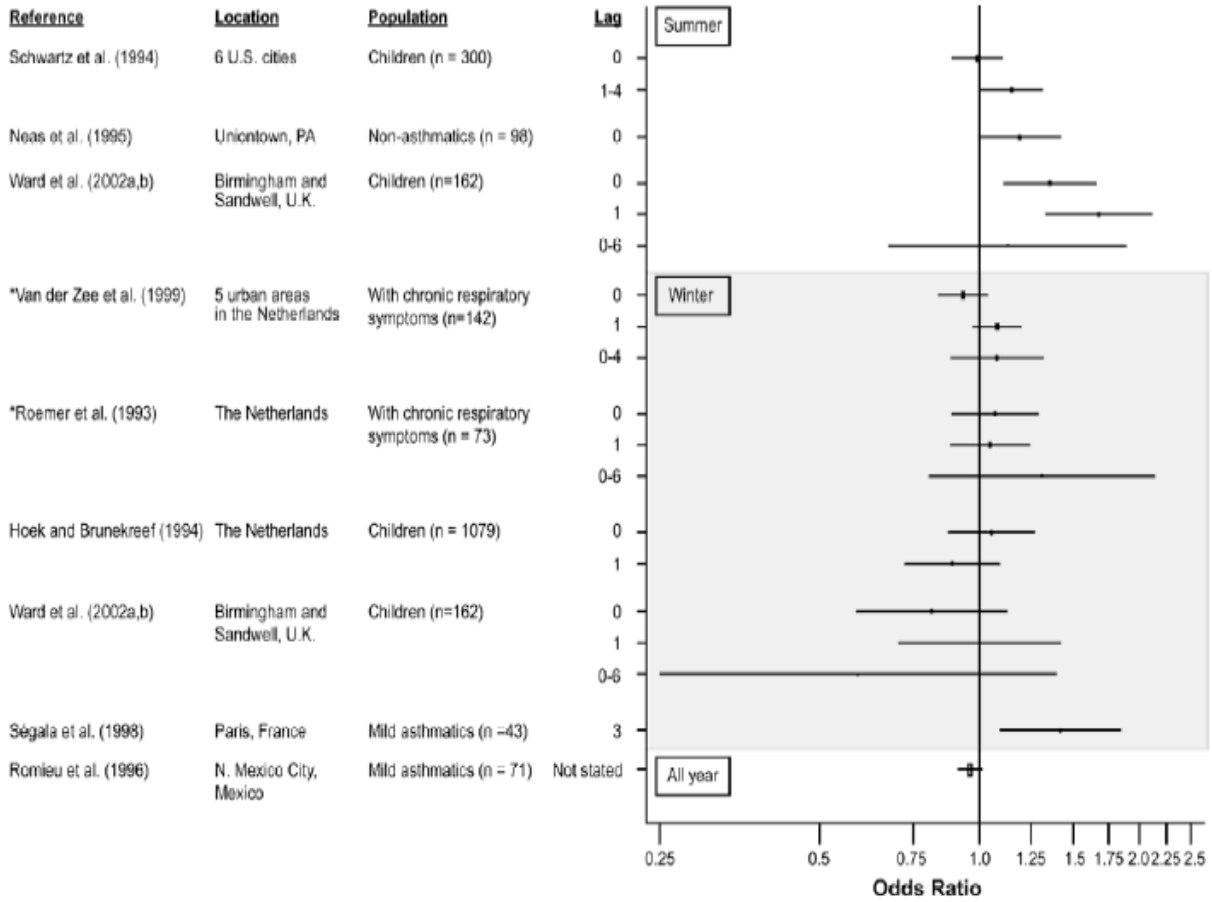

Figure 3-4. Odds ratios (95% CI) for incidence of cough among children, grouped by season. For single-day lag models, current day and/or previous day SO_2 effects are shown, except for Ségala et al. (1998), which only presented results for a 3-day lag. Multiday lag models represent the effect of the mean concentration from the range of days noted. Risk estimates are standardized per 10 ppb increase in 24-h avg SO_2 level. The size of the box of the central estimate represents the relative weight of that estimate based on the width of the 95% CI.

*Note that van der Zee et al. (1999) and Roemer et al.(1998) presented results for prevalence of cough.

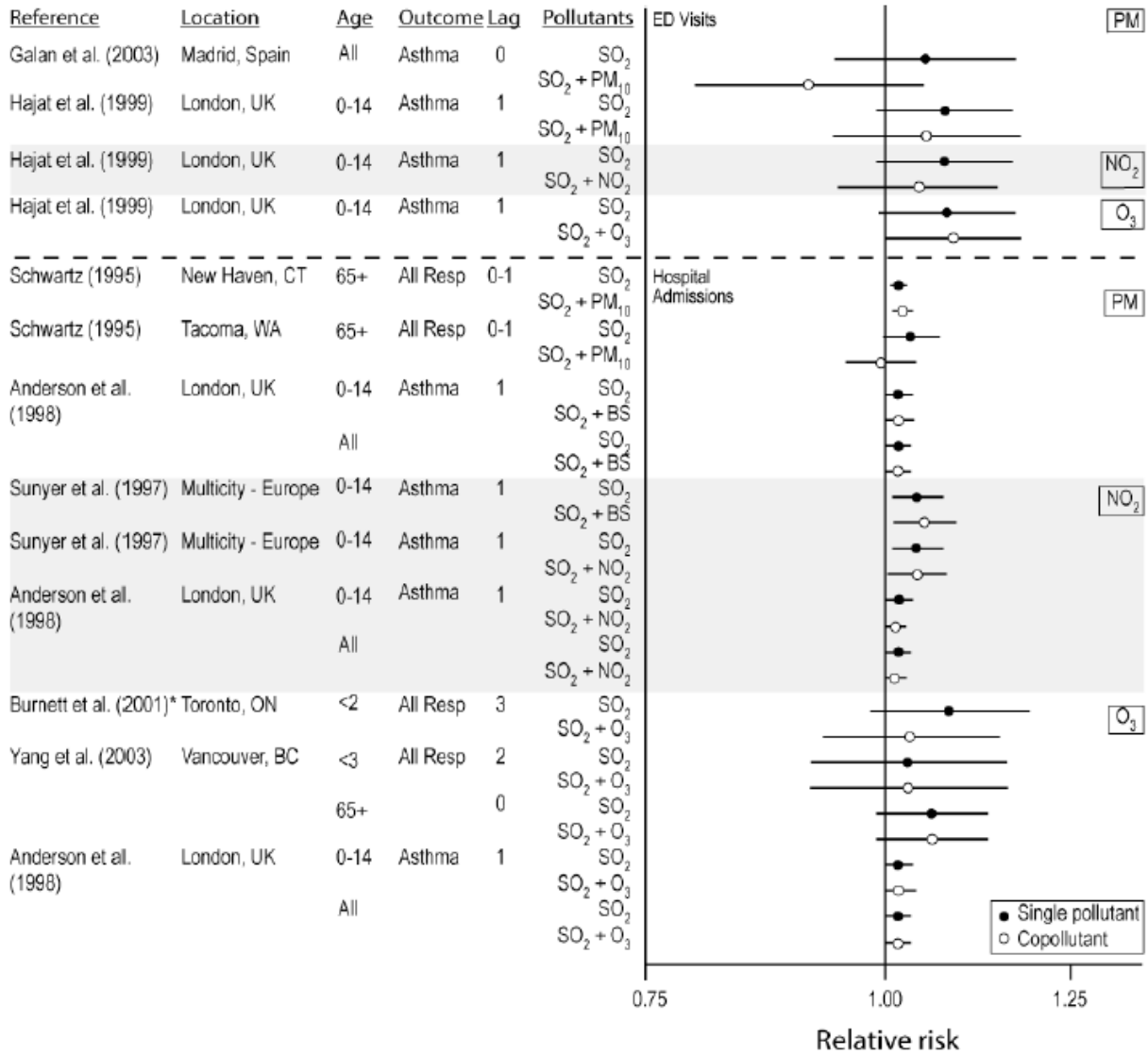

Figure 3-8. Relative risks (95% CI) of SO_2-associated emergency department visits and hospitalizations for all respiratory causes and asthma, with and without copollutant adjustment. Risk estimates are standardized per 10 ppb increase in 24-h avg SO_2 concentrations or 40 ppb increase in 1-h max SO_2. In Burnett et al. (2001), analyses were performed using default convergence criteria for Poisson GAM with a nonparametric LOESS prefilter applied to air pollution and hospitalization data.

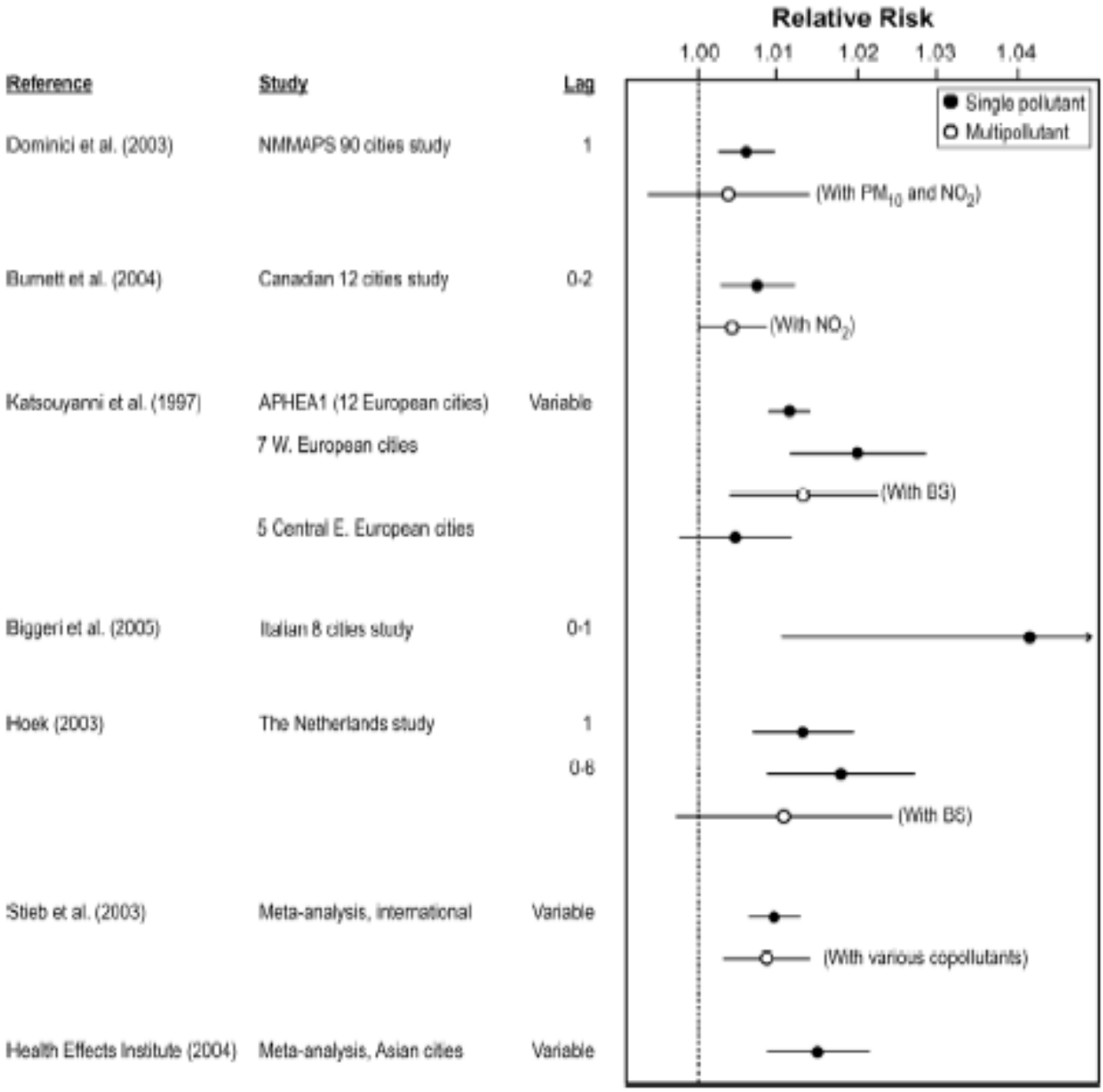

Figure 3-11. Relative risks (95% CI) of SO_2-associated all-cause (nonaccidental) mortality, with and without copollutant adjustment, from multicity and meta-analysis studies. Effect estimates are standardized per 10 ppb increase in 24-h avg SO_2 concentrations. For multipollutant models, results from the models that resulted in the greatest reduction in the SO_2 effects are shown. (NMMAPS = National Morbidity, Mortality, and Air Pollution Study; APHEA = Air Pollution and Health: a European Approach)

Biographical Summary

Julie E. Goodman, Ph.D., DABT
Principal Scientist

Dr. Julie Goodman, Director of Gradient's Epidemiology Services, is an expert in epidemiology and toxicology. Her primary responsibilities include the design, oversight, analysis, and interpretation of epidemiology studies, the evaluation of chemical toxicology data, the analysis of apparent disease clusters, and the evaluation of specific environmental chemical exposures to assess potential human health risks.

Before joining Gradient, Dr. Goodman was a Cancer Prevention Fellow at the National Cancer Institute, where she conducted a number of molecular epidemiology studies analyzing relationships between inflammatory gene polymorphisms and colon cancer risk. She was also instrumental in the development of Polymorphism Interaction Analysis, a powerful statistical tool for cancer risk assessment. While a pre-doctoral student at Johns Hopkins University, Dr. Goodman conducted epidemiological and toxicological research in estrogen metabolism and breast cancer risk. Dr. Goodman has authored several peer-reviewed toxicology and epidemiology articles.

Representative Projects

Cancer Cluster Analysis: At the request of a municipality and in response to citizens' concerns, independently investigated whether there was an increased incidence rate of cancer in residents living near a municipal landfill. Communicated findings with city officials and residents at public meetings.

Cross-sectional Study: Reviewed scientific literature on cancer and noncancer toxicity of trichloroethylene and perchloroethylene in a class-action lawsuit. Managed the quantitative analysis of ingestion and showering exposure of these solvents in groundwater. Determined whether health effects were comparable to those in communities with no known solvent exposures based on questionnaire data.

Efficacy and Toxicity Analysis: For a pharmaceutical company whose patent was being challenged, performed an independent analysis of efficacy and toxicity data to determine if claims in the patent could be challenged.

Regulatory Comment: For a trade association, contributed to comments on the 2006 "National Ambient Air Quality Standards for Particulate Matter: Proposed Rule" based on an examination of whether epidemiological research supports a 24-hour thoracic coarse particle standard ($PM_{10-2.5}$) for "urban" but not "rural" PM-coarse or a mass-based fine particulate matter standard without reference to chemical species.

Weight-of-Evidence Analysis: For a trade association, conducted a comprehensive critical weight-of-evidence review of studies bearing on the ability of very low exposures to Bisphenol A to affect reproduction and development via endocrine disruption. Testified before several state legislative committees regarding potential restrictions on Bisphenol A.

Benchmark Dose Calculations: For a pesticide company, analyzed US EPA's use of the lower confidence limit on the BMD_1 ($BMDL_1$) to determine a point of departure for cancer risk of dimethylarsenic acid in humans in a white paper submitted to US EPA.

Historical State-of-Knowledge Review: Reviewed general and company-specific historical knowledge of human and ecological toxicity of smelter contaminants in an insurance cost recovery case involving two US smelters.

Practice Areas & Expertise

- Epidemiology
- General and Molecular Toxicology
- Occupational Exposures
- Product Safety
- Carcinogenesis
- Endocrine Disruptors

Education

Ph.D., Toxicology, Johns Hopkins University

Sc.M., Epidemiology, Johns Hopkins University

S.B., Environmental Engineering, Massachusetts Institute of Technology

Diplomate of the American Board of Toxicology

Selected Publications

Goodman, JE; Gaylor, D; Beyer, LA; Rhomberg, LR; Beck, BD. 2008. "Effects of MTBE on the reported incidence of Leydig cell tumors in Sprague-Dawley rats: Range of possible Poly-3 results." *Regul. Toxicol. Pharmacol.* 50: 273–284.

Goodman, JE; Witorsch, RJ; McConnell, EE; Sipes, IG; Slayton, TM; Yu, CJ; Franz, AM; Rhomberg, LR. 2008. "Weight-of-Evidence Evaluation of Reproductive and Developmental Effects of Low Doses of Bisphenol A." *Critical Reviews Toxicol.* In press.

Goodman, JE; Mechanic, LE; Luke, BT; Ambs, S; Chanock, S; Harris CC. 2006. "Exploring SNP-SNP interactions and colon cancer risk using polymorphism interaction analysis." *Int. J. Cancer* 118(7):1790-7.

Goodman, JE; Bowman, E; Chanock, S; Harris, CC. 2004. "Arachidonate lipoxygenase (ALOX) and cyclooxygenase (COX) polymorphisms and colon cancer risk." *Carcinogenesis* 25(12):2467-72.

Goodman, JE; Jensen, LT; He, P; Yager, JD. 2002. "Characterization of human soluble high and low activity catechol-O-methyltransferase catalyzed catechol estrogen methylation." *Pharmacogenetics* 12:517-28.

Gradient CORPORATION

20 University Road
Cambridge, MA 02138
Phone (617) 395-5000
Fax (617) 395-5001

jgoodman@gradientcorp.com
www.gradientcorp.com