COMMENTS BY THE TEXAS COMMISSION ON ENVIRONMENTAL QUALITY REGARDING THE DRAFT INTEGRATED REVIEW PLAN FOR THE NATIONAL AMBIENT AIR QUALITY STANDARDS FOR PARTICULATE MATTER

EPA DOCKET ID NO. EPA-HQ-OAR-2015-0072

I. Summary of Proposed Action

On April 19, 2016, the United States Environmental Protection Agency (EPA) published in the *Federal Register* (81 FR 22977) notice of the availability and public comment period for the Draft Integrated Review Plan for the National Ambient Air Quality Standards for Particulate Matter.

The Integrated Review Plan (IRP) outlines the EPA's plan for reviewing and analyzing available scientific literature related to particulate matter (PM) in order to determine whether a revision of the current primary and secondary national ambient air quality standards (NAAQS) are necessary. The draft IRP includes a review of decisions made in the setting of the previous PM NAAQS, key guiding questions and issues to be evaluated in the upcoming review, and a schedule for the subsequent technical documents that will support either the retention of the existing NAAQS or the setting of a new NAAQS.

II. General Comments

The Texas Commission on Environmental Quality (TCEQ) appreciates the EPA's plan for approaching the upcoming review of the PM NAAQS. The draft IRP is generally written in a way that indicates a balanced approach will be taken when considering the available evidence. This unbiased approach is essential to the review and setting of all NAAQS, but is especially important in the setting of the PM NAAQS. The dominant justification for many rules that the EPA has promulgated in recent years has been the reduction in health effects attributed to PM exposure and the resulting many thousands of lives and billions of lost dollars saved [1-4]. This PM review cycle is the EPA's opportunity to break ground in our understanding of PM, not just by summarizing new studies and drawing conclusions that are caveated by the same concerns and uncertainties articulated in the last review, but by directly addressing previously identified issues and uncertainties. Because of the importance and far-reaching impacts of the PM NAAQS, it is crucial that this review be complete, balanced, transparent, and objective. The anticipated schedule provided in the draft IRP is ambitious and should not be expedited. In the end, a standard based on a thorough, balanced, and objective evaluation is far more protective of public health than a standard released to meet a deadline that contains numerous errors and an incomplete discussion of uncertainty and bias.

A balanced review should begin with a balanced overarching question that informs and governs the entire review. The current overarching question presented in the draft IRP ignores the possibility that new information could weaken the support for the current NAAQS, focusing instead on only maintaining or strengthening the NAAQS by detailing that the review will assess whether "the currently available scientific evidence and exposure-/risk-based information *support[s] or call[s] into question the adequacy of* …the current primary $PM_{2.5}$ or PM_{10} standards". Instead, a balanced overarching question should be "[d]oes the currently available scientific evidence and exposure-/risk-based information detract from, support, or call into question the adequacy of the public health protection afforded by the current primary $PM_{2.5}$ or

PM_{10} standards?" A truly balanced review should address both positive and negative associations relating to the current NAAQS.

The TCEQ applauds the EPA's commitment to addressing key uncertainties identified during the last review of the PM NAAQS. Issues related to component toxicity, multipollutant mixtures, exposure measurement error, nature and magnitude of risk at low concentrations of PM, heterogeneity of responses within and between cities and regions, and general uncertainty regarding potential impacts of PM on reproductive and developmental endpoints are particularly important to understanding what level is requisite to protect public health and what is an adequate margin of safety. Evaluation of the scientific literature without pre-suppositions about important aspects such as causality and shape of concentration-response functions is crucial to giving this review the sound scientific foundation it needs. The TCEQ strongly encourages the EPA to better use the IRP to more clearly articulate how the EPA will evaluate these topics in its assessment documents and, whenever possible, quantitatively evaluate the impact these uncertainties have on final risk values. To address some of these concerns, the TCEQ encourages the EPA to consider the following methods, many of which are based on existing peer-reviewed work, including:

- quantitative methods for causal analysis [5-12];
- transparent evidence-based integration methods using state-of-the-field systematic review techniques that include the documentation of clear inclusion/exclusion criteria, risk of bias analysis, and study quality scoring criteria [13-15];
- frank and transparent consideration of weight of evidence, including a system for evaluating and weighing results from studies showing positive, negative, and no effects;
- focused assessment of mode of action and biological plausibility of health effects using studies done at ambient or near-ambient concentrations (far less than the proposed 2 mg/m^3);
- quantitative uncertainty analysis methods, as recommended by the National Research Council (NRC) [16] and discussed in, though absent from, previous NAAQS reviews [17];
- methods using confidence bounds on risk estimates that do not just include statistical uncertainty, but also other uncertainties discussed in these comments;
- quantitative methods for consideration of exposure measurement error in epidemiology study models [18, 19], ;
- quantitative methods for evaluation of population-level concentration-response thresholds in the framework of standard risk assessment methods to include consideration of biological mode of action and shortcomings in the data (exposure measurement error, confounders) that can bias the statistical model; and
- evaluation methods using region-specific concentration-response coefficients and uncertainty bounds, rather than a national concentration-response coefficient that is known to not represent all areas of the country [20-25].

Finally, the TCEQ encourages the EPA to articulate its commitment to transparent policy decisions in the IRP. Decisions based on scientific judgment are made at many stages of the NAAQS review process. This includes the interpretation of the primary literature, determining inclusion/exclusion study criteria, judgments regarding causality and various forms of uncertainty, as well as in the design of the subsequent risk assessment. These decisions, as well as the underlying policy determinations, are often critical in setting the final NAAQS and are amplified when later providing for an additional margin of safety. Because both policy and

scientific judgment decisions are integral to the NAAQS setting process [26, 27], the TCEQ encourages the EPA to clearly articulate when and how these judgments will be made in the upcoming PM NAAQS review.

III. Specific Technical Comments

Evidence Integration

The TCEQ encourages the EPA to provide greater clarity and objectivity regarding the criteria for study selection and inclusion. The draft IRP indicates that study inclusion will be "based on the extent to which the study is potentially policy-relevant and informative" (Draft IRP page 3-11 [28]). The EPA also notes that "conclusions about the strength of inference from study results will be made by weighing the authors' conclusions and independently evaluating study quality" (Draft IRP page 3-11 [28]). A more transparent and objective approach would use systematic review methods to outline specific search terms and document inclusion/exclusion criteria, including clarification on the inclusion/exclusion of studies with different PM size and component measurements [13-15]. The EPA should then clearly articulate how study quality conclusions will be used; for example, if studies will be excluded from further analysis or how they will be further critically evaluated. The TCEQ also encourages the EPA to clarify how the epidemiology data will be integrated with controlled exposure, and toxicological data in order to determine the plausibility of effects reported for low levels of ambient $PM_{2.5}$.

The EPA should also use the IRP to clarify the treatment of studies that reanalyze existing data. The TCEQ agrees that there should be a careful evaluation of new evidence that has been published since the last review. However, the available epidemiologic literature largely consists of reanalysis of existing cohorts, which does not necessarily constitute *new* evidence. For example, the last PM review largely relied on results from the Harvard Six Cities studies as well as the American Cancer Study. Specifically Krewski *et al.* 2009 and Lepeule *et al.* 2012 were used as key studies, which are updates of previous research conducted by Pope *et al.* 2002 and Laden *et al.* 2006 [29-32].

The TCEQ supports the plan put forth in the draft IRP to evaluate studies based on their design, methods, conduct, and documentation, rather than whether the study results were positive, negative, or null. The TCEQ encourages equal weighing of the available evidence, regardless of whether that evidence is positive, negative, or null.

Particulate Matter Heterogeneity

Particle Composition

The EPA should place significant focus on the impact of individual PM components on measured health endpoints. The heterogeneity of PM composition across cities, regions, and seasons is well-documented. PM component data also informs biological plausibility and mode of action; therefore, PM composition likely has significant effects on health endpoints noted in published literature, even if individual species are not identified. It is generally acknowledged that not all species of PM are equally toxic, with metals being generally more toxic (reviewed in Chen and Lippman 2009 [33]) and nitrates generally less so (reviewed in Schlesinger 2007 [34]). Generally, organic and elemental carbon, iron, and ions such as nitrate and sulfate, are the most abundant PM components, which are less toxic species [33, 35]. In particular, Valberg (2004) noted that "none of the chemicals potentially present in PM has an [inhalation reference concentration] sufficiently low to predict serious health effects at the relevant concentrations" [38].

Careful consideration of PM composition, then, is critical to accurately understanding health effect associations and population-level risks for the setting of a standard, even if the body of

literature is insufficient to support individual PM species standards. Numerous studies have emphasized the problem with treating all particles with an aerodynamic diameter of ≤ 2.5 μm or ≤ 10 μm as equally toxic. For example, Valberg (2004) noted that approximately 90% of the $PM_{2.5}$ mass from St. Louis, MO, samples had components that were not toxic enough to lead to mortality [36]. The author suggested that oxygen accounted for the other 10% of the $PM_{2.5}$ mass composition, but did not specifically discuss the potential toxicity of it or the metallic oxides it likely formed. Further, in a study of PM and mortality from 1987 to 2000, Dominici *et al.* (2007) found that "the day-to-day association between particulate matter and mortality is getting weaker over time, possibly as a result of changes in the composition and toxicity of the particulate matter" [37]. This change in association may continue to change, as future mitigation efforts include PM reductions from non-point sources that are major sources of nitrate PM [38]. Most data suggest that there are likely no adverse health effects from current levels of nitrate PM [34, 36, 39]. Therefore, not all PM mass reductions would provide the same public health benefits. For these reasons, directly addressing the issue of PM composition is critical to attaining meaningful ambient reductions and health benefits, even though setting a speciated PM standard is admittedly problematic.

In the draft IRP, the EPA asks if "the evidence support[s] an alternative approach for defining particle pollution, including in terms of … specific components." While we are aware of the difficulties with regulating PM components separately, targeting specific PM components or emission sources would provide better public health protection and more efficient regulation than regulating total PM [40]. This suggestion is consistent with previous advice from the Clean Air Scientific Advisory Committee (CASAC) that urged the EPA to investigate new indicators that may be more directly linked to the health and welfare effects [41]. Additionally, the TCEQ encourages the EPA to set a more precise indicator for the standard, making it the portion of PM most likely to cause health effects. The form of the standard should be determined by subtracting or giving lower weight to known nontoxic species from total PM mass measurements. If in this review the EPA again finds it impossible to set a speciated PM NAAQS, the TCEQ encourages the EPA to at least begin the process of quantifying the impact of speciated PM based on available data. Such actions may in turn encourage more research in this crucial field and help guide more meaningful understanding and regulation of local ambient PM conditions.

Particulate Matter Concentration

In addition to PM composition, ambient PM concentrations are quite heterogeneous across regions and seasons of the year. As detailed in the section below, information about this heterogeneity should be applied to the question of whether PM is *causing* premature mortality and what effect, if any, PM composition has on the regional differences in concentration-response coefficients. Further, the IRP should also include the key question of whether seasonal changes in PM concentration or composition has any effect on health effect endpoints.

The TCEQ also encourages the EPA to take this opportunity to expand the analysis of regional heterogeneity of effect estimates into a more sophisticated approach to better account for uncertainty across regional estimates. Because BenMAP already models pollutant concentrations by geographic location, the most logical solution would be to utilize region-specific concentration-response coefficients (and their confidence bounds) to obtain more accurate information about health risks. Should the EPA decide to model national versus regional results using BenMAP, an in-depth discussion of the transferability of concentration-response coefficients between cities and regions should be provided.

Causality

Causal Framework
The TCEQ strongly encourages both the CASAC and the EPA to reevaluate the strength of the existing causal framework. Of particular concern is whether a single, positive result is adequate to make the determination of "suggestive of a causal relationship," as is currently represented in the framework. Recently, various groups have proposed a number of suggestions for improvement in this area, including inclusion and exclusion criteria, study quality scoring criteria, risk of bias, and integration of evidence across multiple types of data [13-15, 42]. An improvement in the inclusion and transparency of these considerations will significantly improve confidence in the EPA's causal conclusions. After this evaluation, the TCEQ encourages a careful and consistent application of the framework for causality to this and other NAAQS reviews.

Quantitative Determinations of Causality
The EPA should clearly and objectively articulate its method for assessing causality, with particular emphasis on quantitative methods, before beginning the literature review for the assessment documents. The draft IRP states that the last PM review found a "strong and generally robust body of evidence of serious health effects associated with both long- and short-term exposure to $PM_{2.5}$... [with an] overall pattern across a broad range of studies reporting positive associations, which were frequently statistically significant." However, the draft IRP does not discuss whether such statistical associations are causal. Further, the draft IRP states that "scientists" will consider a number of issues in judging causality, but does not indicate specifically who those "scientists" will be, what points the EPA will require the "scientists" to consider, or how those judgments will be made (preferably through analysis of stated objective criteria). The EPA should consider recently proposed quantitative techniques, such as Granger causality tests and Quasi-Experimental approaches, to determine whether the association between exposure and effect is indeed causal in nature prior to making such statements characterizing the strength of the literature [5-12]. Failure to do so will compromise the objectivity and reliability of subsequent analyses.

Biological Plausibility
The EPA should pay particularly close attention to the biological plausibility of the health endpoints it evaluates in the assessment documents and better articulate how biological plausibility will be addressed in the IRP. For example, the draft IRP indicates that diabetes as well as other metabolic diseases and/or endocrine system effects will be evaluated, but does not include discussion of potential mechanisms whereby *ambient* concentrations of inhaled PM could cause such effects. Moreover, the draft IRP indicates that inhaled PM can translocate to the brain and cardiovascular system and cause systemic health effects. It will be critical to fully review and evaluate this information in the context of toxicokinetics/toxicodynamics (e.g. particles crossing the blood-brain barrier) in order to address the biological plausibility of these findings. Systemic inflammation is also often cited as a possible mechanism whereby inhaled PM could have extra-thoracic effects. This evidence should be evaluated within the framework of the Bradford Hill considerations [43], especially the consideration for a given health effect to be specific to the exposure in question.

The TCEQ also encourages the EPA to reconsider the range of PM concentrations it considers relevant to evaluating modes of action and the impact this range has on biological mechanisms underlying health endpoints. The draft IRP indicates that PM concentrations of 2 **mg**/m^3 or higher may be used when evaluating modes of action. It is unclear that the effects observed in experimental animals at these very high concentrations would reasonably be anticipated to occur in humans exposed to ambient levels of PM, which are orders of magnitude lower. Also, it

is possible that such high concentrations would act via mechanisms that may be quite different than those relevant to ambient levels of PM (this is a concept called dose-dependent transitions in mechanisms of toxicology [44]). The TCEQ encourages the EPA to include a robust discussion of how dose-response and mechanisms of action may differ across the full range of exposure concentrations considered in the Integrated Science Assessment. Because of the extensive dataset of PM mechanisms, the EPA should separately consider the mechanisms demonstrated by those studies done at lower, ambient-relevant concentrations to determine applicable mode of action for PM epidemiology studies.

Study Limitations
The TCEQ encourages the EPA to use the IRP to document how it will treat studies with significant limitations in study design. For example, in the previous PM NAAQS review, evidence supporting revision of the PM NAAQS was drawn mainly from observational studies. However, this type of data is not intended to be used alone for inferring causal relationships and it is highly sensitive to statistical modeling choices and confounding [45-47]. The TCEQ strongly encourages the EPA to exclude studies of insufficient quality from causal determinations.

Shape of the PM dose-response curves

For those health effects with enough data to suggest a causal link, the EPA should discuss at length the shape of the PM dose-response curves. This is crucial for appropriately assigning the risks and benefits ascribed to this pollutant, and is the foundation for the remainder of the NAAQS review and resulting choice of alternative standards. The TCEQ encourages the EPA to use the IRP and subsequent assessment documents to fully discuss not just whether there is a threshold for $PM_{2.5}$ effects, but also any evidence for linear or supralinear concentration-response functions, and to model alternative choices. Presenting this information will more clearly and transparently communicate assumptions affecting risks and benefits in the absence of other confounding uncertainties.

The TCEQ also strongly encourages the EPA to redirect a key question in the draft IRP regarding whether epidemiologic studies indicate departures from linearity at low concentrations (i.e. levels below the existing standards) (Draft IRP page 2-16 [28]). Restricting this evaluation to epidemiology studies is inappropriate for two reasons. First, in the traditional risk assessment process, a non-carcinogenic chemical is considered to have an effect threshold unless the mode of action demonstrates otherwise. Therefore, supporting mode of action data must be available to justify the choice of a non-threshold model. The epidemiology studies in previous PM reviews, however, generally assumed a linear dose-response shape [30]. Relying only on epidemiology studies, then, tacitly assumes a linear dose-response shape without consideration of the necessary mode of action data to support such a model choice. As discussed above, it will be very important in this review to use mechanistic data obtained at relevant ambient or near-ambient concentrations to demonstrate this mode of action, particularly for serious health effects such as mortality. The second reason this restricted analysis is inappropriate is that epidemiology studies themselves contain errors and biases that can mask the presence of a threshold, and can bias the shape of the dose-response curve. These include the presence of regional heterogeneity [17] and exposure measurement error [48]. Exposure measurement error, in particular, can result in an exaggeration of risks at low concentrations and tend to make a linear response appear supralinear [49, 50]. Again, the EPA should base decisions on the known specific mechanisms of PM within the framework of risk assessment practice. The EPA should also include a thorough discussion of whether a supralinear concentration response function is biologically plausible for PM, and if so, what specific mechanisms support this determination.

One of the key uncertainties identified by CASAC in the last PM NAAQS review was the range of concentrations associated with observed health effects in the epidemiological studies, and the degree of certainty in effects at the lower concentrations, which can be informed by confidence intervals around such relationships. EPA has previously stated that "these analyses do not provide evidence of a concentration below which the confidence interval becomes notably wider and uncertainty in a concentration-response relationship substantially increases" [51]. However, CASAC disagreed with that assessment and encouraged EPA to integrate the available concentration-response confidence bounds with concentration distributions from available studies when arriving at a range of standard levels for consideration [41]. We applaud the significant improvements made in this area in the final PA, and note that the final PA was updated to state that confidence intervals around concentration-response coefficients "can help inform at what $PM_{2.5}$ concentrations we have appreciably less confidence in the nature of the underlying [concentration-response] relationship" [52]. However, it was determined that the available evidence was too limited to serve as the basis for identifying alternative standards. We anticipate this issue will also be raised during the current review, and encourage the EPA and CASAC to provide in-depth peer review and lengthy discussion on this important topic.

Exposure Measurement Error

The draft IRP indicates that there are uncertainties when extrapolating from stationary monitoring to personal exposure. A wide range of estimated correlations between personal and ambient $PM_{2.5}$ concentrations have been reported [53], leading to a substantial potential for biased effect estimates in studies of $PM_{2.5}$-mediated health effects [54, 55]. Exposure measurement error is a complex, multi-faceted problem. The TCEQ highly recommends that the EPA use adjustment factors or another quantitative method in its analysis to better account for the uncertainties it correctly identifies with regard to exposure measurement error and available epidemiology literature. We recommend against the temptation to simplify the results of this error to the statement that it biases risk estimates toward the null [56]. This is only the case with simple single-pollutant studies where (1) the concentration-response is *genuinely* linear [57], (2) measured concentrations are good surrogates for personal exposure, and (3) differences between the measured and the personal exposures are constant [49]. Instead, we strongly recommend that the EPA meet this long-standing error head-on, by undertaking analytical methods designed to begin to address this problem [18, 19, 58].

The TCEQ also strongly encourages the EPA to give close consideration to the differences between monitored and modeled PM data. Given that the eventual risk analysis and benefits calculations will rely, at least in part, on modeled data in BenMAP, whereas much of the observational data uses monitored data, it is crucial to understand how the two relate to one another as well as how they relate to actual personal exposure, in order to improve human PM exposure models.

Finally, the TCEQ urges the EPA to consider the benefits of conducting a quantitative, microenvironmental exposure assessment during this review cycle. No previous PM NAAQS has included this level of exposure assessment. In so doing, the EPA should update many of the elements that would be covered in this exposure assessment (e.g. time-activity patterns, data on indoor and outdoor PM concentrations and trends, and information on sensitive populations) because they are over ten years old. Each of these elements is vital to a scientifically sound risk assessment. Further, because PM is so heterogeneous, it is important that this assessment specifically include evaluation of certain microenvironments that are likely to have disproportionately high PM (e.g. dense urban canyons susceptible to elevated vehicular PM and locations near large point or natural PM sources), as well as environments that are removed from these sources. Failure to include this assessment will lead to a biased risk assessment that will inappropriately influence the NAAQS review analysis.

Risk Assessment

The TCEQ strongly encourages the EPA to resolve known issues with the BenMAP software prior to utilizing the tool in the present NAAQS review. The EPA relies heavily on BenMAP to model population exposure, health risk, and health benefits to Americans as part of the NAAQS review process. However, the EPA discloses some of the software's known issues on its website[1], including an acknowledgement that monitoring data included in BenMAP-CE is badly out of date. Since states are required to certify air quality monitoring data annually, recent data are readily available to the EPA and should be used in this evaluation. The EPA requires that states utilize the most recent monitor and emission inventory data in developing state implementation plans, and should hold itself to the same standard. Many of the remaining issues are similarly resolvable and the TCEQ encourages the EPA to proactively improve this model for the important risk assessment phase of this NAAQS review.

As noted in a previous section, the TCEQ also strongly recommends that EPA include all available risk estimates in BenMAP. BenMAP presently contains a narrow range of selected concentration-response coefficients from a few key studies, rather than incorporating all of the available estimates. This produces a range of risk estimates that is narrower than would be derived using the full literature [59] and can lead to faulty conclusions about true PM exposure risk. Use of all available data will assist in ensuring a balanced review.

Uncertainty Analysis

The TCEQ encourages the EPA to use the IRP to more fully characterize how uncertainty will be addressed in the upcoming NAAQS review. As the draft IRP describes, the NAAQS risk assessment process characterizes not only the quantitative magnitude of selected risks, but also informs the public and decision makers about the confidence that alternative standards will reduce this risk. In its 2002 review, the NRC highlighted the importance of quantifying and communicating uncertainty surrounding these estimates [16]. It has also been noted that setting a NAAQS involves policy judgments as well as scientific input, and uncertainties are a key part of the context that the Administrator requires for making such decisions [60]. The TCEQ encourages a robust, quantitative consideration of uncertainty, as was recommended by the NRC, and accurate communication of all such information to the public, CASAC, and the Administrator throughout the review cycle, including using both statistical and uncertainty bounds on point estimates of risk.

[1] https://www.epa.gov/sites/production/files/2015-03/documents/benmap_ce_1.1_known_issues_-_03-04-2015_0.pdf

REFERENCES

1. OMB, *2012 Report to Congress on the Benefits and Costs of Federal Regulations and Unfunded Mandates on State, Local, and Tribal Entities*, O.o.I.a.R. Affairs, Editor 2012.
2. USEPA, *Regulatory Impact Analysis of the Final Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone*, 2015: September. p. 480 p.
3. USEPA, *Regulatory Impact Analysis for the Clean Power Plan Final Rule* O.o.A.a. Radiation, Editor 2015: August. p. 343 p.
4. USEPA, *Regulatory Impact Analysis for the Final Mercury and Air Toxics Standards*, H.a.E.I.D. Office of Air Quality Planning and Standards, Editor 2011.
5. Cox, L.A., Jr. and D.A. Popken, *Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States?* Ann Epidemiol, 2015. **25**(3): p. 162-73.
6. Cox, L.A., Jr., D.A. Popken, and D.W. Berman, *Causal versus spurious spatial exposure-response associations in health risk analysis*. Crit Rev Toxicol, 2013. **43 Suppl 1**: p. 26-38.
7. Freedman, D.A., *Graphical models for causation, and the identification problem*. Eval Rev, 2004. **28**(4): p. 267-93.
8. Eichler, M. and V. Didelez, *On Granger causality and the effect of interventions in time series*. Lifetime Data Anal, 2010. **16**(1): p. 3-32.
9. Gilmour, S., et al., *Using intervention time series analyses to assess the effects of imperfectly identifiable natural events: a general method and example*. BMC Med Res Methodol, 2006. **6**: p. 16.
10. Dash, D. and M.J. Druzdzel, *A note on the correctness of the causal ordering algorithm*. Artif Intell, 2008. **172**: p. 1800-8.
11. Lipsitch, M., E. Tchetgen Tchetgen, and T. Cohen, *Negative controls: a tool for detecting confounding and bias in observational studies*. Epidemiology, 2010. **21**(3): p. 383-8.
12. Dominici, F., M. Greenstone, and C.R. Sunstein, *Science and regulation. Particulate matter matters*. Science, 2014. **344**(6181): p. 257-9.
13. Goodman, J.E., et al., *Evaluation of the causal framework used for setting national ambient air quality standards*. Crit Rev Toxicol, 2013. **43**(10): p. 829-49.
14. Rhomberg, L.R., et al., *A survey of frameworks for best practices in weight-of-evidence analyses*. Crit Rev Toxicol, 2013. **43**(9): p. 753-84.
15. Collaboration, T.C., *Cochrane Handbook for Systematic Reviews of Interventions*, ed. J. Higgins and S. Green. Vol. Version 5.1.0. 2011.
16. NRC, N.R.C., *Estimating the Public Health Benefits of Proposed Air Pollution Regulations*, 2002, The National Academies Press: Washington, D.C.
17. USEPA, *National Ambient Air Quality Standards for Ozone (Proposed rule)*. 2014. p. 40 CFR Parts 50, 51, 52, 53 and 58.
18. Lash, T.L., et al., *Good practices for quantitative bias analysis*. Int J Epidemiol, 2014. **43**(6): p. 1969-85.
19. Sheppard, L., et al., *Confounding and exposure measurement error in air pollution epidemiology*. Air Qual Atmos Health, 2012. **5**(2): p. 203-216.
20. Samet, J.M., et al., *The National Morbidity, Mortality, and Air Pollution Study. Part II: Morbidity and mortality from air pollution in the United States*. Res Rep Health Eff Inst, 2000. **94**(Pt 2): p. 5-70; discussion 71-9.
21. Dominici, F., et al., *Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases*. JAMA, 2006. **295**(10): p. 1127-34.
22. Bell, M.L., et al., *Seasonal and regional short-term effects of fine particles on hospital admissions in 202 US counties, 1999-2005*. Am J Epidemiol, 2008. **168**(11): p. 1301-10.

23. Samoli, E., et al., *Acute effects of ambient particulate matter on mortality in Europe and North America: results from the APHENA study*. Environ Health Perspect, 2008. **116**(11): p. 1480-6.
24. Franklin, M., P. Koutrakis, and P. Schwartz, *The role of particle composition on the association between PM2.5 and mortality*. Epidemiology, 2008. **19**(5): p. 680-9.
25. Zanobetti, A. and J. Schwartz, *The effect of fine and coarse particulate air pollution on mortality: a national analysis*. Environ Health Perspect, 2009. **117**(6): p. 898-903.
26. USEPA, *National Ambient Air Quality Standards for Particulate Matter (Final Rule)*, 2012.
27. *State of Mississippi v Environmental Protection Agency*, in *Mississippi v EPA*2013, District of Columbia Circuit Court.
28. USEPA, *Draft Integrated Review Plan for the National Ambient Air Quality Standards for Particulate Matter*, N.C.f.E. Assessment, Editor 2016: Research Triangle Park.
29. Krewski, D., et al., *Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality*. Res Rep Health Eff Inst, 2009(140): p. 5-114; discussion 115-36.
30. Lepeule, J., et al., *Chronic exposure to fine particles and mortality: an extended follow-up of the Harvard Six Cities study from 1974 to 2009*. Environ Health Perspect, 2012. **120**(7): p. 965-70.
31. Pope, C.A., 3rd, et al., *Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution*. JAMA, 2002. **287**(9): p. 1132-41.
32. Laden, F., et al., *Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study*. Am J Respir Crit Care Med, 2006. **173**(6): p. 667-72.
33. Chen, L.C. and M. Lippmann, *Effects of metals within ambient air particulate matter (PM) on human health*. Inhal Toxicol, 2009. **21**(1): p. 1-31.
34. Schlesinger, R.B., *The health impact of common inorganic components of fine particulate matter (PM2.5) in ambient air: a critical review*. Inhal Toxicol, 2007. **19**(10): p. 811-32.
35. Aneja, V.P., et al., *Characterization of major chemical components of fine particulate matter in North Carolina*. J Air Waste Manag Assoc, 2006. **56**(8): p. 1099-107.
36. Valberg, P.A., *Is PM more toxic than the sum of its parts? Risk-assessment toxicity factors vs. PM-mortality "effect functions"*. Inhal Toxicol, 2004. **16 Suppl 1**: p. 19-29.
37. Dominici, F., et al., *Particulate air pollution and mortality in the United States: did the risks change from 1987 to 2000?* Am J Epidemiol, 2007. **166**(8): p. 880-8.
38. Zhang, H., et al., *Source apportionment of sulfate and nitrate particulate matter in the Eastern United States and effectiveness of emission control programs*. Sci Total Environ, 2014. **490**: p. 171-81.
39. Green, L.C. and S.R. Armstrong, *Particulate matter in ambient air and mortality: toxicologic perspectives*. Regul Toxicol Pharmacol, 2003. **38**(3): p. 326-35.
40. Krall, J.R., et al., *Current Methods and Challenges for Epidemiological Studies of the Associations Between Chemical Constituents of Particulate Matter and Health*. Curr Environ Health Rep, 2015. **2**(4): p. 388-98.
41. Samet, J.M., *CASAC Review of the Policy Assessment for the Review of the PM NAAQS - Second External Review Draft (June 2010)*, in *EPA-CASAC-10-015*, C.A.S.A.C. (CASAC), Editor 2010.
42. Goodman, J.E., D.G. Dodge, and L.A. Bailey, *A framework for assessing causality and adverse effects in humans with a case study of sulfur dioxide*. Regul Toxicol Pharmacol, 2010. **58**(2): p. 308-22.
43. Hill, A.B., *The Environment and Disease: Association or Causation?* Proc R Soc Med, 1965. **58**: p. 295-300.

44. Slikker, W., Jr., et al., *Dose-dependent transitions in mechanisms of toxicity: case studies*. Toxicol Appl Pharmacol, 2004. **201**(3): p. 226-94.
45. Moolgavkar, S.H., et al., *Time-series analyses of air pollution and mortality in the United States: a subsampling approach*. Environ Health Perspect, 2013. **121**(1): p. 73-8.
46. Katsouyanni, K., et al., *Air pollution and health: a European and North American approach (APHENA)*. Res Rep Health Eff Inst, 2009(142): p. 5-90.
47. Cox, L.A., *Miscommunicating risk, uncertainty, and causation: fine particulate air pollution and mortality risk as an example*. Risk Anal, 2012. **32**(5): p. 765-7; author reply 768-70.
48. Yoshimura, I., *The effect of measurement error on the dose-response curve*. Environ Health Perspect, 1990. **87**: p. 173-8.
49. Zeger, S.L., et al., *Exposure measurement error in time-series studies of air pollution: concepts and consequences*. Environ Health Perspect, 2000. **108**(5): p. 419-26.
50. Crump, K.S., *The effect of random error in exposure measurement upon the shape of the exposure response*. Dose Response, 2005. **3**(4): p. 456-64.
51. USEPA, *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards, Second External Review Draft*, H.a.E.I.D. Office of Air Quality Planning and Standards, Editor 2010.
52. USEPA, *Policy Assessment for the Review of the Particulate Matter National Ambient Air Quality Standards (Final)*, H.a.E.I.D. Office of Air Quality Planning and Standards, Editor 2011.
53. Avery, C.L., et al., *Estimating error in using ambient PM2.5 concentrations as proxies for personal exposures: a review*. Epidemiology, 2010. **21**(2): p. 215-23.
54. Avery, C.L., et al., *Estimating error in using residential outdoor PM2.5 concentrations as proxies for personal exposures: a meta-analysis*. Environ Health Perspect, 2010. **118**(5): p. 673-8.
55. Rhomberg, L.R., et al., *Measurement error in environmental epidemiology and the shape of exposure-response curves*. Crit Rev Toxicol, 2011. **41**(8): p. 651-71.
56. Jurek, A.M., et al., *Proper interpretation of non-differential misclassification effects: expectations vs observations*. Int J Epidemiol, 2005. **34**(3): p. 680-7.
57. Fuller, W.A., *Measurement Error Models*1987, Chicester: John Wiley & Sons.
58. Alexeeff, S.E., R.J. Carroll, and B. Coull, *Spatial measurement error and correction by spatial SIMEX in linear regression models when using predicted air pollution exposures*. Biostatistics, 2016. **17**(2): p. 377-89.
59. Smith, A.E. and W. Gans, *Enhancing the Characterization of Epistemic Uncertainties in PM2.5 Risk Analyses*. Risk Anal, 2015. **35**(3): p. 361-78.
60. McClellan, R.O., *Role of science and judgment in setting national ambient air quality standards: how low is low enough?* Air Qual Atmos Health, 2012. **5**(2): p. 243-258.