

Inconsistencies in Risk Analyses for Ambient Air Pollutant Regulations

Anne E. Smith*

This article describes inconsistencies between health risk analyses that the U.S. Environmental Protection Agency (EPA) uses to support its decisions on primary National Ambient Air Quality Standards (NAAQS), and in the associated Regulatory Impact Analyses (RIAs) that accompany each NAAQS rulemaking. Quantitative risk estimates are prepared during the NAAQS-setting deliberations using inputs derived from statistical associations between measured pollutant concentrations and health effects. The resulting risk estimates are not directly used to set a NAAQS, but incorporated into a broader evidence-based rationale for the standard that is intended to demonstrate conformity with the statutory requirement that primary NAAQS protect the public health with a margin of safety. In a separate process, EPA staff rely on the same risk calculations to prepare estimates of the benefits of the rule that are reported in its RIA for the standard. Although NAAQS rules and their RIAs are released simultaneously, the rationales used to set the NAAQS have become inconsistent with their RIAs' estimates of benefits, with very large fractions of RIAs' risk-reduction estimates being attributed to populations living in areas that will already be attaining the respective NAAQS. This article explains the source of this inconsistency and provides a quantitative example based on the 2012 revision of the fine particulate matter (PM_{2.5}) primary NAAQS. This article also demonstrates how this inconsistency is amplified when criteria pollutant co-benefits are calculated in RIAs for non-NAAQS rules, using quantitative examples from the 2011 Mercury and Air Toxics Standards and the currently proposed Clean Power Plan.

KEY WORDS: Benefits; co-benefits; NAAQS; ozone; PM_{2.5}; regulatory impact analysis

1. BACKGROUND

When the primary particulate matter (PM_{2.5}) National Ambient Air Quality Standards (NAAQS) were first established in 1997 (one for annual average and one for daily average ambient PM_{2.5} concentrations), the principal basis for those standards was epidemiological evidence of positive statistical associations between ambient PM_{2.5} levels and adverse health effects, including premature death risk.

These reported associations, combined with a presumption that they represented a causal relationship, were also used to calculate quantitative public health risk estimates to supplement reasoning on setting the NAAQS. Quantitative risk analyses based on epidemiological evidence have continued to be a central feature of the review process for revisions of the PM_{2.5} NAAQS since then, and have also been a salient consideration in revisions of the NAAQS for ozone. This article focuses on a quantitative inconsistency that has emerged between the rationale that U.S. Environmental Protection Agency (EPA) Administrators use for setting a NAAQS when relying primarily on epidemiologically-based health risk evidence, and the estimates of public health benefits

*Address correspondence to Anne E. Smith, Senior Vice President, NERA Economic Consulting, 1255 23rd Street NW, Suite 600, Washington, DC 20037, USA; anne.smith@nera.com.

from those rules that EPA staff produces in its Regulatory Impact Analyses (RIAs).¹

2. THE RATIONALE FOR SETTING A PRIMARY NAAQS

The Clean Air Act requires EPA² to set the primary NAAQS for each criteria pollutant at levels that "are requisite to protect the public health" while "allowing an adequate margin of safety."⁽¹⁾ This determination must be made without regard to the potential cost of meeting the standard,⁽²⁾ and legal rationales for choosing a NAAQS traditionally involved a balanced consideration of three attributes: (1) size of affected population, (2) severity of effect, and (3) certainty of effect.⁽³⁾ However, the evolution since 1997 towards greater reliance on epidemiological evidence in setting a NAAQS forced a shift in how the rationale could be constructed, particularly for PM_{2.5}. This was because the available epidemiological studies on several clearly adverse types of health effects (such as premature death) have not been able to identify a "threshold" or any other less sharp delineation of a point where the risk per unit increment of concentration appears to attenuate.³ This situation eliminates the first two of the three above-mentioned considerations that EPA had typically relied on in

NAAQS-setting rationales. That is, (1) the entire U.S. population is now implicated as at risk at every potential NAAQS level, and (2) the severity of effect can no longer be seen to be changing as lower potential NAAQS levels are considered. As a result, consideration (3)—uncertainty about the reliability of the epidemiologically estimated association—has become the only consideration remaining available to EPA for setting a primary NAAQS above zero that can be argued to be adequately protective of the public health as required by the statute.

This shift in the nature of the scientific evidence for setting a NAAQS was so profound that the U.S. Court of Appeals ruled that the setting of a NAAQS under these circumstances amounted to an unconstitutional delegation of legislative power to the Administrator unless she would first articulate an "intelligible principle" for how to draw that line.⁽⁶⁾ However, the Supreme Court overruled this finding,⁽⁷⁾ with the result being that since then EPA's rationales for at least two of the NAAQS (i.e., PM_{2.5} and ozone) have largely emphasized identifying a level at which continuation of the nonthreshold statistical health associations becomes too uncertain to indicate an actionable level of further public health risk.

The preamble for the 2012 PM_{2.5} NAAQS decision provides an example. It starts by noting that setting a standard based on epidemiological studies that cannot identify a population threshold requires a decision-making approach that "includes consideration of how to weigh the uncertainties in the reported associations across the distributions of PM_{2.5} concentrations in the studies and the uncertainties in quantitative estimates of risk, in the context of the entire body of evidence before the Agency."⁽⁸⁾ Later, the document states, "[i]n reaching decisions on alternative standard levels to propose, the Administrator judged that it was most appropriate to examine where the evidence of associations observed in the epidemiological studies was strongest and, conversely, where she had appreciably less confidence in the associations observed in the epidemiological studies,"⁽⁹⁾ and after a detailed discussion of the epidemiological information states, "[t]he Administrator views this information as helpful in guiding her determination as to *where her confidence in the magnitude and significance of the associations is reduced to such a degree* [emphasis added] that a standard set at a lower level would not be warranted to provide requisite protection that is neither more nor

¹A separate point of discussion regarding the quantitative risk estimates is whether the full body of scientific evidence is sufficient to give confidence that these epidemiological associations reflect a causal relationship between the pollutant and health endpoint studied. This article does not attempt to add to that discussion.

²Formally, under the Clean Air Act, the responsibility for deciding where to set a NAAQS is vested specifically in the Administrator. Throughout this article, when I use the term "EPA," I am referring to the EPA Administrator. When not referring to the Administrator specifically, I use the terms "EPA staff" or "Agency."

³EPA staff and others often refer to this as a "threshold" for effects, but the phenomenon being sought to help identify a protective level for a particular adverse effect need not be a point of sharp delineation where all population-wide effects end. Even evidence of diminishment in the slope of the association would be helpful but has not been consistently found. Lack of detection of such a diminishment in an association, even if the detected association is causal at relatively high concentrations, does not mean one does not exist at some relatively low concentration (see Ref. 4, p. 382). This is because the epidemiological techniques available have very limited ability to reliably discern the shape of a potential concentration-response relationship, and thus to inform the question of where or whether the association may end. It is theoretically established that unavoidable inaccuracies in measurement of an explanatory variable (e.g., pollutant exposure) make it difficult to statistically detect a threshold or other nonlinearity at low concentrations even when it actually exists.⁽⁵⁾

less than needed to provide an adequate margin of safety.”⁽¹⁰⁾

Similarly, in 2008 EPA used lack of confidence in continuation of the epidemiological associations to lower levels as its rationale for not setting the ozone NAAQS lower than 0.075 ppm despite clinical evidence in the record of health responses at yet lower concentrations. The ozone NAAQS preamble states: “A standard set at a level lower than 0.075 would only result in significant further public health protection if, in fact, there is a continuum of health risks in areas with 8-hour average O₃ concentrations that are well below the concentrations observed in the key controlled human exposure studies and if the reported associations observed in epidemiological studies are, in fact, causally related to O₃ at those lower levels. Based on the available evidence, the Administrator is not prepared to make these assumptions. Taking into account the uncertainties that remain in interpreting the evidence from available controlled human exposure and epidemiological studies at very low levels, the Administrator notes that *the likelihood of obtaining benefits to public health with a standard set below 0.075 ppm O₃ decreases* [emphasis added], while the likelihood of requiring reductions in ambient concentrations that go beyond those that are needed to protect public health increases.”⁽¹¹⁾ The U.S. Court of Appeals for the District of Columbia Circuit accepted this rationale and upheld the standard in 2013.⁽¹²⁾

Although the NAAQS rationales are not written to conform to the terminology of probability or expected values, readers with decision analytic or other risk analysis training would be inclined to interpret the above quotes as expressing subjective judgments about the probability that the health relationships apparent in statistical associations cease to exist at some point on the continuum of lower and lower ambient pollutant concentrations. A decision-analytic interpretation of the above statements might be as follows. In order for a selected NAAQS level to be deemed as requisite to protect the public health, EPA’s subjective probability that the relationship exists at and below the selected NAAQS level must, logically, be very nearly zero. (Indeed, the subjective probability of continued effects must fall to nearly zero at an ambient concentration somewhere *above* the selected NAAQS level. This is because the NAAQS needs to include at least some margin of safety, and thus must be set at least somewhat lower than the level where expected risk is deemed to

become too small to be considered a public health concern.)

3. THE RESULTING INCONSISTENCY IN BENEFITS ESTIMATES FOR A NAAQS

Thus, in setting NAAQS using epidemiological evidence, EPA has deemed quantitative estimates of health risks for concentrations below the NAAQS far less reliable and more inaccurate than the numerical precision with which they are reported. In essence, the NAAQS rationales give little or no weight to the subset of the quantitative risk estimates the Agency has placed in the record that have been calculated for pollutant concentrations below the selected NAAQS level. This lack of confidence in risk estimates from that below-the-NAAQS range does not, however, make its way into the RIAs that accompany the release of the final rules.

RIAs are documents that report on the benefits and costs of each major new regulation, such as a revised NAAQS. Federal regulatory agencies are required to prepare RIAs by Executive Order of the President.^(13,14) Although this requirement is unrelated to the legal requirements of the statute that motivates the regulation (such as the Clean Air Act in the case of air pollutant regulations), EPA’s RIAs for air regulations adopt the same epidemiologically-based method of quantifying health risks used when deliberating where to set the NAAQS.⁴ The consistency ends there, however. At the same time that EPA is setting NAAQS at levels where it has minimal confidence that the public health is affected at lower concentrations, the Agency’s RIAs are giving the same weight to risks calculated for population exposures *below* the NAAQS level as they do to risks calculated for population exposures *above* the NAAQS level. That is, RIAs assume elevated hazards exist with 100% certainty for all ambient pollutant exposure levels down to a zero concentration, inconsistent with EPA’s judgments (formed when assessing those pollutants’ hazards), which imply nearly 0% certainty. EPA does not explain or try to justify why data that are too uncertain to use in the NAAQS preamble context are certain enough to use in the RIA context. Although different certainty standards may be

⁴While the “benefits” in an RIA are stated as a monetary value to be compared to the regulation’s costs, they are directly derived from quantitative estimates of physical health effects.

justified in the context of decisions with different consequences, the contexts of a NAAQS preamble and that NAAQS's RIA are not very different at all.

This inconsistency was not always as pronounced as it is now. Until 2009, risk reduction calculations used in air RIAs were at least truncated for pollutant concentrations below the lowest concentration level measured in the epidemiological study being used to make the risk estimates. RIAs would still include risk reduction estimates below the prevailing NAAQS level, as NAAQS levels have always been set at levels above the lowest levels measured in the studies. However, from 2009 onwards, RIAs eliminated even that truncation, which resulted in a sudden and large increase in RIA benefits estimates for $PM_{2.5}$ and ozone pollutant changes.⁽¹⁵⁾ The fact that RIAs calculate health risk reductions below the NAAQS, and now down to zero, is widely known but the following examples quantify the extent to which this practice results in upward-biased risk and benefits estimates. This author recommends that EPA staff more clearly communicate subjective epistemic uncertainty in its RIA benefits estimates. More specifically, the author recommends that the Agency's central estimates of benefits in its RIA be made consistent with the science-policy judgments EPA makes in setting the criteria pollutant standards. This recommendation is in line with the need for more effective sensitivity analysis capabilities for health risk analyses, as described by Smith and Gans.⁽¹⁶⁾

4. OVERSTATEMENT OF EXPECTED BENEFITS OF THE 2012 $PM_{2.5}$ PRIMARY NAAQS REVISION

The implications of this inconsistency are illustrated using as an example the RIA for the 2012 PM NAAQS rulemaking.⁽¹⁷⁾ In this rulemaking, the annual primary standard for $PM_{2.5}$ was tightened from an annual average of 15 to $12 \mu g/m^3$. In the associated RIA, a range of 460 to 1,000 fewer premature deaths per year was estimated from tightening the standard to $12 \mu g/m^3$. This range was derived by applying two different concentration-response functions to the Agency's standard risk calculation formula. The concentration-response coefficient for the lower end of the range was derived using a coefficient from Krewski *et al.*,⁽¹⁸⁾ and the upper end of the range was derived using a coefficient from Lepeule *et al.*⁽¹⁹⁾ A yet wider range of uncertainty in potential mortality risk reductions exists, as explained in Ref. 16, but the following discussion addresses only how

the Agency's own range changes when the assumptions of the RIA's risk analysis are made consistent with EPA's reasoning when choosing how stringently to set the standard.

Calculations were performed using EPA's BenMAP model, which is a PC-based program that enables users to compute health risks associated with criteria pollutants using the standard formulas that EPA uses in its own RIAs, and using EPA's or their own input files and other assumptions.⁽²⁰⁾ The air quality input files that had been used for this RIA's calculations were obtained from EPA staff. After confirming that BenMAP does indeed replicate the mortality reduction estimates reported in the RIA using those data, the same files were then used to assess the portion of the RIA's premature mortality estimates that are associated with the linear, no-threshold assumption that assumes that the risk relationship continues to exist below the selected NAAQS. This analysis found that 70% of the benefits for the standard of $12 \mu g/m^3$ were due to reductions in $PM_{2.5}$ from baseline levels that were already attaining (i.e., lower than) that standard.

Given that the choice of a NAAQS level of $12 \mu g/m^3$ meant that EPA assigned too little confidence in the continuation of health effects below $12 \mu g/m^3$ to warrant setting the NAAQS at a lower level, standard decision analysis would assign negligible probability to calculations of benefits from reductions that would be occurring from levels below that NAAQS. That is, the *expected* values for 70% of the Agency's risk calculations should be approximately zero. When a threshold is assumed at $12 \mu g/m^3$, BenMAP calculates that the expected risk reduction of that NAAQS would be 138 to 313 fewer premature deaths per year, considerably lower than the 460 to 1,000 deaths reported in the RIA. (Dollar values of the benefits also fall proportionally.)

As noted above, the rationale for the NAAQS arguably implies that some of the benefits derived from locations with concentrations just above $12 \mu g/m^3$ also should be given less than 100% weight because of EPA's assurance that exposures to annual average concentrations of $12 \mu g/m^3$ are protective *with an adequate margin of safety*. EPA rarely if ever defines the magnitude of its margin of safety quantitatively. However, ranges for its magnitude could be tested with sensitivity analyses. If, for example, the margin of safety is taken to be about $1 \mu g/m^3$, and a threshold is assumed in the risk relationship $13 \mu g/m^3$, BenMAP calculates the expected benefits associated with the selected NAAQS of $12 \mu g/m^3$ are

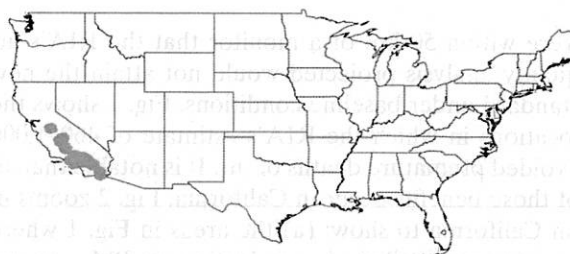
Table I. Estimates of Avoided Premature Deaths in 2020 for the 12 $\mu\text{g}/\text{m}^3$ PM_{2.5} NAAQS: RIA Assumptions Compared to Alternative Views Suggested by EPA's Rationale for that NAAQS

Confidence Category (baseline PM _{2.5} concentration)	NAAQS-Based Risk Reduction Estimate	RIA-Based Risk Reduction Estimate (% of total)
Already attaining ($\leq 12 \mu\text{g}/\text{m}^3$)	Approximately 0	318 (70%)
Not attaining/in margin (e.g., >12 to $13 \mu\text{g}/\text{m}^3$)	0–117	117 (26%)
Not attaining/above margin (e.g., $>13 \mu\text{g}/\text{m}^3$)	21	21 (5%)
Confidence weighted		
Total risk reduction estimate	21–117	456

only 21 to 48 deaths, less than 5% of the RIA's estimate of benefits from that standard.

Whether the particular assumptions in this analysis about where the concentration-response relationship begins to exist are reasonable or should be refined, its point is that the RIA's benefits estimates are very sensitive in the downward direction to expressions of declining confidence in continuation of the association at or just above the selected NAAQS level. The result is that the RIA benefits are substantially overstated compared to those that would more appropriately reflect the subjective weights expressed by EPA in its rationale for setting the standard at 12 $\mu\text{g}/\text{m}^3$. Table I contrasts the results of the RIA with judgments about confidence in those risk calculations that one might infer from the NAAQS rationale, and illustrates one way that RIAs could be enhanced to better communicate to the public the implications of the judgments made in setting the NAAQS for the rule's benefits estimates.

For simplicity, Table I summarizes only the lower-bound benefits estimate of 460 deaths (which BenMAP calculates more precisely as 456 deaths).⁵ In this table, the risk estimates are divided into three "confidence categories." The lowest confidence category is for risk reductions attributed to populations already residing in areas of attainment (i.e., with annual average concentrations less than 12 $\mu\text{g}/\text{m}^3$). Given the NAAQS rationale, the public health risk is *de minimis*, and in weighted terms, would be nearly zero, while in the RIA, which gives 100% weight to all such risk calculations, benefits equal to about 318 deaths per year are assigned. The middle confidence category is for risk reductions attributed to populations in areas that are just above the NAAQS before the standard is implemented, but close

**Fig. 1.** Areas projected in the PM_{2.5} NAAQS RIA to experience health benefits under the selected NAAQS of 12 $\mu\text{g}/\text{m}^3$ (456–1,033 avoided premature deaths, rounded to nearest death).

enough to attainment that they might be viewed as being within the (undefined) "margin of safety." (For purposes of constructing the illustrative tabular summary, the margin of safety is assumed to be about 1 $\mu\text{g}/\text{m}^3$, meaning that less than the NAAQS-based weights would be declining or perhaps nearly zero even within this category of baseline exposures.) To reflect risk estimates that fall in this category, the NAAQS-based risk reduction estimate is listed as being somewhere between 0 and 117, while the RIA would assign it 117 with 100% confidence.

Finally, there are 21 avoided premature deaths estimated for populations living in areas well above the NAAQS. For this third category, the RIA's benefits estimates can be considered consistent with the NAAQS-based rationale. Note that for the PM_{2.5} NAAQS RIA, this category accounts for only about 5% of the total RIA benefits estimate. It is recommended that RIAs provide their benefits estimates for criteria pollutants in a format such as Table I, and more explicitly provide weighted benefits estimates for confidence categories that are defined *with respect to the NAAQS level*.

Geographical representation of where these health benefits are expected to occur is also interesting to explore. The PM_{2.5} NAAQS RIA calculated reductions in premature mortality only for areas that

⁵The upper-bound risk estimates would fall into the three rows in the table in the same proportions as seen for the lower-bound estimates in the table.

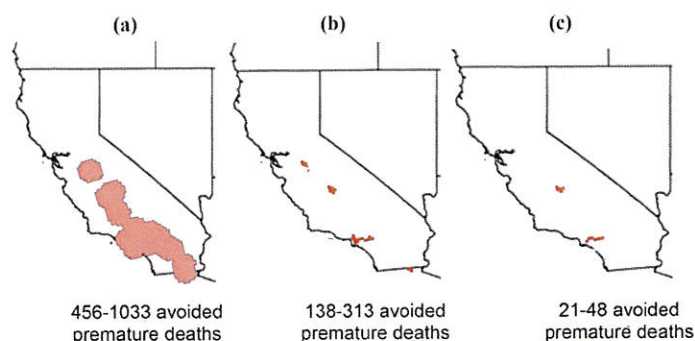


Fig. 2. Sensitivity analysis of areas projected to experience health benefits under the 12 $\mu\text{g}/\text{m}^3$ NAAQS: (a) assuming benefits for all baseline $\text{PM}_{2.5}$ levels; (b) assuming risks exist only if baseline $\text{PM}_{2.5}$ is above 12 $\mu\text{g}/\text{m}^3$; (c) assuming risks exist only if baseline $\text{PM}_{2.5}$ exceeds the selected standard by more than 1 $\mu\text{g}/\text{m}^3$.

were within 50 km of a monitor that the RIA's air quality analysis projected would not attain the new standard under baseline conditions. Fig. 1 shows the locations in which the RIA's estimate of 460–1,000 avoided premature deaths occur. It is notable that all of those benefits occur in California. Fig. 2 zooms in on California to show: (a) the areas in Fig. 1 where benefits are attributed to reductions in $\text{PM}_{2.5}$ at any level (i.e., showing the same areas as in Fig. 1); (b) the more limited areas projected to experience a health benefit when only reductions in $\text{PM}_{2.5}$ that start above the 12 $\mu\text{g}/\text{m}^3$ NAAQS are considered; and (c) the even more limited areas if a 1 $\mu\text{g}/\text{m}^3$ margin of safety is assumed to be associated with the selected standard of 12 $\mu\text{g}/\text{m}^3$. That is, Fig. 2(c) only gives weight to risks below 13 $\mu\text{g}/\text{m}^3$. Both Figs. 2(b) and (c) reveal a far smaller area of at-risk populations than assumed in the RIA (i.e., than in Fig. 2(a)).

This example from the $\text{PM}_{2.5}$ NAAQS RIA brings to light another important uncertainty in its mortality benefits. All of the benefits estimates for the NAAQS of 12 $\mu\text{g}/\text{m}^3$ are based on $\text{PM}_{2.5}$ changes in California. The risk calculations for changes in $\text{PM}_{2.5}$ in California are performed using relative risk estimates derived from the entire United States, yet the epidemiological evidence that an association between $\text{PM}_{2.5}$ and all-cause mortality risk exists in California is tenuous.⁶ Hence all of the above risk estimates might actually be zero, even if one does

not wish to discount risks in areas already below the NAAQS. In other words, the much tighter 2012 $\text{PM}_{2.5}$ NAAQS was set on the basis of projected mortality reductions that occur only in a part of the United States where the evidence of heightened mortality risk from $\text{PM}_{2.5}$ appears to be weaker than in other parts of the United States.

5. OVERSTATEMENT OF CRITERIA POLLUTANT CO-BENEFITS IN NON-NAAQS RULEMAKINGS

As explained in Ref. 15, epidemiologically-based estimates of co-benefits from coincidental reductions of ambient criteria pollutants (especially $\text{PM}_{2.5}$) have also driven statements about regulatory benefits for a majority of non-NAAQS air rulemakings in recent years. The upward bias in RIA benefits estimates becomes even more pronounced when co-benefits are calculated from coincidental criteria pollutant reductions under regulations that do not relate to the NAAQS or regulations to help attain a NAAQS. Prominent examples are the RIAs for the Mercury and Air Toxics Standards (MATS) for electricity-generating units promulgated in December 2011⁽²¹⁾ and the Clean Power Plan (CPP) proposed in June 2014.⁽²²⁾

The MATS RIA projected $\text{PM}_{2.5}$ co-benefits in the hundreds of billions of dollars per year, based almost entirely on estimates of reduced premature mortality from reductions in $\text{PM}_{2.5}$: 4,200 to 11,000 deaths per year. The reductions in $\text{PM}_{2.5}$ in the MATS RIA are projected to occur when generating units are forced to install controls to reduce acid gas emissions, which will also reduce SO_2 emissions, a precursor to ambient $\text{PM}_{2.5}$ formation. A figure in the MATS RIA reveals that over 99% of those projected benefits are projected to occur in areas where the $\text{PM}_{2.5}$ levels will already be below the $\text{PM}_{2.5}$ NAAQS

⁶The $\text{PM}_{2.5}$ RIA⁽¹⁷⁾ cites seven California-specific $\text{PM}_{2.5}$ cohort studies with all-cause risk estimates and notes that four have insignificant associations while three have larger coefficients (Ref. 17 at p. 5, A-13). However, one of the three positive findings cited (i.e., Ostro *et al.*, 2010) was erroneous, according to an erratum published the following year (Ostro *et al.*, 2011), and the corrected estimate of association was found to be insignificant. The remaining two positive findings cited were from the same cohort, one estimate being just an update of the other. Thus, the evidence for an all-cause mortality association in California alone consists of five null findings and one cohort with a positive finding.

of $12 \mu\text{g}/\text{m}^3$ (Figure 5–15 on p. 5–102 of Ref. 21). If the MATS rule's co-benefits are calculated probabilistically, accounting for the very low subjective probability that EPA assigned to the existence of the $\text{PM}_{2.5}$ -health effects relationships at levels below the NAAQS, the resulting estimate of expected benefits from the MATS rule becomes nearly zero.

The fraction of the $\text{PM}_{2.5}$ co-benefits calculated below the NAAQS is much higher in the MATS RIA than the already high level of 70% that we have found for the benefits calculated for the $\text{PM}_{2.5}$ NAAQS rule itself. This is due to the fact that benefits in the RIA for the NAAQS rule were calculated only in areas within 50 km of a monitor that was projected to be out of attainment. By letting projected nonattainment constrain the geographical area over which benefits will be calculated, one ensures that a larger fraction of the resulting benefits will indeed be from areas above the NAAQS. However, when co-benefits of some other rule are assessed using $\text{PM}_{2.5}$ risk relationships, no such constraint is applied. In the MATS rule, co-benefits were calculated across the entire nation, and furthermore, the units where acid gas controls were incremental to baseline controls were more likely to be in areas already attaining the NAAQS. As a result, nearly all of the $\text{PM}_{2.5}$ co-benefits are projected in NAAQS-attaining areas. For these reasons, the bias in $\text{PM}_{2.5}$ co-benefits estimates in RIAs for non- $\text{PM}_{2.5}$ rulemakings will tend to be much greater than the bias in the direct benefits estimates in RIAs for $\text{PM}_{2.5}$ regulations.

The same magnitude of overstatement of co-benefits is apparent in the RIA for the proposed CPP RIA, which includes co-benefits for both $\text{PM}_{2.5}$ and ozone. In the CPP RIA (focusing, for simplicity, on its Option 1 with state-level implementation) the $\text{PM}_{2.5}$ co-benefits of the rule are estimated to be up to 4,100 deaths in 2020 and up to 6,200 deaths in 2030, and the ozone co-benefits are estimated to be up to 170 and 440 in those respective years (Tables 4–16 through 4–18 on pp. 4–34 to 4–36 of Ref. 22). Unlike the MATS RIA, the CPP RIA does not provide any information on the fraction of these co-benefits that are calculated for areas already attaining those two NAAQS, but they can be inferred by replicating the co-benefits calculations from other data in the RIA.⁷ Recalling that the $\text{PM}_{2.5}$ NAAQS RIA indicates that only California will be exceeding

the $\text{PM}_{2.5}$ NAAQS in 2020, only California-based $\text{PM}_{2.5}$ co-benefits estimates could be associated with exposures in the above-the-NAAQS category: less than 1% of the CPP RIA's $\text{PM}_{2.5}$ co-benefits are attributable to changes in emissions in California in 2020. Furthermore, the $\text{PM}_{2.5}$ NAAQS is supposed to be fully attained by 2020, so even that sliver of the $\text{PM}_{2.5}$ co-benefits attributable to California are supposedly in an attainment area. Although California is not projected to attain the ozone NAAQS before 2030, less than 0.5% of the ozone-related co-benefits are associated with changes in ozone precursors in California. Thus, in the CPP RIA as well in the MATS RIA, more than 99% of the co-benefits would be discounted if health risks below the NAAQS are assigned a much lower probability (or confidence weight) than risks above the NAAQS.

6. CONCLUSION

In conclusion, we find that a large majority of the Agency's estimated health benefit from the 2012 $\text{PM}_{2.5}$ NAAQS are attributable to reductions of $\text{PM}_{2.5}$ in areas that are already in attainment of the $\text{PM}_{2.5}$ NAAQS. RIA calculations of risk reduction in areas already attaining the new NAAQS are given the same weight (i.e., subjective confidence level) as projected benefits from areas that would be exceeding the NAAQS. These RIA calculations are based on assumptions that are inconsistent with the rationale for that NAAQS. The above sensitivity analyses show that this causes RIAs' benefits estimates to be much larger than estimates of the expected benefits that can be reasonably inferred from EPA's NAAQS-setting rationale. The overstatement becomes nearly 100% for co-benefits from criteria pollutants in RIAs for non-NAAQS regulations, such as the MATS rule and the proposed CPP rule. RIAs should be written to reflect consistency with EPA's NAAQS policy judgments. Precise confidence weights will likely never be articulated, but this article has shown that the quantitative importance of such policy judgments for benefits estimates can be communicated to RIA readers in simple formats. It is the opinion of this author that such quantitative disclosure is important to maintaining credibility and trust in the Agency's RIAs.

ACKNOWLEDGMENTS

This work was conducted with funding from the Electric Power Research Institute. The author thanks Ms. Reshma Patel for her analytical support

⁷This involves using data on emissions reductions of the $\text{PM}_{2.5}$ and ozone precursor emissions in the RIA's Table 4–10, and multiplying them by the incidence-per-ton estimates in Tables 4A–5 through 4A–7.

in preparing the analyses used in this article. The author also thanks three anonymous reviewers for their comments and suggestions. Any errors remain the author's sole responsibility.

REFERENCES

1. Section 109 (b)(1), Clean Air Act; 42 USC §7409.
2. *Whitman v. American Trucking Associations*, 531 U.S. 473–476 (2001).
3. *Lead Industries Association Inc v. Environmental Protection Agency*, 647 F.2d 1130 (1980).
4. Smith AE. Response to commentary by Fann *et al.* on “Enhancing the characterization of epistemic uncertainties in PM_{2.5} risk analyses.” *Risk Analysis*, 2015; 35(3):381–384.
5. Brauer M, Brumm J, Vedal S, Petkau AJ. Exposure misclassification and threshold concentrations in time series analyses of air pollution health effects. *Risk Analysis*, 2002; 22(5):1183–1193.
6. *American Trucking Associations v. EPA*, 175 F. 3d 1027, 1034–1037 (D.C. Cir. 1999), rehearing granted in part and denied in part, 195 F. 3d 4 (D.C. Cir. 1999), affirmed in part and reversed in part, *Whitman v. American Trucking Associations*, 531 U.S. 457 (2001).
7. *Whitman v. American Trucking Associations*, 531 U.S. 473–476 (2001).
8. 78 Fed. Reg. 3086 at 3098.
9. 78 Fed. Reg. 3086 at 3139.
10. 78 Fed. Reg. 3086 at 3161.
11. 76 Fed. Reg. 16436 at 16483.
12. *Mississippi v. EPA*, 744 F.3d 1334 (D.C. Cir. 2013).
13. Executive Order 12866. 1993. “Regulatory Planning and Review.” 58 Fed. Reg. 51735, October 4. Available at: www.whitehouse.gov/omb/inforeg/eo12866.pdf.
14. Executive Order 13563. 2011. “Improving Regulation and Regulatory Review.” 76 Fed. Reg. 3821, January 18. Available at: http://www.regulations.gov/exchange/sites/default/files/doc_files/President%27s%20Executive%20Order%2013563_0.pdf.
15. Smith, AE. An evaluation of the PM_{2.5} health benefits estimates in regulatory impact analyses for recent air regulations. Report prepared for the Utility Air Regulatory Group, December 2011, at p. 24. Available at: http://www.nera.com/content/dam/nera/publications/archive2/PUB_Smith_Ozone_NAAQS_0711.pdf.
16. Smith AE, Gans W. Enhancing the characterization of epistemic uncertainties in PM_{2.5} risk analyses. *Risk Analysis*, 2015; 35(3):361–78.
17. EPA. Regulatory impact analysis for the final revisions to the National Ambient Air Quality Standards for particulate matter, EPA-452/R-12-003, December, 2012.
18. Krewski D, Jerrett M, Burnett RT, Ma R, Hughes E, Shi Y, Turner MC, Pope CA 3rd, Thurston G, Calle EE, Thun MJ, Beckerman B, DeLuca P, Finkelstein N, Ito K, Moore DK, Newbold KB, Ramsay T, Ross Z, Shin H, Tempalski B. Extended follow-up and spatial analysis of the American Cancer Society Study linking particulate air pollution and mortality. *Research Report Health Effects Institute*, 2009; 140:5–114.
19. Lepeule J, Laden F, Dockery D, Schwartz J. Chronic exposure to fine particles and mortality: An extended follow-up of the Harvard Six Cities Study from 1974 to 2009. *Environmental Health Perspectives*, 2012; 32(1):81–95.
20. BenMAP version 4.0.67 Available at <http://www.epa.gov/air/benmap/download.html>.
21. EPA. Regulatory impact analysis for the final mercury and air toxics standards, EPA-452/R-11-011, December 2011.
22. EPA. Regulatory impact analysis for the proposed carbon pollution guidelines for existing power plants and emission standards for modified and reconstructed power plants, EPA-542/R-14-002, June 2014.