John Engler

President and CEO

October 9, 2007

Ms. Molly A. O'Neill
Assistant Administrator, Office of Environmental Information
And Chief Information Officer
United States Environmental Protection Agency
Office of Environmental Information
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Dear Ms. O'Neill:

On behalf of the National Association of Manufacturers, I am hereby submitting the attached Request for Correction (RFC) in accordance with the procedures set forth in Section 8.5 of EPA's Information Quality Guidelines. Our RFC concerns the Notice of Proposed Rulemaking on the ozone NAAQS revision, and certain scientific documents disseminated by EPA in support of that NPRM. For the Agency's convenience, we have also submitted this RFC as a public comment directly to the docket for this rulemaking.

We understand from EPA's Guidelines that this RFC will be treated as a public comment to the docket for that rulemaking, and that the Agency will issue a final response on or before the date on which it promulgates the final rule.

Sincerely,

John Engler

JE/qr

Table of Contents

I. Summary

This Request for Correction (RFC) is submitted in accordance with administrative procedures established by the U.S. Environmental Protection Agency (EPA) to ensure and maximize the quality of information the Agency disseminates:

> The Environmental Protection Agency (EPA) is committed to providing public access to environmental information. This commitment is integral to our mission to protect human health and the environment. One of our goals is that all parts of society - including communities, individuals, businesses, State and local governments, Tribal governments - have access to accurate information sufficient to effectively participate in managing human health and environmental risks. To fulfill this and other important goals, EPA must rely upon information of appropriate quality for each decision we make (U.S. Environmental Protection Agency 2002).

EPA is committed to the principles of information quality. It is Agency policy that:

- Disseminated information should adhere to a basic standard of quality, including objectivity, utility, and integrity.
- The principles of information quality should be integrated into each step of EPA's development of information, including creation, collection, maintenance, and dissemination.
- Administrative mechanisms for correction should be flexible, appropriate to the nature and timeliness of the disseminated information, and incorporated into EPA's information resources management and administrative practices.

This RFC concerns EPA's Notice of Proposed Rulemaking (NPRM) proposing to revise the National Ambient Air Quality Standard (NAAQS) for ozone, and several supporting documents, each of which contains influential scientific information crucial for regulatory decision making under §§ 108 and 109 of the Clean Air Act. Nothing in this RFC contests the statutory authority of the Administrator to make this decision; to make it promptly; or the nature of the criteria he is required to take into account in making his decision. Indeed, it is the petitioners view that adherence to the information quality standards EPA has committed to uphold is the best and surest way to fulfill this statutory mandate.

Section II identifies the specific information EPA has disseminated which is the subject of this RFC; the identity of the affected party submitting the RFC; the information quality principles implicated; and the procedures EPA has

established to administratively manage correction requests. Section III identifies a series of areas in which petitioners believe that the primary scientific information identified in Section II.A exhibits manifest information quality errors that will, if not corrected, have a material effect on the ability of the Administrator to exercise his statutory responsibilities. Sections IV extends that review to analogous problems with EPA's human health risk assessment. Section V identifies concerns with the manner in which EPA has managed information and policy recommendations it received from the Clean Air Scientific Advisory Committee (CASAC). Sections VI and VII focus on information quality issues related to the "rollback" and Policy Relevant Background assumptions EPA has used to predict the risk reduction that would be achieved by lowering the standard.

II. Introduction

This document is a formal Request for Correction (RFC), submitted in accordance with government-wide requirements related to information quality (Office of Management and Budget 2002; Section 515, Treasury and General Government Appropriations Act for Fiscal Year 2001 [2000]) and procedures established by EPA (U.S. Environmental Protection Agency 2002), concerning certain information disseminated by the Agency in association with its recent proposed rulemaking on the ozone National Ambient Air Quality Standard (NAAQS) (Docket ID EPA-HQ-OAR-2005-0172).

A. Information Subject to this Petition

EPA's proposed revision to the ozone NAAQS is published as a Notice of Proposed Rulemaking:

(U.S. Environmental Protection Agency 2007b)	"National Ambient Air Quality Standards for Ozone: Proposed Rule," 72 Fed. Reg. 37818-37919 (July 11).

This document includes influential scientific, technical, statistical and economic information that is subject to the Information Quality Act (Section 515, Treasury and General Government Appropriations Act for Fiscal Year 2001 [2000]), OMB's government-wide guidelines (Office of Management and Budget 2002), and EPA's agency-specific guidance (U.S. Environmental Protection Agency 2002).

EPA also has disseminated several secondary documents containing covered influential information:

(U.S. Environmental Protection Agency 2006a)	"Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume I of III," EPA 600/R-05/004aF, February 2006.
(U.S. Environmental Protection Agency 2006b)	"Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume II of III," EPA 600/R-05/004bF
(U.S. Environmental Protection Agency 2006c)	"Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume III of III," (EPA 600/R-05/004cF)
(U.S. Environmental Protection Agency 2007c)	"Ozone Health Risk Assessment for Selected Urban Areas," EPA-452/R-07-009, July 2007.
(U.S. Environmental Protection Agency 2007f)	"Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information: OAQPS Staff Paper," EPA-452/R-07-007, July 2007.
(Brown 2007)	"The effects of ozone on lung function at 0.06 ppm in healthy adults." EPA-HQ-OAR-2005-0172-0175, June 2007.
(Langstaff 2007)	"Analysis of Uncertainty in Ozone Population Exposure Modeling." January 31. (http://www.epa.gov/ttn/naaqs/standards/ozone/data/2007-01_o3_exposure_uncertainty.pdf, EPA-HQ-OAR-2005-0172-0174.

We have tried to cover the major information quality issues in the risk assessment components of these documents, but the task of being comprehensive is impossible give then the length of the public comment period on the NPRM, and it would be formidable even with much more time. The quantity of text that EPA has produced is astounding.

Some important documents have only recently been posted to the online docket. We have not been able to review them, but reserve the right to supplement or amend this RFC as appropriate.

On August 2, 2007, EPA released its Regulatory Impact Analysis (RIA) for this proposed rule:

(U.S. Environmental Protection Agency 2007e)	"Regulatory Impact Analyses: 2007 Proposed Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone," released August 2, 2007 (www.epa.gov/ttn/ecas/ria.html#ria2007)

Given the short amount of time available since its publication, we have not yet reviewed the RIA to identify material information quality errors that might lie within it. The RIA does not appear to have been distributed for notice and comment – it is not part of the standard-setting process and we cannot locate a relevant Federal Register notice requesting public comment -- so it is not covered by this part of EPA's information quality guidelines (U.S. Environmental Protection Agency 2002, Section 8.5).

Nonetheless, the RIA incorporates scientific information from the other documents listed above. Thus, our challenges to the scientific information in the aforementioned documents also apply to the RIA to the extent that the RIA contains materially equivalent information quality errors.

This RFC concerns influential scientific, technical, statistical, and economic information contained or referenced in these documents. It does not include material that is strictly policy in nature; such information is excluded from the definition of "information" because it is an expression of values or preferences, and not of facts or data (Office of Management and Budget 2002).[1]

B. Affected Party

The National Association of Manufacturers (NAM) is the nation's largest industrial trade association representing small and large manufacturers in every industrial sector and in all 50 states. Headquartered in Washington, D.C., the NAM has more than 11,000 corporate members, representing a sector that employs more than 14 million American workers. The NAM's mission is to enhance the competitiveness of manufacturers and improve American living

[1] "'Information' means any communication or representation of knowledge such as facts or data, in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms. This definition includes information that an agency disseminates from a web page, but does not include the provision of hyperlinks to information that others disseminate. This definition does not include opinions, where the agency's presentation makes it clear that what is being offered is someone's opinion rather than fact or the agency's views." See Section V(5) at 8460.

standards by shaping a legislative and regulatory environment conducive to U.S. economic growth.

As the leading voice of manufacturing in the United States, the NAM is deeply concerned that crucial decisions on air pollution control policy reflect the best, unbiased scientific information possible. Our members, and their employees and families, deserve that these important policy decisions be grounded in science.

C. Applicable Error Correction Procedures

Under OMB's government-wide information quality guidelines (Office of Management and Budget 2002), every agency must issue its own implementing guidelines, taking account of its specific needs and characteristics. The OMB guidelines provide a procedural and substantive floor; agencies are allowed to establish more rigorous procedures and substantive information quality standards, but they may not establish procedures or standards that are less demanding.

1. *EPA's preferred procedures for RFC challenging information contained in a proposed rule*

EPA's Information Quality Guidelines (U.S. Environmental Protection Agency 2002) follow the OMB Guidelines in most respects that are material to this petition. For example, EPA's definitions of relevant terms are the same. EPA's agency-specific procedures for affected parties to follow in submitting requests for correction (RFCs) are set forth in Section 8. Where an affected party seeks a correction to information disseminated as part of a Notice of Proposed Rulemaking, EPA says such RFCs should be submitted as part of the public comment process (Section 8.5, page 32):

> When EPA provides opportunities for public participation by seeking comments on information, the public comment process should address concerns about EPA's information. For example, when EPA issues a notice of proposed rulemaking supported by studies and other information described in the proposal or included in the rulemaking docket, it disseminates this information within the meaning of the Guidelines. The public may then raise issues in comments regarding the information.

EPA has committed to implement OMB's directive to respond in a timely manner by responding within 90 days (Section 8.4, page 31). For RFCs submitted as part of a public comment process, EPA has committed to respond prior to taking final agency action.

2. Limits on EPA's procedures

Affected parties are not required to submit RFCs before the conclusion of the public comment period because information dissemination is ongoing and does not stop once a public comment period closes. Nor is an affected party precluded in any way from submitting an RFC concerning a promulgated final rule, for which the public comment period is obviously closed.[2] By submitting an RFC during the public comment period, an affected party can be assured, according to EPA's Guidelines, of receiving a timely final (not interim) response.

EPA cannot reject an RFC on the ground that the information contested was, at some time in the past, subject to a public comment period that is now closed. As OMB says in its government-wide information quality guidelines:

> The agency's administrative mechanisms … shall apply to information that the agency disseminates on or after October 1, 2002, *regardless of when the agency first disseminated the information* (emphasis added) (Office of Management and Budget 2002, 8459, Section III.4) (emphasis added).

D. Relevant Information Quality Principles

Each of the documents that is designated a subject of this RFC is *influential*, as that term is defined by both OMB and EPA. The specific information quality principles at issue are (a) utility, (b) integrity, and (c) objectivity. Objectivity comes in two subspecies: (i) substantive objectivity and (ii) presentational objectivity. Related to but distinct from the twin objectivity principles is a requirement that influential information be transparent and capable of being substantially reproduced. Transparency is essential for reproducibility, and reproducibility often is necessary for affected parties to be able to detect information quality errors.

1. Failure to adhere to the objectivity standards

In this RFC, petitioners claim that information within the listed documents does not satisfy the information quality principles of objectivity (both subspecies). In particular, EPA's ozone risk assessment is neither substantively

---

[2] "The agency's administrative mechanisms, under paragraph III.3., shall apply to information that the agency disseminates on or after October 1, 2002, *regardless of when the agency first disseminated the information*" (emphasis added) (Office of Management and Budget 2002)

objective nor presented in an objective manner. These defects are pervasive and systemic. In some cases they are obvious, and in other cases quite subtle. Because EPA's Regulatory Impact Analysis (RIA) relies on this information as a critical input for the estimation of health benefits, these estimates also are not substantively objective.

2. Failure to adhere to the utility standard

Petitioners also claim that because of these systemic and material defects in objectivity, the documents subject to this RFC do not satisfy the utility standard. Utility requires that information that is disseminated be useful for the purpose to which it is intended. In the case of the RIA, the purpose of the document is to accurately, fully, and clearly inform the public concerning the costs, benefits, distributional consequences, and other effects attributable to a more stringent ozone NAAQS. Pervasive and systemic information quality errors in EPA's risk assessment render the Agency's benefit estimates systematically biased, and thus neither valid nor reliable for informing the public. Substantively "accurate, reliable, and unbiased" benefit estimates require, at a minimum, "accurate, reliable, and unbiased" estimates of risk. It is impossible for a benefit estimate to satisfy the substantive objectivity standard if it must rely on crucial information that is materially defective with respect to substantive objectivity. For that reason alone, benefit estimates in the RIA also fail to satisfy the substantive objectivity standard, and by failing that standard they cannot have utility for their intended purpose.

The purpose of the Criteria Document and Staff Paper were to accurately, fully, and clearly inform the Administrator concerning the health risks posed by ozone at levels below the current standard, the incidence of health effects resulting from these risks assuming attainment of the current standard, and the change in incidence resulting from alternative, lower standards. Due to EPA's pervasive and systemic failure to adhere to the substantive and presentational objectivity standards in its risk assessment, it is impossible for the Staff Paper and Criteria Document to have utility for the Administrator so long as he is committed to set the standard in accordance with the criteria established by law. The law does not authorize the Administrator to base his decision on inaccurate scientific information.

The purpose of the preamble to the Notice of Proposed Rulemaking is to articulate, and communicate to the public, the scientific information that the Administrator considered, and the reasoned basis for determining what standard to propose to set. The Administrator has substantial policy discretion provided by law to decide where to set the standard, and the reasoned basis set forth in the preamble explains how the Administrator incorporated the scientific information

he was provided. But this scientific information is fundamentally flawed because it systematically violates the objectivity standards. For that reason, the Administrator's reasoned basis for decision-making almost certainly relies on inaccurate scientific information. In this RFC, petitioners do not challenge the Administrator's reasoned basis for decision-making, for such a challenge is impermissible under both OMB's and EPA's information quality guidance. Rather, petitioners challenge the scientific and statistical information provided to the Administrator. It is highly likely that the Administrator's reasoned basis for decision-making would have been different if he had been provided scientific and statistical information that adhered to applicable information quality principles.

This RFC does not take any position concerning what reasoned determination the Administrator might make if given scientific and statistical information that adheres to applicable information quality principles. Petitioners' interest is to ensure that the scientific and statistical information, *upon which the Administrator must rely to exercise his statutorily delegated policy discretion*, constitutes the best available scientific evidence. To meet that standard, the evidence and its presentation must adhere to the objectivity, integrity, and utility standards.

E. Relief Sought

Petitioners have identified these defects in information quality to assist EPA in developing unbiased estimates of human health risk posed by ozone inhalation. This requires that EPA take the following steps:

- Disseminate for independent examination the pre-dissemination review which the Agency is required to have performed pursuant to its own information quality guidelines (U.S. Environmental Protection Agency 2002).
- Describe and characterize scientific and statistical information in an objective manner without embedded risk management policy preferences, whether disclosed or undisclosed, taking care to ensure that this information is accurate, reliable, and unbiased, and presented in an accurate, clear, complete, and unbiased manner.
  - All information describing, characterizing, or quantifying human health risk must be presented without regard for the air pollution control policy preferences of EPA staff or management, the Bush administration or Members of Congress, nongovernmental organizations, or industry.
  - Where inferences are drawn from scientific and statistical information and characterized as scientific, they must be supported solely by *scientific* logic and reasoning, free of the air pollution control policy preferences all

scientists have concerning what decision the Administrator, in their view, ought to make in exercising his statutory discretion.

- The proper role of scientists is to describe and predict human health risks, not to leverage their position as scientists to advocate air pollution control policy based on the values and preferences they happen to hold.

❖ Ensure that the estimates of benefits, costs, distributional consequences and other similar phenomena contained in the RIA are derived from descriptions, characterizations, and estimates of human health risk (and health effects valuations) that adhere to the substantive and presentational objectivity principles.

- All information describing, characterizing, or quantifying the value of human health risks avoided must be presented without regard for the air pollution control policy preferences of EPA staff or management, the Bush administration or Members of Congress, nongovernmental organizations, or industry.
- Where inferences are drawn from scientific and statistical information and characterized as economic, these inferences must be supported solely by economic logic and reasoning, free of the air pollution control policy preferences all economists have concerning what decision the Administrator ought to make in exercising his statutory discretion.
- The proper role of economists is to describe and predict the costs, benefits, and other effects from implementing a lowered ozone NAAQS, not to leverage their position as economists to advocate air pollution control policy based on the values and preferences they happen to hold.
- An RIA performed in accordance with information quality principles has critical utility for the public, the audience for whom it is prepared. An RIA that does not adhere to these principles has negative utility insofar as it leads the public to believe incorrect statements about benefits, costs, and other effects.

III. Information Quality Errors in the Description, Analysis and Presentation of Scientific Information

Bias takes many forms. Certain biases affect the scientific information upon which EPA relies. Other biases arise from how EPA chooses to utilize this information. Bias per se is not a violation of the information quality standard of objectivity because it is an evitable fact when dealing with uncertain quantities that have to be estimated. However, *purposeful* bias – the dissemination of information that is known or intended to over- or understate uncertain quantities

– is unambiguously a violation of the objectivity standard. Information containing a series of purposeful biases systematically violates the objectivity standard.

A. EPA Begins with an Inevitably Biased Scientific Database

Although independently and externally peer reviewed scientific information enjoys a presumption of adequate objectivity (Office of Management and Budget 2002), this presumption is a weak one that can be rebutted by persuasive evidence that the information is not, in fact, objective. To succeed in rebuttal, a petitioner need only show that objectivity is not met, not that the petitioner "knows" the right answer. Nor must a petitioner prove that alternative scientific information is objective to satisfy his burden of proof. Federal agencies, not petitioners for correction or research scientists, are subject to the strictures of information quality. This system is designed to instill incentives that reward constant improvements in scientific quality. The goal is for more-objective information to supplant less-objective information at every opportunity. An agency cannot justify the dissemination of less-objective information because such information more conveniently conforms to its policy mission, or to staff risk management preferences.

Scientific information provided to EPA may be biased for several reasons. We discuss three such reasons below.

1. Effective control by a party with a risk management objective

Health-effects studies have been funded by government, industry, and sometimes jointly. Because industry has well-defined policy interests, it is widely presumed that it would endeavor to control the research it funds to ensure that it yields desired results. For that reason, it is commonplace for industry to fund research through arm's-length contracts that insulate researchers from such interference.

Government agencies also have well-defined policy interests, however, and thus they have similar incentives to control how research is performed. It is not clear that they devote the same effort to insulate their grantees and contractors from sponsor interference.

Federal information quality guidelines deal with this bias problem two ways. First, they place a high value on full disclosure sufficient to ensure reproducibility. Irrespective of the motives of research sponsors, reproducibility is believed to be the best way to discern whether interference occurred. When a research sponsor declines to make information available, that which it does disclose becomes presumptively suspect. Second, as long as information is

capable of being substantially reproduced, information quality principles emphasize quality attributes and not the source of research sponsorship per se.

If these principles are adhered to, then biases resulting from sponsor control over research should be rare or nonexistent. Unfortunately, EPA's risk assessment documents are nontransparent with respect to scientific information that the Agency itself generated or sponsored.

2. Publication bias

It is well known that the published literature is not a random sample of scientific information, and that journal editors favor papers that report statistically significant results.[3] Nevertheless, it has been argued from theoretical first principles that despite statistical significance, most published research findings are false (Ioannidis 2005). Though it may be true that all numbers are interesting (Martin 2002), not all numbers are published because some numbers are more interesting than others. More interesting numbers include those that suggest statistically significant positive associations between ozone and health effects and endpoints.[4]

This phenomenon has three subspecies: *positive-results bias*, *outcome-reporting bias*, and *inferential exaggeration.* Positive-results bias occurs because studies that do not show positive associations are published less frequently. This arises because researchers and editors (and sponsors, such as pharmaceutical companies and regulatory agencies) earnestly desire to find positive effects. They can be motivated by the value of drug patents, air pollution control policy preferences, an affinity for subpopulations such as asthmatics or children, or just the honest desire to find what they are looking for.

Positive-results bias is hard to show empirically because the array of studies with nonpositive results that could have been published but were "left in the file drawer" can never be known with certainty. However, Ioannidis (2005) has shown that positive results bias is endemic in epidemiological literature, and it is magnified by manipulation of the analyses or the reporting of findings. He uses a model that captures various assumptions about power, the ratio of true to

---

[3] Writing 40 years ago, Sir Austin Bradford Hill bemoaned the insistence on statistical significance as a selection criterion. See III.C.6.

[4] Nonpositive results are inherently less interesting because they do not reject the null hypothesis. Negative results may be interesting because they are evidence of error -- positive effects from ozone exposure being biologically implausible -- but journal editors are not interested in publishing error.

not-true relationships in the literature, and bias (defined as the proportion of probed analyses that would not have been "research findings" but were presented as such anyway) to calculate the positive predictive value (PPV). PPV reaches 0.85 under plausibly best-case conditions of high power (0.80), a 1:1 ratio of true to not-true results in the literature, and bias of just 0.1. For a meta-analysis of small inclusive studies each with high power (0.80), a 1:3 ratio of true to not-true studies in the literature, and moderate bias (0.40), PPV declines to 0.41, making them likely to be false. PPV is much lower when the purpose of performing meta-analysis is to overcome the low power of individual studies. Values for PPV drop well below 0.01 for exploratory epidemiology. Thus, most research results are false for most research designs.

Even in a research area stipulated to have no effects, false positive results can be expected to dominate and reflect publication bias:

> For example, let us suppose that no nutrients or dietary patterns are actually important determinants for the risk of developing a specific tumor. Let us also suppose that the scientific literature has examined 60 nutrients and claims all of them to be related to the risk of developing this tumor with relative risks in the range of 1.2 to 1.4 for the comparison of the upper to lower intake tertiles. Then the claimed effect sizes are simply measuring nothing else but the net bias that has been involved in the generation of this scientific literature. Claimed effect sizes are in fact the most accurate estimates of the net bias. It even follows that between "null fields," the fields that claim stronger effects (often with accompanying claims of medical or public health importance) are simply those that have sustained the worst biases (Ioannidis 2005, 700).

Outcome-reporting bias occurs when researchers report results with the highest apparent association, an almost universally observed phenomenon. Sometimes, dozens of models will have been examined but only the handful with the strongest association will be reported. A variety of motivations can be at play here, ranging from overt or covert policy preferences in favor of more stringent air pollution control regulations to a policy-neutral desire for professional advancement, which is enhanced by publishing positive associations.

From an information quality perspective, agencies that utilize scientific information known to contain outcome-reporting bias have a minimum obligation to ascertain and report its severity. This is not as burdensome as it might seem, because in any specific situation (such as the ozone NAAQS revision) only a handful of the hundreds of studies published are critical for EPA's synthesis. For each critical study, EPA should determine the extent to

which nonpositive outcomes were not reported and include that information in its presentation. Failing to do so leads decision-makers and the public to believe that the reported results are representative rather than extraordinary.

Outcome-reporting bias is evident in several of the studies on which EPA relies heavily. For example, important studies on respiratory symptoms report only the most statistically significant results (Gent et al. 2003; Korrick et al. 1998; Mortimer et al. 2002). The extent to which these results are representative of all the models analyzed is not clear, nor is it known how many different models the authors examined before settling on the ones they published.

Inferential exaggeration occurs when scientists draw (and editors accept) conclusions that are not supported by the data and analysis actually performed. Because they are in the business of conducting research, scientists as a group are predisposed to be cautious about drawing inferences that go beyond their data and analyses. However, because they also have opinions about policy and face other incentives, sometimes they do not follow these professional norms and instead exaggerate the strength or certainty of their results, or the implications of their results for public policy.

Two examples of inferential exaggeration are notable in the ozone epidemiology literature. Gent et al. (2003) concluded that asthmatic children using maintenance medication are particularly vulnerable to ozone, even after controlling for exposure to fine particles, at levels below the current standard. But this conclusion goes beyond what can be inferred from the reported data and analysis. Peak ozone exposures, which logically drive the results observed, exceeded current standards.

Mortimer et al. (2002) also concluded that ozone below current standards has adverse effects on asthmatic children. This is based on selective reporting of model results: an odds ratio of 1.16 for4-day average ozone, 1.26 for 2-day average PM_{10}, 1.32 for 2-day average SO2, and 1.48 for 6-day average NO2, all obtained from single pollutants models. Odds ratios were barely statistically significant (lower 95% CIs ranged from 1.00 to 1.03). It cannot be discerned whether any of these ratios would have retained statistical significance, or declined in magnitude, if estimated using a multipollutant model or if departures from random sampling were taken into account. Finally, there is no evidence that the fundamental data quality problems with self-reported respiratory testing found by Kamps et al. (2001), discussed below, were even recognized, much less addressed.

3. Methodological error

No research study is perfect, and information quality guidelines neither require nor expect perfection as the standard of objectivity. But some methodological errors are so severe that that they have a material effect on utility, particularly for regulatory decision-making.

Prominent examples of methodological error can be found in several important studies. For example, in studies of respiratory symptoms:

- Repeated statistical tests are performed without apparent regard for the resulting increase in the rate of false positives (Korrick et al. 1998; Mortimer et al. 2002)
- Statistically significant but biologically implausible lags are reported (Mortimer et al. 2002)
- Single rather than multipollutant models are emphasized (Korrick et al. 1998; Mortimer et al. 2002)
- Known confounders are inadequately controlled (Gent et al. 2003; Korrick et al. 1998)

Methodological errors also arise with respect to several data measures, including provably unreliable self-administered peak expiratory flow rate (PEFR) testing (Mortimer et al. 2002); selective use of data from different times of day (Mortimer et al. 2002); and the reliance on (subjective) symptoms rather than (objective) signs (Gent et al. 2003). Korrick (1998) uses a nonstandard measure of incremental change (% per 50-100 ppb ozone) that is too large to have any practical utility for making policy decisions over small exposure ranges.

The data quality problems associated with PEFR tests deserve special attention because they are methodologically crucial to several of the epidemiological studies. To obtain reliable data, the procedure requires both training of the person administering the test and practice by the subject, who also must be willing and able to cooperate. Because of the learning effect, multiple tests are necessary to obtain clinically reliable information. However, clinical practice appears to utilize the data in a largely arbitrary manner (Queen's University Clinical Skills Education Centre 2007).[5]

[5] An unspecified degree of consistency in readings is sought, and the precision of measurement is technology dependent, with manual devices having less precision than electronic units. It is also unclear how much precision is in the test irrespective of measurement technology. Once obtained, data can be

Air pollution studies typically are conducted outside clinical settings, where personal data are not collected by medical professionals or technicians but by subjects themselves (or their parents, if the subjects are children). Subjects must be trained to self-administer tests correctly and motivated to report data accurately for data collected in home settings to have the same information quality as data collected clinically. For that reason, it is vital that PEFR and similar measurements obtained outside clinical settings be validated for reliability and validity.

Kamps et al. (2001) studied 40 asthmatic children aged 5-16 years to ascertain the validity of PEFR data self-reported over four weeks by patients and their parents. Data were obtained by diary and, unbeknownst to subjects, microchip memory recorders within the PEFR meters. The simultaneous collection of self-reported and automated data from the same individuals over the same time period provided a powerful test of validity and reliability. Stated compliance with the data collection protocol was 96%, but actual compliance averaged 77%, declining significantly over the course of the study. For 12.5% of the subjects, actual compliance was less than 50%, meaning that Nonresponse was systematic and not random. Further, self-reported data were correct (as measured by microchip) only about half of the time. In about one-fourth of all cases, self-reported data were invented. Self-reported data were biased toward understating electronically measured respiratory performance.[6]

Kamps and coworkers reasonably concluded that self-reported PEFR data were unreliable, and that electronic meters should be used instead of diaries. Similar data obtained by ozone researchers could have been much more reliable

---

averaged -- presumably arithmetically, though the actual shape of the distribution of test values may not be Gaussian. Low values might be discarded on the ground that they are trials, but this could upwardly bias the recorded data. Queens University recommends recording the highest of three measurements with what appears to be two significant figures, though it is not clear why the highest is obviously "correct" or that the second figure is actually meaningful.

[6] This degree of nonresponse, and the problem of manufactured data, likely would have prevented EPA from obtaining permission to collect such data or sponsor its collection. Response rates below 70% trigger a requirement for nonresponse analysis to ascertain the extent of bias and enable the development of a plan to reduce and manage it (Office of Management and Budget). Without such plans, approval to collect the information likely would not be granted.

than what Kamps et al. found, but based on this evidence it is clear that ensuring elementary data quality in respiratory effects studies is a challenging task. In none of the studies cited by EPA is there persuasive evidence reported that these problems were prevented, or that the researchers had a good grasp of their magnitude. This is especially worrisome because errors from self-reporting could easily exceed the incremental differences in performance (10%) that researchers are looking for.

4. Peer review practices

Journal and governmental peer review have several fundamental differences. For example, when researchers submit manuscripts for journal publication, they neither control the selection of peer reviewers. On the contrary, agencies directly or indirectly control the selection of scientists who peer review their work products. Members of the Clean Air Scientific Advisory Committee (CASAC) are selected by the EPA Administrator, based on advice he receives from EPA career staff.

Scientists submitting manuscripts to journals also do not control the decisions of editors whether to publish. In contrast, government peer review panels rarely, if ever, can decline to accept for publication a manuscript they review. For that reason, they can only recommend that changes be made, and they have no power to enforce their recommendations. Agencies can, and often do, decline to implement the recommendations they receive from peer reviewers.[7]

Journals also differ from government with respect to data disclosure. Agencies are required to disclose all data, assumptions, models and methods sufficient to make their work products "capable of being substantially reproduced" (Office of Management and Budget 2002, 8460), but often fail to do so. Journals are not subject to this requirement, but sometimes choose to follow it anyway.[8]

---

[7] In its June 2006 letter to the Administrator, CASAC recommended that significant changes be made in EPA's ozone exposure analysis. See (Clean Air Scientific Advisory Committee 2006b). EPA did not implement these recommendations.

[8] As a condition of acceptance, the *American Economic Review* requires that authors make all their data available upon request. See the AER's policy on data availability at http://www.aeaweb.org/aer/data_availability_policy.html. This is not common practice among biomedical journals that publish health effects

In Section V below, we discuss the peculiar characteristics of the Clean Air Scientific Advisory Committee (CASAC) as a peer review body. We show that its statutory charge includes both reviewing EPA's risk assessment and providing policy advice. CASAC's policy advice function confounds its scientific review, making it difficult – and, in some cases, impossible – to discern when it is performing scientific review and when it is delivering policy advice. Nevertheless, EPA is obligated under information quality principles to make this discernment in its review and synthesis of CASAC reports and commentary – an obligation that would have been much easier to meet if the Agency had explicitly asked CASAC to adhere to information quality principles from the outset.

B. Non-disclosure of critical studies and analyses

EPA has excluded certain studies relevant to determining Policy Relevant Background (PRB) even though they were published in time for consideration (Ortmans et al. 2006; Vingarzan 2004). These studies suggest that PRB is higher, and more variable, than the values EPA uses, so their exclusion is evidence of bias.

A more noticeable instance of bias is EPA's failure to disclose at least one critical collection of information within its control: its reanalysis of the data obtained by Adams (2006a). The information is summarized in an internal staff memorandum dated just six days before the Administrator signed the NPRM (Brown 2007, "Brown Memorandum"). Even though this reanalysis is a crucial element of Agency staff policy recommendations, EPA did not disclose enough information to make it reproducible.

This reanalysis is fully subject to information quality standards and does not benefit from the weak rebuttable presumption of objectivity because it has not been peer reviewed. Moreover, because it reaches conclusions opposite of the researcher, it is equivalent to a new study inserted into the record in a discriminatory fashion. It is beyond dispute that EPA would not have accepted a new analysis of the Adams data submitted by a third party on June 14, 2007, unless perhaps it supported the staff's policy recommendations. EPA clearly displays a discriminatory preference for data and analyses that support staff risk management preferences, an obvious information quality defect.

---

studies. To the extent that an economist wishes to publish a study that utilizes health effects data as inputs, but such data have not been publicly disclosed, he cannot publish in the flagship journal of his profession.

The Brown Memorandum itself exemplifies multiple types of violations of the information quality standard of objectivity. First, it is a post hoc statistical analysis conducted on data whose initial analysis did not support the declared policy preferences of Agency staff. Second, it was prompted by an admittedly low-quality analytic review ("visual comparison" and "cursory evaluation," p. 3). Third, it was structured in a fundamentally different way than other research studies: for the express purpose of justifying a stated policy preference for minimizing Type II error (failing to reject the no-effect hypothesis when in fact it is false, pp. 5ff). The description of the reanalysis is so wedded to this policy preference that the line between science and policy is not merely blurred, but obliterated.

Ioannidis (2005) shows that positive results are usually false even when researchers exercise normal restraint with respect to Type I error (rejecting the no-effect hypothesis when in fact it is true). In the reanalysis summarized in the Brown Memorandum, EPA has discarded that restraint in favor of a statistical approach that has the peculiar quality of dramatically increasing Type I error, the likelihood of concluding that ozone exposure below the current standard poses health risks when it does not.

In the NPRM, EPA also acknowledges implicitly that Agency staff used the Adams data for purposes that were never intended by the study design. It is reported that two of the 30 healthy adult subjects experienced exercise-adjusted FEV_1 decrements exceeding 10% at 0.060 ppm using one of the two exposure patterns examined (but not the other) (U.S. Environmental Protection Agency 2007b, 37828, n. 16). Of course, the number could well have been even larger by chance if the sample size had exceeded 30. This implies that the optimal sample size for EPA is whatever number is needed to observe the effect the Agency is looking for. It is inappropriate to obtain a sample, subject its members to a well-designed test, learn that the sample does not yield hoped-for outcome, and in response, abandon the sample in favor of focusing on selected individuals within it. If EPA can find a reputable statistical authority for this procedure, the agency should make its identity known.[9]

Federal information quality guidelines require transparency. Moreover, without transparency the public is placed at an overwhelming disadvantage insofar as it cannot reasonably be expected to provide informed public comment on scientific information that the Agency has not disclosed. When, as in this case, the disclosed portion of information shows the hallmarks of purposeful bias, the

[9] In the Brown Memorandum, the source of statistical authority is EPA staff.

only responsible default inference is that the reanalysis does not meet applicable information quality standards. Unless EPA can show otherwise, the reanalysis should not be disseminated and the Brown Memorandum should be withdrawn.

C. EPA Interprets and Presents Scientific Information in a Systematically Biased Manner

The Criteria Document, Staff Paper, and Notice of Proposed Rulemaking all collect, summarize and synthesize scientific evidence, much of it published in peer-reviewed journals. The challenge under applicable information quality guidelines is ensure that this information is accurate, reliable, and unbiased, and presented in an accurate, clear, complete, and unbiased manner. Each document displays evidence of both substantive and presentational bias, and it appears to increase in the progression from Criteria Document to Staff Paper to NPRM.

Interpretative bias arises in several forms. We discuss several below.

1. *The inclusion or exclusion of data or studies based on the extent to which they support stated or unstated risk management objectives*

The inclusion of EPA's reanalysis of the Adams data also is evidence of purposeful bias because the reanalysis supports staff policy recommendations and Adams' own analysis does not. It is a violation of information quality principles to choose a conclusion first, then fill in behind with data and analysis that support it. A risk assessment performed this way cannot be unbiased, either in substance or presentation.

2. *The inclusion or exclusion of data or studies based on post hoc or non-transparent criteria*

So far, we have set aside whether the 10% threshold for FEV_1 decrements cited in the Brown Memorandum is even appropriate. It turns out that this threshold differs significantly from the clinical thresholds recommended by the American Thoracic Society (15%) and CASAC (15-20%). In short, the 10% decrement is a post hoc threshold apparently chosen for compatibility with EPA staff policy recommendations. In the Staff Paper, EPA says CASAC agrees with its 10% threshold, but there is strong evidence suggesting that any such agreement was at best a minority view based on policy considerations, not science. CASAC panel member Dr. Sverre Vedal was succinct:

> While most attributions to CASAC are correct, I don't believe it was a written opinion of CASAC that "more emphasis should be placed on numbers of subjects in controlled human exposure studies with FEV1 decrements greater than 10%, which can be clinically significant, rather

> than on the relatively small average decrements" (p. 6-43). While this may have merit in some (or even many) situations, for example when noting that 26% of individuals had > 10% FEV1 decrements at 0.08 ppm… , in other cases, such as the specific case of 0.060 or 0.040 ppm exposures, this approach amounts to attempting to find effects in a very few individuals when the statistical tests are not significant, which is a dangerous precedent – especially in this case where we are looking at small effects in 3 of 30 vs. 1 of 30, a pitiful number on which to attempt to base policy… (Clean Air Scientific Advisory Committee 2007, C-30, [internal citations omitted]).

EPA did not correct the misstatement in the final Staff Paper (U.S. Environmental Protection Agency 2007f, 6-43).

3. Mischaracterization of results

Scientific results can be misrepresented many ways, and several of these ways are evident in EPA's risk assessment documents.

a) Characterizing a study as "new" since the last ozone NAAQS review when in fact it was part of the last review

EPA cites a long list of clinical studies performed to estimate the effect of ozone at different doses and under various conditions. Many of these studies are not new, however, and were cited in the Agency's risk assessment documents supporting the 1997 NAAQS revision. It is misleading to report "old" studies as if they were new.[10]

The Administrator's statutory duty under Clean Air Act §§ 108-109 is to utilize the best available *new* scientific information. To maximize the utility of its scientific reports to the Administrator, EPA should segregate "old" from "new" science to ensure that the two categories are not confused, and discuss "old" studies only to set the stage for its review of "new" studies. None of the studies cited in this round of review that also was cited in the previous round has utility for the Administrator's task of deciding whether to change the standard – with one important exception, discussed in b) below.

[10] An "old" study can become "new" if, since the date of the last NAAQS review, new analyses of the data were performed. In that case, however, it is the new study or reanalysis that is relevant to the current review, not the original.

b) Interpreting an "old" study differently than it was interpreted in the last ozone NAAQS review without providing a credible explanation for the discrepancy

An old study also can become "new" if, since the date of the previous review, the Agency has learned about a material error. This could be error in the original data (e.g., transcription error), error in a statistical analysis (e.g., GAM convergence issues), or error in its interpretation. It does not matter whether the Agency learned about an error on its own, through a subsequent publication (e.g., (Health Effects Institute 2003)), or a Request for Correction under applicable information quality error correction procedures.

However, EPA is obligated to present an objective explanation for the discrepancy. It cannot simply ignore errors it knows about in previously published research. Also, it cannot simply describe, characterize, or analyze an "old" study in a contrary manner without acknowledging and explaining why its previous review was wrong.

c) Characterizing a study as reporting something about which it is silent

In the NPRM, EPA states that results from "numerous" multi-city and single-city studies show that the association between ozone and mortality "do not appear to be changed in multipollutant models including PM_{10} or $PM_{2.5}$ (U.S. Environmental Protection Agency 2007b, 37839). These "numerous" studies consist of the NMMAPS studies, none of which has daily $PM_{2.5}$ data. The associations in these studies "do not appear to be changed" primarily because they only measure PM_{10}.

In the ozone risk assessment, there are two superlative examples of this bias. The first involves EPA's interpretation of studies by Suresh Moolgavkar and coworkers (Moolgavkar 2000; Moolgavkar et al. 1995). EPA represents the results of these studies in ways that the corresponding author says are incorrect (Moolgavkar 2007, 4-5). It is possible that EPA is right and Moolgavkar is wrong. However, it is Moolgavkar who enjoys the weak presumption of objectivity under applicable information quality standards. EPA cannot legitimately interpret Moolgavkar's work in ways with which he disagrees and simultaneously claim the mantle of presumptive objectivity that Moolgavkar's studies enjoy by virtue of having been peer reviewed. To rebut Moolgavkar, EPA is required to persuasively show that he is wrong. In none of EPA's discussions of Moolgavkar's work, however, does the Agency provide such a rebuttal.

EPA has done a similar thing with respect to the studies by William Adams and coworkers (Adams 2002, 2003a, 2006a, 2006b). EPA has altered

Adams' results and reanalyzed them (Brown 2007; U.S. Environmental Protection Agency 2007b, 2007f) in ways with which he has publicly disapproved (Adams 2007), and retained its modified results despite his objections (U.S. Environmental Protection Agency 2007b). Federal information quality guidelines confer on Adams' peer reviewed published studies a weak presumption of objectivity. EPA is entitled to make the effort to rebut this presumption, but to do so the Agency must make a persuasive showing that Adams has misreported or misinterpreted the results of his own work. EPA did not do this in the Staff Paper (U.S. Environmental Protection Agency 2007f), and so it did not satisfy even this limited burden of proof.

In both cases, the practical effect is to increase EPA's estimated magnitude of ozone health risks (leading to upwardly biased risk perception), and narrowed the estimated confidence interval (leading to the perception that the higher risk estimate is more rather than less certain). Both phenomena are evidence of bias, which is incompatible with the objectivity standard.

d) Characterizing a study as reporting something when it reports the opposite

Sarnat et al. (2001) measured personal multi-pollutant exposures and corresponding ambient exposures for 56 subjects in or near Boston. Ambient and personal exposures were uncorrelated for all subgroups. Sarnat et al. (2005) measured personal multi-pollutant exposures in 43 subjects in or near Baltimore. Ambient and personal exposures were uncorrelated in the winter, correlated in the summer for eight of 21 subjects but uncorrelated for the others. As in Boston, correlations were observed with $PM_{2.5}$. In 2001, the authors concluded that ambient concentrations did not provide a surrogate for personal exposures, and in 2005 they reconfirmed that result.[11]

[11] "Ambient concentrations of gaseous pollutants cannot be considered as surrogates for their respective personal exposures without site-specific evidence to support that assumptions"(Sarnat et al. 2001); "Our results support the earlier findings that summertime gaseous pollutant concentrations may be better surrogates of personal $PM_{2.5}$ exposures (especially $PM_{2.5}$ of ambient origin) than they are surrogates of personal exposure to the gases themselves" (Sarnat et al. 2005).

In the Staff Paper, the study sites were reversed and the results reported incorrectly:

> The first study conducted in Baltimore, MD [*sic,* actually Boston] observed no relationship between ambient concentrations and personal exposures in both the summer and the winter. However, the second study conducted in Boston, MA [*sic*, actually Baltimore], found statistically significant associations between ambient O_3 concentrations and personal exposures to O_3.
>
> … Collectively, these studies observed that the daily averaged personal O_3 exposures from the population were well correlated with ambient O_3 concentrations despite the substantial variability that existed among the personal measurements. Averaging likely removes the noise associated with other sources of variation. *These studies provide supportive evidence that ambient O_3 concentrations from central monitors may serve as valid surrogate measures for mean personal exposures experienced by the population*, which is of most relevance for time-series studies (U.S. Environmental Protection Agency 2007f, 3-40, emphasis added).

In the NPRM, EPA propagated its error:

> Studies by Sarnat et al. (2001, 2005) that included this susceptible group ["those who suffer from chronic cardiovascular or respiratory conditions"] reported mixed results for associations between ambient O_3 concentrations and personal exposures to O_3. Collectively, these studies observed that the daily averaged personal O_3 exposures tend to be well correlated with ambient O_3 concentrations despite the substantial variability that existed among the personal measurements. These studies provide supportive evidence that ambient O_3 concentrations from central monitors may serve as valid surrogate measures for mean personal exposures experienced by the population, which is of most relevance for time-series studies. A better understanding of the relationship between ambient concentrations and personal exposures, as well as of the other factors that affect relationship will improve the interpretation of concentration-population health response associations observed (U.S. Environmental Protection Agency 2007b, 37838).

e) Selective and misleading citation

Among sources of external authority for EPA staff, CASAC ranks at the top because of its statutory role in reviewing the science (and giving policy advice on the standard). In several places the NPRM cites supporting statements from CASAC, and this conveys the impression that CASAC endorsed all or

virtually all of the EPA staff analysis. CASAC also made comments that do not support the EPA staff interpretation of the science, but most of these statements do not appear in the NPRM. Selective citation is a form of presentational bias.

Below we list the few examples we can find in the NPRM in which a CASAC statement appears and EPA has cited it. The complete paragraph is provided; the underlined text is quoted in the NPRM:

- Since it is unlikely that each of these pollutants will have similar short-term effects on mortality, these findings suggest that while the time-series study design is a powerful tool to detect very small effects that could not be detected using other designs, it is also a blunt tool. The Clean Air Act requires that NAAQS be set for individual criteria air pollutants using the best available science. Because results of time-series studies implicate all of the criteria pollutants, findings of mortality time-series studies do not seem to allow us to confidently attribute observed effects specifically to individual pollutants. This raises concern about the utility of these types of studies in the current NAAQS-setting process and could serve to motivate interest in taking a broader perspective on regulating air pollution that incorporates the entire mixture of community air pollutants (Clean Air Scientific Advisory Committee 2006b, 3).
- Time-series studies typically make use of data from available air pollution monitoring network sites in which concentrations of various subsets of the criteria pollutants are measured. Study findings focus on identification of associations between day-to-day variation in these concentrations and daily mortality. Not only is the interpretation of these associations complicated by the fact that the day-to-day variation in concentrations of these pollutants is, to a varying degree, determined largely by meteorology, the pollutants are often part of a large and highly-correlated mix of pollutants, only a very few of which are measured. For the ozone and other photochemical oxidant NAAQS, this pollutant mix includes a large number of both gas- and particle-phase photochemical oxidant pollutants. Unfortunately, we have only limited information on the specific chemical composition, toxicity and, equally importantly, the population exposure of oxidant pollutants other than ozone (Clean Air Scientific Advisory Committee 2006a, 3).
- From the perspective of the epidemiological data, the Ozone Panel judged the selection of: respiratory symptoms in moderate/severe asthmatic children (ages zero [birth] to 12); hospital admissions for respiratory illness among asthmatic children; and premature total non-accidental and cardiorespiratory mortality for inclusion in the quantitative risk

> assessment to be appropriate. However, the CASAC believes that several other endpoints should be discussed qualitatively to support the findings that these endpoints indicate that significant adverse effects are occurring at exposure concentrations well below the current standard. Other endpoints deemed worthy of additional discussion included respiratory emergency department visits among asthmatics and patients with other respiratory diseases, increased medication usage, and increased symptomatology reported at exposure levels well below the current standard. Taken together, members of the Ozone Panel felt strongly that these findings preclude including the current standard as a scientifically defensible option for the Administrator (see discussion about Chapter 6 found in the main portion of the letter above). (Clean Air Scientific Advisory Committee 2006b, 12)

We have also found important scientific comments from CASAC that do not support EPA's interpretation of the data but which do not appear in the NPRM. For each we have identified below in *italics* the critical element in the EPA risk assessment that is undermined:

- **Error in Estimating Exposure to Ozone:** The Ozone Staff Paper should consider the problem of exposure measurement error in ozone mortality time-series studies. It is known that personal exposure to ozone is not reflected adequately, and sometimes not at all, by ozone concentrations measured at central outdoor monitoring sites. Typically, personal exposures are much lower than the ambient concentrations, and can be dramatically lower depending on time-activity patterns, housing characteristics and season. In addition, and of particular importance for the ozone time-series studies, there can be no correlation between personal concentrations of ozone measured over time and concentrations measured at central outdoor sites. The population that would be expected to be potentially susceptible to dying from exposure to ozone is likely to have ozone exposures that are at the lower end of the ozone population exposure distribution, in which case this population would be exposed to very low concentrations of ozone indeed, and especially so in winter. Therefore it seems unlikely that the observed associations between short-term ozone concentrations and daily mortality are due solely to ozone itself (Clean Air Scientific Advisory Committee 2006a, 3-4).
  - *Data from ambient monitors yield upwardly biased estimates of risk.*
  - *Personal ozone exposure cannot be correlated with ambient exposure, especially in subpopulations of concern, so associations between data from ambient monitors and mortality are spurious.*

- *Short-term studies do not provide supporting evidence that positive associations in time series results are true.*

- Another implication of ozone measurement error that is relevant to the NAAQS-setting process is that this degree of measurement error would be expected to have a substantial impact on the ability to detect a threshold of the concentration-response relationship below which no ozone effects are discernible. Pollutant exposure measurement error obscures true thresholds in the concentration-response relationship, and this effect worsens with increasing degrees of measurement error. Since threshold assumptions are incorporated in the Agency's risk assessment and risk analyses, this issue will need to be addressed (Clean Air Scientific Advisory Committee 2006a, 4).

  - *The inability to detect a threshold for ozone may be the result of measurement error.*

- Another problem in the health effects calculations (see Table 5-5 and 5-11) is that they are based on computations of the form $R_x - R_{PRB}$, where R_x is the risk at a given concentration x of O_3 and R_{PRB} is the corresponding risk at policy-relevant background (PRB) for O_3. As discussed at the Ozone Panel's August meeting, the PRB is highly-problematic to calculate and is, in some sense, "unknowable." One can avoid this problem by calculating the $\Delta = R_{0.8} - R_x$ for various concentrations x. This form would allow focus on the question, "What is the difference in the expected number of health effects that will occur at various concentrations of O3, relative to the current standard of 0.08?" A key advantage of Δ is that it does not depend on the choice of PRB, and thus is free of the uncertainties surrounding estimation of PRB (Clean Air Scientific Advisory Committee 2006b, 13).

  - *Incremental risk reductions can be estimated without imposing a scientifically dubious and policy-laden Policy Relevant Background constraint.*

- At least two questions arise from these observations that are relevant to the ozone NAAQS-setting process: (1) What chemical agent or agents are at least partly responsible for the observed associations between ozone and mortality in the time-series studies? And (2) Do we require an immediate answer to the question of whether ambient ozone adequately serves as a surrogate marker, that, when controlled, effectively mitigates health impacts of this entire mix of pollutants? One possible explanation for the observed associations of ozone is that ozone itself serves as a marker for other agents that are contributing to the short-term exposure effects on mortality. This would require that outdoor concentrations of these agents are correlated over time with outdoor ozone concentrations,

which is to be expected if they are products of the same process that lead to ozone formation, and that these outdoor pollutant concentrations are better correlated with personal exposures than is the case for ozone itself (Clean Air Scientific Advisory Committee 2006a, 4)

- *Attributing the observed associations between ozone and mortality in time series studies masks the underlying factor(s) actually responsible.*
- *If ozone is not the relevant factor, and its control will not serendipitously reduce the relevant factor(s), then the risk reductions predicted by EPA will not be realized.*

f) Drawing inferences from a study that are not supported by the data and analysis reported

EPA says controlled human exposure studies provide compelling evidence that ozone exposure below the current ozone NAAQS causes lung function decrements, inflammation, and respiratory infection. However, the vast majority of the studies that EPA cites involve exposures at or above the current standard. EPA provides only a quasi-policy rationale for this inference, but that is impermissible under information quality principles. Policy officials have discretion over policy statements, but scientific statements must be supported by science.[12]

Five clinical studies have been performed since the last review, but only three involve exposures below the current standard (Adams 2002, 2003b, 2006a). EPA's interpretation of Adams' work diverges, and is inconsistent with, that of the author (Adams 2007). Adams reports no statistically significant effects from ozone at 0.040 ppm, but EPA finds them in its reanalysis. This poses special information quality issues. As indicated above, EPA's interpretation of Adams' work does not enjoy a rebuttable presumption of objectivity. Moreover, to utilize Adams' work in contrary ways, the Agency must first rebut the presumption of objectivity that attaches to these studies.

Furthermore, EPA's reanalysis of Adams' data in the final Staff Paper (U.S. Environmental Protection Agency 2007f) and Brown Memorandum (2007) is neither transparent nor capable of being substantially reproduced. This is an

[12] The statement "Chemical X causes cancer in humans" is scientific, and thus must be supported by high-quality, best available scientific evidence. EPA cannot disseminate this statement supporting it only with agency policy. EPA lacks the statutory authority to redefine as policy any "representation of knowledge such as facts or data" and thus evade the requirements of the Information Quality Act. See § V.5, (Office of Management and Budget 2002).

essential prerequisite for adherence to the objectivity standard and inconsistent with EPA's information quality guidelines.[13] The NPRM is even less transparent because of the complexity of the single-sentence description of the reanalysis and EPA's use of passive voice, which leaves entirely nontransparent its motivation, protocols, procedures, and execution:

> A combined data set including individual level data from the Folinsbee et al. (1988), Horstman et al. (1990), and McDonnell et al. (1991) studies, used in the previous risk assessment, and more recent data from Adams (2002, 2003, 2006) *have been used to estimate* probabilistic exposure-response relationships for 8-hour exposures under different definitions of lung function response (i.e., ≥10, 15, and 20 percent decrements in FEV1). As discussed in the Staff Paper (EPA, 2007, p. 5–27), while these specific controlled human exposure studies only included healthy adults aged 18–35, findings from other controlled human exposure studies and summer camp field studies involving school age children in at least six different locations in the northeastern United States, Canada, and Southern California indicated changes in lung function in healthy children similar to those observed in healthy adults exposed to O3 under controlled chamber conditions (emphasis added) (U.S. Environmental Protection Agency 2007b, 37857).

The public is limited to only the limited information EPA discloses, not what is required to reproduce EPA's analysis. EPA's reanalysis of the Adams' data is among the most influential items of information in the entire risk assessment. As such, it warrants the highest level of transparency, reproducibility, and independent scrutiny.[14]

---

[13] "EPA intends to ensure reproducibility for disseminated original and supporting data according to commonly accepted scientific, financial, or statistical standards… In addition, these Guidelines provide for the use of especially rigorous 'robustness checks' and documentation of what checks were undertaken. These steps, along with transparency about the sources of data used, various assumptions employed, analytic methods applied, and statistical procedures employed should assure that analytic results are "capable of being substantially reproduced" (U.S. Environmental Protection Agency 2002).

[14] There is no public evidence that this reanalysis was subjected to open, rigorous, independent and external peer review. It appears to originate in the July 2007 edition of the final Staff Paper (U.S. Environmental Protection Agency 2007f).

EPA's analysis of clinical data on cardiac effects is similarly problematic with respect to information quality standards. The published studies show no statistically significant increases in dozens of endpoints examined, with one exception. In a study of 10 nonmedicated[15] hypertensive patients and six healthy adult males, approximately two dozen cardiac measures were obtained (Gong et al. 1998). Only two statistically significant differences were observed: a clinically nonsignificant 6% reduction in FEV_1 and a greater than 10 mm Hg increase in alveolar-to-arterial PO2 gradient ($AaPO_2$). In a study with 20 separate group comparisons, the chance of finding one relationship that is statistically significant by chance is 100%. Nevertheless, the NPRM, EPA emphasizes the increase in AaPO2 and interprets it as evidence that ozone exposure "result[s] in an overall increase in myocardial work and impairment in pulmonary gas exchange" (U.S. Environmental Protection Agency 2007b, 38734). EPA says nothing about the relevance of the exposure level, which was 0.3 ppm -- 3.75 times greater than the current NAAQS, or the uncertainties implied by extrapolating to the population clinical data obtained from a sample of 16.

g) Utilizing for a more demanding purpose data that were collected for another purpose

The entire structure of the Information Quality Act, and the Paperwork Reduction Act in which it is codified, argues for taking great care to use and disseminate information for the purpose for which its collection was intended. A corollary to that principle is that agencies must be careful not to use or disseminate information in a manner that conflicts with these purposes.

The Adams' clinical studies (Adams 2002, 2003a, 2003b, 2006a, 2006b) were performed for the purpose of validating personal exposure test methods and using these test methods to examine different wave patterns in ozone exposure. They were not designed or intended for the purpose of estimating individual or population variability in response. Yet, EPA has used these data for precisely this alternative, unintended purpose.

Whether data collected for one purpose are appropriate for use in another depends on a careful examination of the alternative use. Whatever limitations apply to the original data must be extended to the alternative use. Neither the Staff Paper nor the NPRM provide any useful analysis on this point. Rather, it

[15] Although the abstract says the hypertensives were "nonmedicated," the text of the study describes them as "treated either pharmacologically for > 1 yr or by nonpharmacologic methods."

appears to have been simply assumed that data, once generated, can be used for any purpose without restriction.

h) Hypothesizing after the results are known

Unfortunately, it has become common practice to wait until data come in before deciding how to model them or what statistical procedures to apply. Hypothesizing after the results are known ("HARKing," in the terminology set forth by (Kerr 1998)) is not considered an approved practice precisely because it leads scientists toward false positives. Often, this is the result of data mining and similar statistical practices. When improper statistical tests are then performed, such as a seemingly unlimited set of pairwise comparisons, it is assured that some associations will be statistically significant.

Properly performed hypothesis-testing research requires researchers to specify *a priori* the hypotheses to be specified and the methods that will be used to test them. Improvisational data collection or statistical analysis after-the-fact are fine, but such research is properly described as hypothesis-*generating* rather than hypothesis-*testing*. The results of hypothesis-*generating* research should only be used to guide future hypothesis-*testing* research, and not be used to draw inferences – especially inferences that have significant public policy implications.

4. Study selection bias

Study selection bias is the selection of critical studies based on direction or strength of association rather than quality of data or analysis. In the documents subject to this RFC, EPA has displayed a systematic preference for studies that show positive associations even among studies that have important information quality limitations.

Several studies are available from which EPA could have chosen for its analysis of ozone with respect to asthmatics. EPA consistently selects studies that show positive associations with ozone (e.g., Gent et al. 2003; Mortimer et al. 2002) over studies that do not (e.g., Schildcrout et al. 2006), but does not establish an information quality basis for its selections.

We have been unable to identify any epidemiological study that EPA has interpreted as dispositive but which did not find a positive association. For epidemiological studies, only studies with positive associations matter. The recent chamber study by Adams (2006a) did not yield statistically significant effects at 0.040 ppm, which led EPA to focus on the two subjects who did and reanalyze the data in order to detect statistical significance.

The evidence is clear that EPA's risk assessment for ozone includes a large measure of study selection bias. The incremental magnitude of its effect on the resulting risk estimate could be infinite.

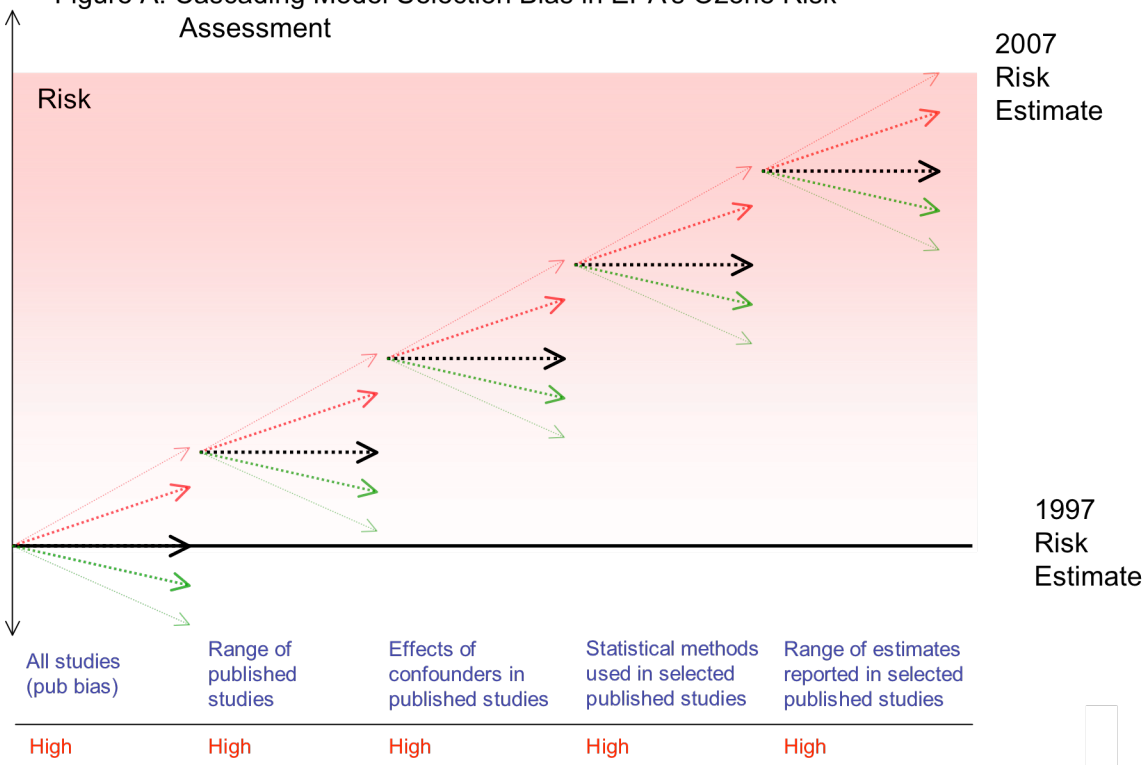
5. Model selection bias

Within any selected study, multiple models typically are tested and results reported. Model selection bias involves choosing which models to emphasize based on criteria other than information quality. Examples of model selection bias include:

- Selecting models based on criteria other than quality of data or analysis
- Selecting models known to yield upwardly biased risk estimates, such as single-pollutant models that do not control for known confounders
- In time series models, choosing lags based on statistically convenient but biologically implausible criteria
- Disseminating results from models known to yield risk estimates that are upwardly biased and more uncertain, such as Generalized Additive Models conducted with insufficient convergence criteria

Each of these biases can be found in a highly significant place in EPA's risk assessment documents. Figure A below illustrates just a few of the biases in model selection, beginning with publication bias (the predominance of positive

Figure A: Cascading Model Selection Bias in EPA's Ozone Risk Assessment



studies reported); the choice of which study to emphasize (with preference given to those with larger risk coefficients); inadequate control for confounders (leading to their effects being attributed to ozone); the choice of statistical methods (with some more likely to reject the no-effect hypothesis than others); and the choice of which estimate from which reported model to rely upon. The result is a cascade of bias such that the resulting risk assessment is much higher than would have been obtained using objective methods. It is also false reported to be more certain.

a) Selecting models based on criteria other than quality of data or analysis

Risk assessment begins with exposure, for without exposure there is no risk. EPA uses data from ambient monitors as proxies for personal exposure in studies attempting to estimate individual effects, even though readings from ambient monitors exceed those from personal monitors and are less variable. Concerns about the scientific merit of this approach have been raised (Clean Air Scientific Advisory Committee 2006b; Moolgavkar 2007), but these concerns have previously not been explained in information quality terms.

The literature consistently shows that personal ozone exposure is generally less than the values obtained from ambient monitors. This is especially true for members of sensitive subpopulations, who spend a disproportionately large fraction of their time indoors, often with filtered and cooled air. The only subpopulation for which ambient monitors may closely track personal exposure are those who work outdoors during the ozone season. However, there is virtually no overlap between this subpopulation and sensitive subpopulations. To adhere to the information quality substantive objectivity standard, EPA must provide unbiased estimates of exposure for each subpopulation of concern. EPA cannot simply assume that data from ambient monitors is unbiased, especially when it is known that they are not.

b) Control for known confounders other than air pollution

In both the literature EPA cites as dispositive and in its analysis of ozone's effects on asthmatic children, the effects of known non-air pollutant confounders are not adequately controlled. Sarafino et al. (2001) list 12 common asthma triggers: (1) air pollution, (2) cigarette smoke, (3), high humidity, (4) high/low environmental temperature, (5) allergens, (6) respiratory infection, (7), exercise, (8) nighttime hours, (9) stress or worry, (10) anger, (11) excitement, and (12) laughter. Several are correlated with air pollution, and air pollution consists of more than ozone. Failing to adequately control for known confounders yields upwardly biased estimates of risk.

In the ozone epidemiology literature, control for confounders such as meteorological conditions is spotty even though they are known triggers for asthma and other respiratory health effects. Ozone is correlated with both temperature and humidity. Failing to control for both leads to upward bias of the estimated risk because ozone and humidity are correlated.[16]

EPA is fond of and relies heavily on the longitudinal study by Gent et al. (2003) as supportive evidence for an effect from ozone independent of PM. However, the only non-air pollution confounder that appears to have been controlled for is temperature. EPA acknowledged this limitation in the Criteria Document (U.S. Environmental Protection Agency 2006a, 7-53),[17] but dropped it from the Staff Paper (U.S. Environmental Protection Agency 2007f, 3-82). In the NPRM, EPA restored (and perhaps expanded) proper notice of these limitations (U.S. Environmental Protection Agency 2007b, 37829),[18] though it is difficult to discern the extent to which these limitations figured in the risk assessment.

Direct control for the confounding effects of allergens appears to be very rare. In our review we did not find a single epidemiological study cited by EPA that did so. Some studies try to control confounding by exclusion, but because the propensity of those with asthma to also have allergies is great, the inclusion of asthmatics inevitably includes those with allergies. Cockroach and house dust mite allergens are both very important, the former particularly so among poor children. This literature is missing from both the Criteria Document and the Staff Paper; indeed, both documents consistently view ozone solely as an additive factor in the environment. The same appears to be true for exercise. In principle, if low-level ozone exposure can be shown to aggravate asthmatic symptoms in a subpopulation of children who live on upper floors of walk-up apartments heavily contaminated with cockroach feces, other allergens and tobacco smoke, ozone is presumed to be the causal factor explaining their condition.

Control for confounders almost always reduces risk estimates. Indeed, the general trend among epidemiological studies is that the better the control for

---

[16] Indoor air conditioning significantly reduces temperature, relative humidity, and pollutant exposure.

[17] "Study limitations include limited control for meteorological factors and the post-hoc nature of the population stratification by medication use."

[18] “Study limitations include the post-hoc nature of the population stratification by medication use. Also, the study did not account for all of the important meteorological factors that might influence these results, such as relative humidity or dew point.”

confounders, the smaller will be the relative risk and the less likely the effect will be statistically significant. Because relative risk ratios are typically small (i.e., < 2), the practical effect of better control for confounders may be to eliminate positive association, statistical significance, or both. It is not hard to imagine that certain confounders could have much larger effects than ozone.

c) Selecting models known to yield upwardly biased risk estimates, such as single-pollutant models and models that do not control for known confounders

For its assessment of risk for asthmatic children, EPA utilizes the study by Get et al. (2003) in ways that systematically overstate risk estimates (U.S. Environmental Protection Agency 2007c, 2007d). Data from ambient monitors were used. Asthmatics are assumed to be exposed at the same level as other children. Symptom data come from locations where ozone levels exceeded the current standard, and are subjectively reported without any validation. Meteorological confounders except for maximum daily temperature are not addressed. (Despite these biases, the association the researchers estimated was small and barely significant [lower 95th percent CI = 1.00]).

A fundamental principle of risk assessment is that the same risk cannot be simultaneously attributed to multiple causes without double counting, and double counting is inherently unjustified. Considering only air pollutants, in principle risk could be attributed to each constituent in the mixture or to an interactive effect resulting from joint effects. Consider a simple model with two pollutants, x_1 and x_2, and a set of confounders, C. All linear effects are fully described by the model:

$$\text{Risk} = \alpha + f(C) + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon \, .$$

The term x_1x_2 is an interactive term representing the effects from joint exposure; the coefficients β_1, β_2, and β_3 are the slopes of these effects per unit of pollutant; and ε is the error term. If this model provides an unbiased estimate of risk, then removing any of the terms gives upwardly-biased estimates of the others. For this reason, all single pollutants models are biased, as are all models that fail to adequately control for confounders.

EPA's decision to disseminate risk estimates derived from single-pollutant models as estimates of risk, even when results from multi-pollutant models are available, is an unambiguous violation of the objectivity standard.

d) In time series models, choosing lags based on statistically convenient but biologically implausible criteria

Time series models take into account the extent to which health effects are delayed (i.e., "lagged") after exposure. The correct lag structure is not known

with certainty. Nevertheless, there are at least three critical biological principles that must be respected:

- All health effects must occur after exposure

It is biologically impossible for the health effects of ozone to materialize prior to exposure. Any lag structure that permits risk to occur prior to exposure is simply wrong.

- For each health effect, the gradient of risk must be biologically appropriate

Risk cannot "bounce" all over the place, but instead must follow a biologically sensible dose-response relationship. Moreover, the proper lag structure cannot be chosen based on statistical criteria that have no biological justification.

- More severe health effects must occur subsequent to more minor effects

Health effects can be arrayed by severity, and the choice of lag structure must be at least ordinally consistent. Typically, respiratory symptoms occur before emergency department visits, which occur before hospital admissions.

The time series studies EPA relies upon do not respect these fundamental biological requirements, and thus they sacrifice the weak presumption of objectivity they otherwise would enjoy under applicable information quality standards. Lags for specific health effects have been selected based on statistical strength without regard for the underlying biology, a procedure that yields upwardly-biased risk estimates (Moolgavkar 2007, 6-7). Moreover, this has led to incoherence in lags across health effects, in which more severe health effects are implied to occur before milder ones.

The study by Mortimer et al. (2002) is symptomatic of this constellation of information quality defects, in large part because of the authors' heroic statistical efforts to find positive associations:

- Multiple lag times (and multiple constructions of the lag) are used without biological justification. Despite the breadth of this effort, statistically significant associations are found for a 4-day lag, but not for lags of 1-, 2-, 3-, 5-, or 6-days.
- Self-reported symptom scores are statistically significant for lags of 2- and 4-days, but not lags of 1-, 3-, 5-, or 6-days.
- Effects were greater in the morning (when ambient ozone levels are lowest) than in the evening (when ambient ozone levels are highest). The authors' explanation for this anomaly is redolent of HARKing.

- ❖ Relative risk declined to nonsignificance when other air pollutants were controlled for.

Models that are biologically implausible, either individually or in combination, cannot satisfy the information quality standard of objectivity. They are scientifically inaccurate.

e) Disseminating results from models known to yield risk estimates that are upwardly biased and more uncertain, such as Generalized Additive Models conducted with insufficient convergence criteria

For years EPA has relied on studies utilizing Generalized Additive Models (GAMs) and software (S-PLUS) that yielded estimates of health effects from air pollution that were upwardly biased and excessively certain (Health Effects Institute 2003). Correcting for these biases reduced the mean estimate of mortality attributed to PM_{10} from 0.41% per 10 $\mu g/m^3$ to 0.27% per 10 $\mu g/m^3$ (using GAM with stricter convergence criteria) and 0.21% per 10 $\mu g/m^3$ (using Generalized Lineal Models [GLM] with natural cubic splines), a reduction in estimated mean risk of about half. Given the uncertainties involved, doubts of course arise concerning the validity of the reported precision even in these re-estimates.

To the extent that risk estimates from time series models are inherently biased, they do not strictly adhere to the information quality standard of objectivity. To comply, EPA must always prefer models plausibly shown to have reduced bias, and not erect barriers to their use just because they have lower risk estimates. To the extent that fully objective estimates cannot be derived, EPA is obliged under the objectivity standard to present these facts in an accurate, clear, complete, and unbiased manner. Indicative of this problem, the convergence issue is documented in the Criteria Document, but ignored in both the Staff Paper and the NPRM. This is fundamentally at odds with the principle of presentational objectivity.

Defects in presentational objectivity also go beyond the omission of relevant information. There is no way to ensure that whichever model is selected is not upwardly biased due to insufficient control for confounding (Moolgavkar 2007, 9). Nevertheless, EPA persists in looking at multiple (upwardly biased) time series models and concluding that, because their results are consistent, the association they describe must also be robust (U.S. Environmental Protection Agency 2007b, 37836). This confuses reliability (i.e., the capacity to generate similar outputs) with validity (i.e., the capacity to generate an unbiased estimate).

Biostatisticians who utilize GAMs to estimate air pollution effects warn that their properties are still too uncertain for widespread use by epidemiologists lacking expertise in both the underlying mathematics and statistical programming (Lumley and Sheppard 2003). The effects to be teased out of the data are smaller than uncertainties in model specification and known confounders. For example, seasonal effects on health endpoints are understood to be much larger than the effects of the air pollutants themselves. Presentational objectivity would be greatly enhanced if EPA placed in context the magnitude of the effects from ozone with the magnitudes of effects resulting from differences and changes in confounders, all objectively estimated.

6. Assumption of causality

Discerning which epidemiological associations are causal is inherently difficult and fraught with uncertainty. Criteria have been proposed as guides, most notably by (Hill 1965): (1) strength, (2) consistency, (3) specificity, (4) temporality, (5) biological gradient (i.e., dose-response), (6) biological plausibility, (7) coherence, (8) experiment, and (9) analogy. EPA asserts that the epidemiological evidence satisfies these criteria, but it does not reveal how it interpreted the Hill (or similar) criteria. Transparency is crucial here, because there may be no greater scientific issue than causality, and without transparency, reproducibility is impossible. Without transparency, it appears that EPA has used policy tools to "deem" scientific statements to be "true."

With regard to tests of statistical significance, Hill advised against either ignoring them or becoming their slaves. Writing in the United Kingdom in the mid-1960s, Hill said:

> Fortunately I believe we have not yet gone so far as our friends in the USA where, I am told, some editors of journals will return an article because tests of significance have not been applied. Yet there are innumerable situations in which they are totally unnecessary – because the difference is grotesquely obvious, because it is negligible, or because, whether it be formally significant or not, it is too small to be of any practical importance. What is worse the glitter of the t table diverts attention from the inadequacies of the fare.

The Hill criteria do not interpret or apply themselves; judgment is clearly required. However, there is a fundamental difference between scientific judgment (judgment founded on scientific data, methods and reasoning used to infer what *is* true) and policy judgment (judgment based on values, preferences and opinions about what *ought to be* true). At the intersection between science and policy, it may be impossible to discern scientific and policy judgment.

However, EPA's approach to judging causality is unambiguously and transparently policy directed.

Figure B: Signal Strength and Statistical Significance in EPA Ozone Risk Assessment

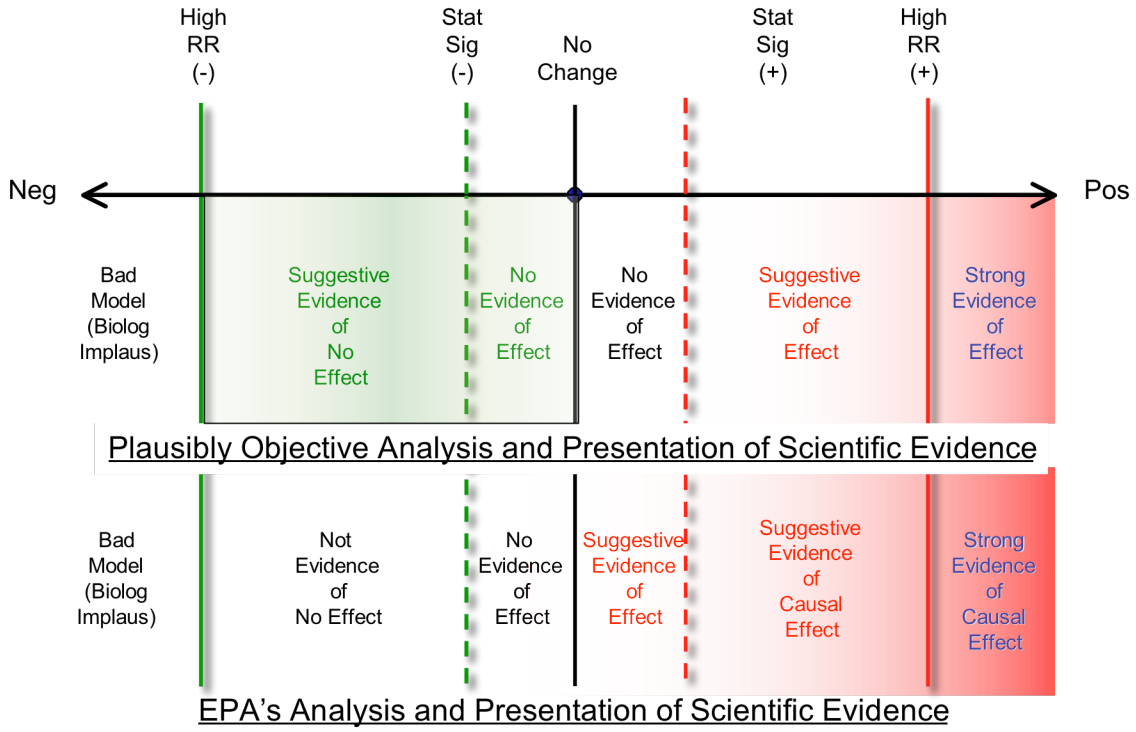


Figure B illustrates how a plausibly objective analysis would incorporate negative, nonpositive, and positive results from a given study, and compares that schema with the method EPA appears to have followed. Vertical lines represent arbitrary thresholds for statistical significance, and higher thresholds (in absolute value) for associations that are strong enough to provide "high" estimates of relative risk. The figure is drawn symmetrically with respect to positive and negative values, and it is assumed that statistical significance is obtained at relative risk levels lower than what is required for biological significance. Highly negative relative risk values are biologically implausible, so in both schemes such results are interpreted as evidence of a badly specified model.

In the plausibly objective approach, a highly positive relative risk ratio is strong evidence that the association is genuine and not the result of chance. As the relative risk declines, confidence that the association is true also declines. Relative risks that are not significant are interpreted as no evidence of an effect regardless of the sign. Negative relative risk ratios that are statistically significant but not biologically implausible are suggestive evidence of no effect.

EPA's approach to utilizing individual studies is very different. First, negative relative risk ratios are never suggestive of the absence of an effect. Second, positive relative risk ratios that are not statistically significant (and well below biological significance) are considered suggestive evidence of an effect. Statistically significant positive relative risk ratios are interpreted as suggestive evidence of a causal effect, and highly positive relative risk ratios are considered strong evidence of a causal effect.

EPA's approach is generous with respect to interpreting positive associations as meaningful and quick to infer causality. This explains how EPA can collect many studies on ozone, each of which has small relative risks with small effects, and some of which are positive, and from this collection draw a "weight of evidence" conclusion that, when taken as a whole, the literature supports or strongly supports an inference of causality.

D. EPA's Risk Assessment Is Biased as a Matter of Policy

Of course, this is hardly an objective weight-of-evidence procedure. It has its origin in the EPA's policy mission, which Agency staff perceive as protecting public health over all other considerations. In a recent report on Agency risk assessment principles and practices, EPA staff acknowledged that it was their practice to produce purposefully biased risk assessments (U.S. Environmental Protection Agency Office of the Science Advisor 2004). The staff report stated a commitment "*to provide the best possible scientific characterization of risks based on a rigorous analysis of available information and knowledge*" (p. 3, emphasis in original), and endorsed the information quality principle of objectivity (pp. 9-10). But the staff then abandoned those principles by declaring that its practices were biased, and purposefully so:

> EPA's risk assessments are conducted in support of its mission to protect public health and the environment. Given the uncertainty, variability, and data gaps encountered when conducting any risk assessment, a key objective for EPA's risk assessments is that they avoid both *underestimation* of risk and *gross overestimation* of risk (p. 11, emphasis added).

"In other words," the staff continued, "EPA seeks to adequately protect public and environmental health by *ensuring that risk is not likely to be underestimated*" (emphasis in original).

In its ozone risk assessment, EPA has been faithful to the Staff Report. It is highly unlikely that the Agency has underestimated risk, and it is almost certain that it has overstated risk. EPA's ozone risk assessment is fundamentally incompatible with information quality principles that the Agency, and the staff who perform risk assessments, say they agree with and fulfill. Comparing the plain language of information quality guidelines and the EPA Staff Report, it is self-evident that the two authorities cannot be reconciled.

E. <u>EPA Assumes Constant Risk per Unit of Ozone for All Levels of Ozone Exposure</u>

EPA's concentration-response function assumes that the contribution to risk from any increment of exposure is essentially constant irrespective of the baseline level of exposure. Thus, the effect on risk is the same for exposure 10 ppb at a baseline of 80 ppb as it is for a baseline of 20 ppb. This assumption is not supported by scientific evidence and represents an upper-bound estimate of risks at lower concentrations.

Nevertheless, the fact that something is biologically implausible is not proof that it is false. EPA thus could disseminate biologically implausible statements and assumptions if the Agency defended them with persuasive scientific evidence that they are, in fact, likely to be true. However, EPA does not cite any scientific evidence supporting that proposition.

Information quality guidelines do not restrict the Administrator's discretion to make his policy determination based on statements that he transparently acknowledges are biologically implausible and without scientific foundation. The information quality guidelines are triggered, however, if EPA asserts or implies that biologically implausible scientific assumptions are in fact true, or if EPA fails to properly characterize such statements as biologically implausible, without scientific foundation, and almost certainly false.

Lack of scientific certainty about any scientific proposition is not a justifiable reason for supplanting what is scientifically known in favor of default values. Applicable information quality guidelines require that scientific and statistical statements and claims reflect the best available evidence, and that they not include purposeful or known biases – particularly those which might be motivated by policy considerations. Default values that conflict with what is known scientifically cannot be compatible with the information quality principles of objectivity. Policy considerations, such as the degree of risk

aversion to incorporate into the NAAQS *because the science is uncertain*, must be characterized and described distinct from the science that EPA believes is uncertain. That is, the Administrator may choose to be as risk averse toward scientific uncertainty as he believes is permissible under the law, but EPA cannot misrepresent the type, degree or magnitude of scientific uncertainty in its communications to the Administrator without violating applicable information quality principles.

F. EPA Attributes to Ozone Risks That It Has Previously Attributed to other Pollutants

EPA uses single-pollutant models to estimate the association between ozone and health effects. For example, the same studies that EPA uses report estimates for multi-pollutants models that include $PM_{2.5}$. Assuming that these models are unbiased, the signal from $PM_{2.5}$ in these models is greater than the signal from ozone. But EPA ignores the contribution to risk from $PM_{2.5}$ and relies on estimates that exclude $PM_{2.5}$. As long as $PM_{2.5}$ and ozone are correlated, the exclusion of $PM_{2.5}$ will have the effect of attributing to ozone risk that is better attributed to $PM_{2.5}$.

An area of concern is a lack for transparency concerning the nature of control for particulate matter, which of course is not the same irrespective of particle size. When EPA says that its estimates of risk from ozone (or a peer reviewed study on which it relies) has "controlled for PM," it is not clear whether that control addressed only PM_{10} or $PM_{2.5}$ as well. $PM_{2.5}$ is correlated in space and time with ozone; PM_{10} is not. Thus, controlling for PM_{10} is, in effect, functionally equivalent to no control at all.

Estimating ozone risks while failing to control for $PM_{2.5}$ results in upwardly biased estimates. In some cases, it attributes to ozone risks that the Agency previously attributed to $PM_{2.5}$.

This is an obvious violation of the substantive objectivity standard. If it is true that the studies EPA relies on constitute the best scientific evidence, then EPA is required to disseminate results from models that control for confounding effects such as the contribution to risk from $PM_{2.5}$. Single-pollutants models yield unambiguously biased estimates of relative risk. (EPA can satisfy the presentational objectivity standard by disseminating estimates from both models; describing results from the single-pollutant model as biased; simultaneously reporting the magnitude of this bias; and describing and quantifying this bias in all subsequent estimates or calculations that rely on biased inputs. But EPA violates the presentational objectivity standard if, at any point onward, it fails to describe and quantify this bias.)

IV. Information Quality Errors in the Assessment of Human Health Risk

In this section, we summarize several general aspects of EPA's ozone risk assessment that conflict with the information quality principle of objectivity.

A. EPA Treats Transient and Reversible Effects as Adverse

Researchers have long struggled to define *adversity* based on scientific criteria, and there is no consensus concerning where to draw the line distinguishing adverse from non-adverse effects, such as effects that are transient and reversible over short periods of time. This problem continues to grow. EPA increasingly treats as adverse mere exposure. When exposure cannot be detected, EPA increasingly looks for biomarkers of exposure -- irrespective of whether they are associated with symptoms or signs; their selectivity with respect to the hazard of concern; or the capacity to detect them without complex analytic techniques. The Agency has issued policy guidance giving Agency staff wide discretion to decide what is implicitly "adverse," such as its definition of the term "key event" (see, e.g., U.S. Environmental Protection Agency 2005a, 1-10, n. 1). Scientists might even be able to agree as to which phenomena are "key events" and which are not, but it is crucial to remember that this scientific task is a very limited one, consisting only of the assignment of phenomena into categories *defined by policy*. Nevertheless, information quality guidelines require that the Agency be transparent in its scientific description of these phenomena and the available data. This is especially so when phenomena are deemed to be "adverse" based on policy criteria because the public rightly infers that something described as "risky" must be "bad" on some objective scale.

As one example, it is easy to see how this problem infects EPA's descriptions of effects to asthmatics. Transient and reversible effects, such as increases in chest tightness or wheezing, have a dozen reported triggers (Sarafino et al. 2001), including laughter. How similar in magnitude are the effects of ozone to the effects caused by exposure to humor? (Kimata 2004; Lianges et al. 2003; Lianges et al. 2004)

B. EPA Uses Important Scientific Terms and Language in Policy-directed Ways

Statements described as having "medium confidence" or being "possible" or "reasonable" are difficult to either support or refute, and so they have little or no utility for decision-making. For that reason, there is always pressure to say more – that a statement is "likely" to be true or an event is "likely" to happen" -- to convey some sense that the balance is tipped by new information. There is a

tension at work: statements that tip the balance a lot do convey more information, but they are harder to justify and easier to refute. Our model of EPA's approach to causality presented in Figure B in Section III.C.6 shows one "workaround": a large number of studies can be assembled, each of which has weak or limited evidence, but in combination they can be transformed into predictions that are "likely" about which EPA staff is "confident." The words "likely" or "unlikely" appear 144 times in volume 1 of the Criteria Document, 177 times in the Staff Paper, and 134 times in the NPRM. In none of these documents are the words themselves defined.[19]

Where EPA uses probabilistic terms to describe statements of fact or knowledge, information quality principles require that the Agency show that its probabilistic terms are founded on science and comport with how decision makers and the public understand these terms. It is not enough merely to show that, once defined, scientists can consistently apply them. The terms and categories themselves must be consistent with scientific principles, objective in design, and have utility for the purpose to which they are used. Thus, it is a violation of the information quality standard of objectivity to use terms such as "likely" or "probably" in ways that conflict with their actual use in an appropriate context or without clear definition.

EPA needs to establish clear rules and procedures for how probabilistic language will be used in risk assessments and similar documents prepared to guide decision-making. Prescriptive consistency in language reduces uncertainty about how language is used in documents prepared by multiple authors or by agency committee and work group process, such as the documents subject to this RFC. Four principles should guide the development of these rules and procedures.

First, because probabilistic statements are semi-quantitative, when scientists, decision-makers and the public use the same words, they should mean roughly the same thing. Without guidance, potential interpretative heterogeneity is unbounded. By assigning quantitative values to statements about likelihood, interpretative heterogeneity should be drastically reduced.

Second, the values assigned by EPA to likelihood statements and probability descriptors must be consistent with both intuition and scientific

---

[19] The terms "robust" and modified variant (e.g., "fairly robust," "generally robust," "statistically robust") are never defined, but they appear 54 times to describe associations in volume 1 of the Criteria Document, 48 times in the Staff Paper, and 28 times in the NPRM.

research about such terms. That is, EPA cannot simply invent a rule that enables it to transform objectively weak scientific information into statements asserting high levels of confidence or likelihood. EPA must look at relevant research literature on the meaning of ambiguous terms and utilize this research in crafting the scales.

Third, the values EPA assigns to probabilistic language must be transparent, and to a great degree, also reproducible with an acceptable degree of imprecision or error (Office of Management and Budget 2002, Sections V.5.a ["transparency"] and V.10 ["reproducibility"]). To adhere to applicable information quality standards, EPA must at a minimum make transparent what it means when it uses likelihood statements and probability descriptors. Further, it must re-examine its use of these statements and descriptors to ensure that the Agency is applying them consistently throughout.

Finally, EPA must be forthcoming with full and complete documentation of what it proposes, and subject its work to pre-dissemination review (such as peer review by qualified psychologists). Applications of this guidance must be challengeable under the Agency's error correction procedures.

C. EPA Confuses Variability and Uncertainty

The distinction between variability and uncertainty is well established and understood in the risk assessment field. Variability is the range of effects estimated or observed in a sample, subpopulation or population. Uncertainty is the domain of what is scientifically unknown, and in some cases, scientifically unknowable. It is crucial to keep these concepts distinct because uncertainty is reducible in part through scientific investigation, variability is not.

EPA's documents consistently confuse these terms – or, more specifically, they frequently use *uncertainty* to refer to both *uncertainty* and *variability,* particularly the documents (and sections of documents) most likely to be read by policy officials. Knowledgeable officials, who understand the difference between the two terms, are misled to believe that it is uncertainty that is being described when in fact it is variability.

This problem extends beyond estimates of the magnitude of risk. Assuming that an unbiased estimator is used to describe risk, confidence intervals capture only variability and, if the data are from a sample rather than a census, the portion of uncertainty resulting from sampling error. Variability can be large, but among sources of uncertainty, sampling error tends to be small. This is certainly true with respect to the measurement and model uncertainties that underline EPA's risk estimates because the strength of the signal is weak even if it is causal. It has long been advocated that risk assessment do more than

acknowledge uncertainty, and that it be made explicit and transparent (Morgan et al. 1990).

Because variability is often mischaracterized as uncertainty, EPA's analysis encourages readers (including the Administrator) to be considerably more confident in the risk estimates than is warranted. This is exacerbated by EPA's emphasis on populations of concern, which by definition are small and in some cases remote percentiles of the population. Estimating the tails of any distribution is much more difficult than estimating the mean, and estimates of the tails are much less certain.

The practical effect of this confusion is that the Administrator is misled to believe that EPA's risk estimates are much more certain than they actually are, and that what is described in EPA's supporting documents is anything other than population variability and sampling error. It is technically incorrect and fundamentally misleading to provide the Administrator information about sample variability but describe that information as characterizing the bounds of scientific uncertainty.[20]

In short, EPA has presented the Administrator data and analyses that led him to be much more confident than is scientifically justified that ozone exposure below the current NAAQS poses actual human health risks. EPA's risk estimates capture only statistical variability for the selected models, not scientific uncertainty about their validity and reliability as estimates of human health risk. Information about variability, which is small relative to the magnitude of variability and uncertainty combined, has no utility *to the Administrator* unless it is placed in proper context with information about uncertainty. The Administrator's statutory assignment is to decide *whether there is a sufficient evidence* that exposures below the current standard poses a sufficient *incremental* risk to warrant revising the NAAQS downward.[21]

---

[20] "A distinction between uncertainty (i.e., degree of potential error) and interindividual variability (i.e., population heterogeneity) is generally required if the resulting quantitative risk characterization is to be optimally useful for regulatory purposes, particularly insofar as risk characterizations are treated quantitatively" (National Research Council 1994).

[21] "[D]eciding whether to be conservative in the face of variability rests on a policy judgment about how far to extend the attempt to provide safety" (National Research Council 1994).

D. EPA Does Not Disclose a Credible Analysis of Uncertainty

Since at least 1994, EPA has been advised by the National Academy of Sciences to perform quantitative uncertainty analysis in its most important risk assessments (National Research Council 1994). The Agency was criticized then for relying on single point estimates, especially when those estimates were described as "plausible upper bounds" without quantitative analysis of the extent to which they were upwardly biased. Such risk estimates were criticized as misleading or untrue. Uncertainties needed to be explicit and presented "as accurately and fully as is feasible and needed for risk management decision-making" (p. 185). Thirteen years later, in a risk assessment supporting perhaps the Agency's most important regulatory initiative, EPA continues to rely on plausible upper-bound point estimates and declines to conduct or disseminate a formal uncertainty analysis.[22]

In its 2002 report to EPA on health benefits estimation, a committee of the National Academy of Sciences commented on the Agency's practices for risk assessment and health benefits analysis (National Research Council 2002). The committee reached several conclusions that also apply to the risk assessment documents prepared in support of the NPRM. Quoting verbatim from the NAS report:

> In its primary analyses of health benefits, EPA reports the uncertainty as a probability distribution. Only one source of uncertainty, the random sampling variability of the estimated concentration-response function, is given with an emphasis on the mean of the probability distribution. The absence of other sources of uncertainty makes the results of the primary analyses appear more certain than they are.
>
> To address other sources of uncertainty, EPA uses ancillary analyses, such as alternative and supplementary calculations and sensitivity analyses. With the exception of concentration-response function estimates, these ancillary analyses usually examine only one source of uncertainty at a time and only for the impact on the mean value of the probability distribution from the primary analysis. As a consequence, though laudable steps in the right direction, these ancillary analyses do not adequately convey the relative or aggregate degree of uncertainty created by the sources of uncertainty addressed in the analyses, nor, of course, do they depict uncertainty from other sources.

---

[22] The only uncertainty analysis provided is an unpublished staff memorandum (U.S. Environmental Protection Agency 2007a).

EPA appears to have made little progress even acknowledging these scientific problems. The documents subject to this RFC continue the same practices that the National Academy criticized five years ago, three years before the Agency published its research plan for the ozone NAAQS review (U.S. Environmental Protection Agency 2005b).

In 1997 when the ozone NAAQS was last revised, the 2002 Academy report was off in the future, and it might be argued that the 1994 Academy report was too new, too novel, or too difficult.[23] That defense is no longer credible in 2007. EPA has adopted few of the Academy's recommendations, and in the documents subject to this RFC the Agency does not explain in these documents why the Academy's advice is scientifically flawed. Reports from the National Academy have a rebuttable presumption of objectivity when the subject being addressed is strictly scientific. Where EPA chooses not to adopt such technical recommendations, information quality principles strongly suggest that it has a duty to at least explain why, and probably to rebut the presumption that they are, in fact, objective.

E. EPA's Particular Use of Default Values Violates Information Quality Principles

The use of "inference guidelines" (National Research Council 1983) and "default options" (National Research Council 1994) has a long and checkered history. Regardless of the terminology used, it refers to a scientific concept, construct or fact which is uncertain, unknown or unknowable, and for which judgment of some sort is required to choose "among several scientifically plausible options" (National Research Council 1983). Ten years later it became clear that there was an irreconcilable difference between those who thought default options *ought* to err on the side of overestimating risk (National Research Council 1994) and those who did not (National Research Council 1994). The committee as a whole nevertheless reached agreement that EPA needed to "provide justification for its current defaults and set up a procedure such as that proposed in the report that permits departures from the default options"

[23] EPA also has had five years' experience integrating the information quality principle of objectivity into its risk assessment practices. Coincidentally, five years is the statutorily mandated interval by which EPA is required to review its National Ambient Air Quality Standards under Section 109 of the Clean Air Act.

(National Research Council 1994). Thirteen years later, EPA has not established that procedure,[24] or any other one.

More importantly, the federal Information Quality Act and its implementing guidance have superseded these debates. Information of a scientific nature now must be objective, both in substance and presentation. Default options consist of scientific information, and thus they are fully subject to these objectivity requirements. Whether to set standards that are health protective (i.e., aim to protect a relatively high percentile of the affected population), and if so, how protective (i.e., which percentile to aim to protect) are policy decisions solely within the discretion of the authorized decision maker – in this case, the Administrator of EPA. The Administrator's obligation is to be transparent and accountable with respect to these judgments, but he cannot do so if the scientific information on which he must depend is infected with default options that implicitly and surreptitiously contain policy judgments that he alone is authorized to make. In the words of Justice Breyer in the *American Trucking* case:

> The statute's words … authorize the Administrator to consider the severity of a pollutant's potential adverse health effects, the number of those likely to be affected, the distribution of the adverse effects, and the uncertainties surrounding each estimate. They permit the Administrator to take account of comparative health consequences. They allow her to take account of context when determining the acceptability of small risks to health. And they give her considerable discretion when she does so (Whitman v. American Truckling Ass'ns, Inc, 531 U.S. 457, 495 (J. Breyer, concurring, internal citations omitted)).

Exercising this discretion requires *accurate, reliable, and unbiased* information about "the severity of [ozone's] adverse health effects, the number of those likely to be affected, the distribution of the adverse effects, and the uncertainties surrounding each estimate." This information must be presented in an *accurate, clear, complete, and unbiased manner*.

The documents subject to this RFC systematically incorporate default options that fail the substantive objectivity test. Moreover, the degree to which policy judgments that belong solely to *the Administrator's* discretion are

---

> "In many cases, the regulated parties may be willing to fund research that will enable health-protective default options in risk assessment to be replaced by more complex and less conservative alternatives" (National Research Council 1994).

subordinated to or restricted by the public policy preferences of *Agency staff* is nowhere made transparent. For that reason, these documents also violate the presentational objectivity test. The Administrator cannot reasonably be expected to discern, from the documents he has been provided, an unbiased estimate of human health risks posed by ozone exposure below the current NAAQS. These documents thus do not satisfy the utility standard of information quality. The Administrator cannot responsibly exercise the full breadth of his statutory authority; he can only exercise that portion of his statutory discretion left over after career EPA staff have misused risk assessment to severely narrow his choices in ways that appeal to them.

F. EPA Assumes Confidence Intervals Adequately Describe Variability and Uncertainty

EPA begins with relative risk estimates from specific epidemiological studies. These estimates are reported as either statistically significant or nonsignificant based on reported confidence intervals. These confidence intervals assume that errors are independently and identically distributed, but generally they are not even when the samples are random, and that means they are biased. Even if the errors are independently and identically distributed, there is measurement error in the independent variables. For example, even if ambient monitors are unbiased predictors of personal exposure – i.e., they are equally likely to over- or under-predict actual personal exposure -- they are less precise and thus less certain. True confidence intervals are correspondingly much wider.

This problem is exacerbated by the tendency of epidemiologists to use convenience samples and other non-random research designs, then use statistical methods which to be valid require that the samples be random. Reported confidence intervals thus represent the "best case" condition in which a non-random sample happens to be equivalent to a random sample, just by chance, and actual confidence intervals are (perhaps much) wider.

The odds against a non-random sample being equivalent to a random sample by chance event are astronomical. Yet it is rare for any biomedical researcher to explore the extent to which the properties of random sampling do not apply to her data. Given the demands of journals (and sometimes sponsors) for statistically significant results, there may be no incentive to even consider undertaking that kind of evaluation. In any case, when epidemiologists report relative risk estimates with lower 95th percent confidence intervals that are close to 1.0, it is very likely that departures from randomization in their data are sufficient to eliminate statistical significance. Presentational objectivity demands at least a transparent acknowledgement of this problem with the advice to interpret such results with caution.

G. EPA Assumes that Ambient Monitors Provide Unbiased Estimates of Personal Exposure

The scientific literature that EPA considered shows convincingly that personal exposures to ozone are lower than what is recorded by ambient monitoring. This is true because ambient monitors measure outdoor exposure and people spend most of their time indoors, and the level of ozone indoors is a small fraction of outdoor levels. By assuming outdoor exposure, EPA exaggerates exposure by a factor of three to ten for all but a subset of persons who work outdoors throughout the ozone season.

This assumption violates the objectivity requirement of information quality because it imparts purposeful and avoidable bias to the risk estimate. Presentational bias could have been avoided if EPA had acknowledged both the existence and the magnitude of this bias, and explained its implications for the consideration of regulatory alternatives, but the Agency did not do so. The result is that EPA denied the Administrator crucial information he needs to exercise his statutory discretion in standard-setting informed by accurate, clear and unbiased scientific information. Risk assessments that include purposefully biased estimates of exposure have no utility for risk management decision-making.

H. EPA Assumes that Associations Observed in Time Series Studies Are Significant but the Absence of Associations in Long-Term Cohort Studies Is Not

EPA acknowledges that epidemiological evidence from long-term cohort studies does not show consistent associations between ozone and premature mortality (U.S. Environmental Protection Agency 2006a); (U.S. Environmental Protection Agency 2007f). However, EPA also says that it believes the time series studies that *do* show positive associations, and the Agency interprets this as supporting evidence that ozone causes premature mortality. In short, EPA says that ozone causes premature mortality in the short-term that cannot be observed over the long-term. No logical argument is advanced to support this claim. EPA arrives at this conclusion by leap of faith, not scientific inference.

I. EPA Assumes No Uncertainty about Causality

In human health risk assessment, ascertaining causal relationships is one of the most difficult tasks that must be undertaken. It is known throughout the scientific community that association is not causation. Rarely, if ever, can a causal relationship be proved between a particular exposure and an effect even when a known causal relationship exists between a pollutant and a health effect. For example, lead is known to cause adverse neurological effects that result in

reduced intellect, as measured by IQ. But it cannot be shown that a specific individual's reduction in IQ was caused by a specific increment of lead exposure.

For this reason, it is essential for an estimate of risk given a fixed exposure level to be substantively objective, it must fully account for uncertainty about causality. For example, if the best available scientific evidence is that there is exactly a 50% chance that a causal relationship exists for any individual, then that risk should be described as having a normal distribution with mean 0.5 times the (objective) estimate of the unit risk value. The risk distribution can be carried through the entire risk analysis, including the estimates of (a) baseline disease incidence, (b) post-regulation disease incidence, and (c) the valuation of avoided disease incidence. (Valuation should be adjusted to the certainty equivalent to appropriately account for risk aversion in the affected population, and the level of risk aversion rises with the magnitude of risk. Individual willingness to pay will exceed the expected value for certain risks, but it will be less than the expected value for risks that are uncertain.)

EPA's approach to causality is fundamentally inconsistent with these information quality principles. Instead of capturing and propagating uncertainty about causality in its risk estimates, EPA simply assumes it away. All uncertain risks – whether their likelihoods are 90%, 50%, 10% or 1% -- are treated as if their likelihoods were 100%. In short, EPA commits the fundamental scientific error of assuming that association is causation. That this makes EPA' s estimates biased is beyond dispute; the only question is how much.

J. EPA Does Not Explain the Effects of Ozone with Reference to Any Non-Air Pollution Context

Information quality principles require that scientific information be "presented in an accurate, clear, complete, and unbiased manner," and "within a proper context" (Office of Management and Budget 2002, 8459, Section V.3.a). A starting point for context would be to compare the severity of respiratory effects from ozone to effects associated with the most relevant of perhaps a dozen confounders (Sarafino et al. 2001). For example, what is the equivalent difference in allergen exposure, exercise, temperature or humidity that has about the same effect as a 0.010 ppm difference in ozone? Academic efforts notwithstanding, the public is more likely to understand the magnitude and severity of the hypothesized risk if they are described in terms of non-air pollution phenomena that they routinely experience.

K. Double-counting

We sympathize with EPA concerning the difficulty of parsing effects into those associated with air pollution and those that are associated with other

factors; and among air pollutants, effects associated with ozone from effects associated with $PM_{2.5}$ and NOx. Still, there is a fundamental risk assessment principle that must be followed: double counting is simply not acceptable. The same unit of risk cannot be attributed now to ozone if it has been previously attributed to PM.[25] Similarly, risks attributed to ozone now cannot in the future be attributed to PM.

In many respects, EPA has practiced an envelope-style method that maximizes what conceivably could be attributed to ozone. Because this approach is so extreme, it is virtually certain that health risk now attributed to ozone have been accounted for previously by EPA -- most prominently, premature mortality risks in its analysis of fine PM.

Each unit of air pollution risk (like risks from other factors) must be fully counted once, but only once. EPA's risk assessment does not perform this allocation. Because of the close environmental and programmatic relationship between ozone and fine PM, this analysis is essential.

L. EPA's Alternative Risk Estimates

Substantive objectivity requires that information be presented in an "accurate, reliable, and unbiased." EPA does not adhere to this requirement in the reporting of alternative risk estimates. EPA characterizes its risk estimates as "primary" or "secondary." This language implies that one set of estimates ("primary") have a stronger scientific foundation and are more likely to be correct than the other set of estimates ("secondary"). However, nowhere does the Agency use logic, science or statistical method to show that this distinction is grounded in either science or probability. Rather, EPA's "primary" risk estimates are those that tend to support a policy preference for a more stringent NAAQS, and EPA's "secondary" risk estimates are those that do not. This distinction is purely nonscientific. It cannot be characterized as "accurate, clear, complete, and unbiased." Accuracy and clarity require that EPA avoid language suggesting any scientific or statistical foundation for claims that cannot be supported with science or statistics. As an organization, of course, EPA is entitled to prefer more stringent air pollution standards. Nevertheless, information quality guidelines prohibit it from mischaracterizing these policy preferences as scientific, or informed by science, when they are not.

[25] Unless, of course, EPA transparently admits that its estimates for PM were wrong and it corrects previous errors before being petitioned to do so.

1. Defects in Substantive Objectivity

To adhere to information quality requirements, EPA must accurately and clearly describe what distinguishes these two classes of risk estimates. It cannot use words that imply an ordinal ranking informed by the application of scientific criteria. This is particularly important because the risk estimates EPA prefers for policy reasons are less likely to be true than the risk estimates that it dislikes.

2. Defects in Presentational Objectivity

Presentational objectivity requires that information be presented in an "accurate, clear, complete, and unbiased" manner. EPA also does not adhere to this requirement in the reporting of alternative risk estimates. Only the Agency's "primary" risk estimates (i.e., those which tend to support lowering the NAAQS) are carried forward into any document decision-makers might be expected to read. Risk estimates that do not tend to support lowering the NAAQS (i.e., EPA's "secondary" risk estimates) are relegated to background documents. Failing to carry forward all risk estimates of similar likelihood – and in this case, giving greater presentational attention to risk estimates of lower likelihood – cannot be reconciled with the presentational objectivity standard.

V. Information Quality Errors in the Consideration of Reports from CASAC

By law, the Clean Air Scientific Advisory Committee (CASAC) is charged with performing both a scientific review and policy advice function.[26] This

[26] Clean Air Act, Section 109(d)(2):

> (A) The Administrator shall appoint an independent scientific review committee composed of seven members including at least one member of the National Academy of Sciences, one physician, and one person representing State air pollution control agencies.
>
> (B) Not later than January 1, 1980, and at five-year intervals thereafter, the committee referred to in subparagraph (A) shall complete a review of the criteria published under section 108 and the national primary and secondary ambient air quality standards promulgated under this section and shall recommend to the Administrator any new national ambient air quality standards and revisions of existing criteria and standards as may be appropriate under section 108 and subsection (b) of this section.

means EPA must be extraordinarily careful in how it listens to CASAC to ensure that it clearly distinguishes scientific insight from policy prescription. Because it is an independent body outside of the Agency's control, CASAC is exempt from federal information quality guidelines. However, EPA is not exempt when it disseminates or uses information provided by CASAC. EPA cannot simply cite CASAC as a scientific authority without regard for whether their content adheres to applicable information quality standards. In any case where EPA disseminates covered information obtained from CASAC in a manner that a reasonable person would construe as Agency agreement, EPA must ensure that the information satisfies applicable information quality standards as if the Agency itself had produced or sponsored the information.[27]

Policy advice provided by CASAC members generally is not subject to information quality principles because it lies outside the boundaries of the definition of *information*. However, EPA must be careful to correctly characterize policy advice it receives from CASAC as policy advice and not, explicitly or implicitly, describe it as science.[28] By doing so, EPA voids the "opinion exemption" in the definition and subjects policy advice to the same level of

---

> (C) Such committee shall also (i) advise the Administrator of areas in which additional knowledge is required to appraise the adequacy and basis of existing, new, or revised national ambient air quality standards, (ii) describe the research efforts necessary to provide the required information, (iii) advise the Administrator on the relative contribution to air pollution concentrations of natural as well as anthropogenic activity, and (iv) advise the Administrator of any adverse public health, welfare, social, economic, or energy effects which may result from various strategies for attainment and maintenance of such national ambient air quality standards.

[27] The information quality definition of *information* "does not include opinions, where the agency's presentation makes it clear that what is being offered is someone's opinion rather than fact or the agency's views (Office of Management and Budget 2002). However, once an agency adopts a third party's scientific statements as its own, then information quality principles apply. "Subsequent agency dissemination of [third-party scientific] information requires that the information adhere to the agency's information quality guidelines (p. 8454, col. 2).

[28] This is true even if CASAC describes its input as scientific when it is in fact policy advice.

scrutiny to which scientific information is subject. Fortunately, this problem is easy to solve, simply by properly distinguishing policy matters from science.

A. CASAC's Scientific Charge

CASAC's primary scientific responsibility is to perform a scientific peer review of EPA's various secondary risk assessment documents, including the Criteria Document and the Staff Paper. CASAC may, and perhaps ought, but is not required to, review the underlying studies cited and summarized in these secondary documents. CASAC is directed to "complete a review of the criteria published under section 108" (§109(d)(2)(B)), which requires that air pollution criteria "accurately reflect the latest scientific knowledge useful in indicating the kind and extent of all identifiable effects on public health or welfare which may be expected from the presence of such pollutant in the ambient air, in varying quantities" (§108(a)(2)). In short, even though Clean Air Act § 109 preceded the Information Quality Act, CASAC's primary duty is to ensure that EPA's risk assessment is accurate, clear and unbiased. It cannot, without violating its charge, ratify a risk assessment that is inaccurate, incomplete, or fails to include the latest scientific knowledge.[29]

It is important to keep in mind that CASAC is not charged with performing a de novo review and synthesis of the science. That is EPA's job, which it discharges through the preparation of the Criteria Document. In principle, the Criteria Document should be completely free of policy considerations, so there is no reason why CASAC, in its review, should ever stray into policy matters.

CASAC's review of the EPA Staff Paper is necessarily different, for the Staff Paper is by design a complex mix of science and policy recommendations from Agency staff. In principle, the design of the Staff Paper should make it relatively easy for CASAC to maintain a clear distinction between its scientific review and policy advocacy roles.[30] CASAC does not seem to have adhered to

---

[29] § 109(d)(2)(C) gives CASAC an important secondary scientific charge related to research needs ("areas in which additional knowledge is required"), disaggregate natural from anthropogenic contributions to ambient air pollution, and the quantification of substitution risks ("any adverse public health, welfare, social, economic, or energy effects which may result from various strategies for attainment and maintenance").

[30] Chapters 2, 4, and 5 should be strictly scientific. Chapters 3, 6, 7, and 8 are a blend of science and policy (U.S. Environmental Protection Agency 2007f).

that principle; it is difficult to discern where it is commenting on science and opining about policy. To take just one obvious example, the list of bullets in the transmittal letter to the Administrator contains both scientific comments and policy advice, often within the same bullet.[31]

B. CASAC's Policy Advice Charge

Second, as suggested above, CASAC's members also are invited to provide the Administrator with their opinions regarding how he ought to exercise his statutory discretion in revising or retaining the NAAQS. Because their principal charge is scientific, the public might reasonably expect CASAC members to limit their advice to matters of a strictly scientific nature, as befitting their technical expertise. However, the law does not limit CASAC to advising on matters of science, nor does it constrain them from providing pure policy advice reflecting their personal values and preferences.

The law invites CASAC to provide policy advice several ways. First, it specifies that one member of the committee must "represent[] State air pollution control agencies" (§109(d)(2)(A)). Like EPA, these agencies are regulatory rather than scientific in nature, function, or organization, and they are populated with personnel who quite reasonably share their agency's (and EPA's) air pollution control mission. Furthermore, the act of *representation* is inherently a stakeholder role. When a person "representing" State air pollution control agencies gives advice, it is presumed that this advice will favor intensifying the stringency of

[31] See (Clean Air Scientific Advisory Committee 2006a, 2-3). Like peer reviewers generally, CASAC members are susceptible to the temptation to assume that their own research is most relevant to the question at hand, irrespective of when it was performed. Several CASAC members are well represented in the literature synthesized by EPA, which raises questions about whether they are being asked to indirectly review their own work. (This practice is expressly permitted under the National Academy of Sciences' conflict of interest rules [(The National Academies 2003)], but with an important limitation ["[A]n individual should not serve as a member of a committee with respect to an activity in which a critical review and evaluation of the individual's own work"]).

In this case, the transmittal letter cites for special emphasis six peer reviewed papers authored or co-authored by CASAC members Drs. Morton Lippman and Frank Speizer, all published between 1988 and 1993. The relevance of these studies to CASAC's charge to focus on *new* research is, at best, unclear.

federal air pollution standards. It would be newsworthy only if this person recommended *against* more stringent federal standards.

Second, CASAC members are asked "recommend to the Administrator any new national ambient air quality standards and revisions of existing criteria and standards as may be appropriate"(§ 109(d)(2)(B)). In short, they are invited to speculate as to how they think they would exercise the Administrator's statutory discretion if they were standing in his shoes. Despite the fact that CASAC members all have scientific training and have distinguished themselves in one or more scientific fields, there is nothing scientific about this assignment.

C. EPA Does Not Adequately Distinguish Between Scientific Insight and Policy Advice It Received from CASAC

EPA's Notice of Proposed Rulemaking (U.S. Environmental Protection Agency 2007b) contains numerous subsections in which the input it received from CASAC is summarized. In some places, this input is clearly described as scientific information, subject to applicable information quality guidelines. But in most instances, EPA does not carefully distinguish CASAC's scientific review from its policy advice. This is entirely understandable insofar as CASAC itself did not makes these distinctions clear. However, adherence to information quality guidelines is EPA's responsibility and not that of CASAC. Moreover, in its charge to CASAC EPA did not ask the committee to clearly distinguish between its scientific review and its policy recommendations. For example, EPA did not ask CASAC to apply the Agency's information quality guidelines, and CASAC didn't do so.[32] Nor did EPA disclose any pre-dissemination review of the input it received from CASAC (or any other third party) to ensure that applicable information quality requirements were met.

To minimize the number of error correction requests they receive, agencies are required by OMB's government-wide information quality guidelines to establish effective procedures for pre-dissemination review:

> As a matter of good and effective agency information resources management, agencies shall develop a process for reviewing the quality (including the objectivity, utility, and integrity) of information before it is disseminated. Agencies shall treat information quality as integral to every step of an agency's development of information, including creation, collection, maintenance, and dissemination. This process shall enable the

[32] Information quality is not discussed by CASAC (Clean Air Scientific Advisory Committee 2006a, 2006b, 2007).

> agency to substantiate the quality of the information it has disseminated through documentation or other means appropriate to the information (Office of Management and Budget 2002).

In EPA's own guidelines the Agency states that it has in place sufficient pre-dissemination review procedures to ensure that information quality error is rare:

> Each EPA Program Office and Region will incorporate the information quality principles outlined in section 6 of these Guidelines into their existing pre-dissemination review procedures as appropriate. Offices and Regions may develop unique and new procedures, as needed, to provide additional assurance that the information disseminated by or on behalf of their organizations is consistent with these Guidelines. EPA intends to facilitate implementation of consistent cross-Agency pre-dissemination reviews by establishing a model of minimum review standards based on existing policies. Such a model for pre-dissemination review would still provide that responsibility for the reviews remains in the appropriate EPA Office or Region.
>
> For the purposes of the Guidelines, EPA recognizes that pre-dissemination review procedures may include peer reviews and quality reviews that may occur at many steps in development of information, not only at the point immediately prior to the dissemination of the information (U.S. Environmental Protection Agency 2002)

The problem, though, is that none of the documents subject to this RFC contain any text suggesting that pre-dissemination review actually occurred. Information quality and its associated concepts and definitions simply don't appear.[33]

VI. Information Quality Errors in the Rollback Assumption

To estimate the change in the incidence of health effects attributable to reducing the NAAQS, one must answer the following questions:

---

[33] The dearth of pre-dissemination review is particularly notable for the one instance in which information quality principles do appear in the Notice of Proposed Rulemaking and Staff Paper: EPA's summary of public comments saying that EPA had not examined "the evidence for both adverse and beneficial effects [of tropospheric ozone from UV-B shielding] with the same objectivity." See (U.S. Environmental Protection Agency 2007b) and (U.S. Environmental Protection Agency 2007f)

1. *What must a State do to meet the new standard?* Under the proposal, States would have to ensure that their 4th highest 8-hour average ozone level at all monitors within a specific jurisdiction does not exceed the new value, which is proposed to be in the range of 0.070 to 0.075 ppb.
2. *Is a State already in compliance?* If so, the State should be expected to take no action. For these jurisdictions, the new NAAQS would not be expected to achieve any reduction in health risks.
3. *If a State is not in compliance, what would it do to reach compliance?* A reasonable expectation is that States would focus on actions that reduce ozone peaks, and it would prefer those actions that reduce ozone peaks at least cost.

A perfectly designed and implemented least-cost strategy would ensure that the 4th highest 8-hour ozone level is barely below the lower NAAQS. States would focus their energies on reducing peaks, but devote little or no attention to reducing ozone levels on days when it is already well below the standard.

EPA's rollback assumption might be accurate, but it does not adhere to this logical framework. EPA assumes that States will take regulatory actions that reduce ozone by a proportionate amount throughout the distribution of 8-hour measurement periods. That is, EPA assumes States are indifferent between reducing ozone levels on days when it is low and on days when it is high. Moreover, they are indifferent between targeting regulatory restrictions where they help achieve attainment and imposing broad new requirements that reduce ozone across the board, even though broad new restrictions that reduce ozone levels across the board do not necessarily help States reach attainment. EPA's approach also implies that States are indifferent between imposing regulations that minimize the cost achieving attainment, and imposing a very different set of regulations that nearly maximizes cost. (There are even more costly options available to a State than the one implied by EPA's proportional rollback assumption For example, a State could choose to enact regulations so stringent that they eliminate all ozone above the Policy Relevant Background, even if that required regional deindustrialization.)

The States might in fact behave this way, but EPA provides no evidence suggesting that they do. To justify its proportional rollback assumption as substantively objective, EPA must show that a proportional reduction in ozone levels is the best, unbiased estimate of the effect on ozone levels reasonably attributable to the actions States are likely to take to comply with a lower NAAQS. One place EPA could look for supporting data is in the State Implementation Plans submitted in response to the 1997 revision. This would be

an easy task inasmuch as the Agency must review and approve all SIPs to ensure that they are reasonably expected to achieve attainment.

VII. Information Quality Errors in the Description of Policy Relevant Background

Policy Relevant Background (PRB) is defined as "the ozone concentrations that would be observed in the U.S. in the absence of anthropogenic emissions of precursors (e.g., VOC, NO2, and CO) in the U.S., Canada and Mexico" (U.S. Environmental Protection Agency 2007b). Despite the word "policy," in its title, PRB is a strictly scientific concept. That is, nothing in the definition speaks to how much ozone one might prefer or hope for; it simply defines how much ozone there actually would be in the absence of these anthropogenic emissions.

A. EPA's *Definition* of Policy Relevant Background Yields Biased Estimates of Baseline Risk and Risk Reduction

EPA's definition does include an important policy judgment, however. EPA assumes that ozone precursors from anthropogenic sources in Canada and Mexico are subject to control by U.S. air pollution policy and regulation. The purpose of PBR is to establish a floor below which US regulatory actions cannot decline. By failing to include anthropogenic emissions from Canadian and Mexico, EPA implicitly asserts that these sovereign nations are within its effective jurisdiction.

This assertion is false, of course, irrespective of the merits of a US policy that could control Canada and Mexico. As an analytic matter, the exclusion of anthropogenic emissions from Canada and Mexico result in systematically biased estimates of baseline risk and risk reduction:

1. In any case where EPA predictions of ozone reductions go beyond what State regulations can achieve *within State boundaries*, the Agency must incorporate a convincing explanation as to how State implementing regulations work to reduce background emissions from nearby States.
2. In any case where there are nontrivial emissions from Canada or Mexico that EPA believes would be reduced by lowering the NAAQS, the Agency must include a convincing explanation why State regulations would reach beyond US borders. Otherwise, EPA's estimate of baseline risks from US-controllable anthropogenic emissions is upwardly biased.
3. If EPA's estimates of ozone reductions resulting from a more stringent NAAQS include reductions presumed to occur in Canada

or Mexico, then its estimates of health risks avoided also are upwardly biased.

1. Defects in Substantive Objectivity

EPA's definition of PRB is inherently biased because of its treatment of anthropogenic emissions from Canada and Mexico. Unless extraordinary steps are taken to remove this bias, its practical effect is to (1) exaggerate human health risks from US-controllable ozone in the baseline, and (2) exaggerate the magnitude of human health risks avoided by establishing and implementing a more stringent standard. Only US-controllable ozone matters.

[[Therefore, EPA's estimates of the PRB, and the question how subsequent regulatory action will affect Canadian and Mexican emissions, are fully subject to the provisions of the Information Quality Act and its subsequent implementation guidance. Law and guidance require that EPA's estimate of PRB be "accurate, reliable, and unbiased" and presented in an "accurate, clear, complete, and unbiased" manner. Because the form of the standard is an extreme value from the distribution (i.e., the 4th highest 8-hour average), PRB must be estimated as a distribution of 8-hour averages.]]

B. EPA's *Estimates* of Policy Relevant Background Are Almost Certainly Biased

All estimates of PRB are uncertain, and uncertainty is not an information quality defect unless it is presented in an inaccurate, unclear, incomplete, or biased manner. EPA's PRB estimates have information quality defects independent of both their definition and uncertainty.

EPA's background range of background estimates, 15-35 ppb are derived from the GEOS-CHEM model. These figures do not appear to have been validated. To ensure that they satisfy the applicable objectivity standard, EPA must show that they neither under- nor overestimate actual background levels.

1. Defects in Substantive Objectivity

Scientific literature may be excluded if two conditions apply. First, the information must fail to satisfy applicable information quality standards. No information is exempt from these standards, irrespective of its source, and agencies may not discriminate in favor of or against information that meets information quality standards because of its source. Second, it must be inferior in quality to the information that the Agency would disseminate in its place. EPA cannot discard scientific information that, although imperfect on information quality grounds, is superior in scientific quality to a default assumption or an

alternative scientific source. Information quality principles are not a barrier to inhibit the production, dissemination and use of new scientific information. Rather, they are a perpetual incentive for scientific advancement.

EPA's characterization of the PRB appears to exclude scientific information for reasons other than defects in information quality relative to the quality of the information that would be disseminated in its place. In particular, EPA appears not to have considered at least one literature review (Vingarzan 2004) and one primary research study (Ortmans et al. 2006). Both articles indicate that there is significant spatial and seasonal variability, with springtime peak.

EPA's description of the PRB, and the process it used to estimate it, does not meet these standards. EPA used a specific, low-resolution model (GEOS-CHEM) to estimates "monthly background daily diurnal profiles for each of the 12 urban areas for each month of the O3 season using meteorology for the year 2001" (U.S. Environmental Protection Agency 2007b).

EPA has chosen a value for the PRB that does not adhere to these standards. The Agency has replaced the uniform PRB of 40 ppb that it used to support its 1997 NAAQS revision with a temporally and spatially variable range derived from the GEOS-CHEM model.

EPA implies that the selection of the PRB is a matter of policy discretion, but the Agency defines the PRB in scientific terms. EPA has the statutory discretion to decide how much protection from health effects should be provided, but it does not have the authority to alter scientific principles and concepts in the service of these policy objectives.

2. Defects in Presentational Objectivity

EPA's effort to estimate PRB appears to exclude relevant scientific literature. A recent survey of that literature (Vingarzan 2004) reported a range of annual average background concentrations in North America of 20-45 ppb and annual medians in Canada of 23-34 ppb, with these figures expected to rise due to increased long-range transport of Asian air pollution. These averages and medians disguise seasonal cycles with springtime peaks. Annual averages at remote sites around the world were reported.

The choice of the PRB dominates EPA's estimates of ozone health risks, the incidence of health effects in the U.S. population, and the reduction in incidence likely attributable to lowering the NAAQS. However, the Agency does not provide an accessible, accurate and unbiased sensitivity analysis of how the incidence of health effects is affected by the choice of PRB.

This defect is exacerbated by the fact that the Agency's reduction in the PRB is largely hidden from public view. Whereas EPA proposes to reduce the standard for peak ozone levels, virtually all of its estimated reductions in adverse health effects are attributable to having lowered the PRB. Presentational objectivity requires that EPA accurately and clearly state the proportion of reduced health effects properly attributable to each of the following factors: (a) the increase in EPA's estimate of unit health risks; (b) the proposed reduction in the allowable peak ozone level; (c) the rollback procedure; and (d) the lowering of the Policy Relevant Baseline. Without such disaggregation, policymakers, decision makers, and the public will be misled to believe that all or virtually all of the projected health benefits are attributable to the reduction in the standard.

C. The Combined Effect of EPA's Rollback and Policy Relevant Background Assumptions

Figure C illustrates the combined effect of these two critical assumptions. Ozone emissions are scaled on the horizontal axis and divided into biogenic and anthropogenic sources, with the latter category further subdivided into U.S., Canada, Mexico, and non-North American sources. The distances between the vertical boundaries are arbitrary.

EPA's Policy Relevant Background (PRB) is shown by the transparent rectangle that ranges from green on the right to red on the left. The colors are selected to represent the feasibility of control. The left side is red for two reasons. First, EPA has no jurisdiction over anthropogenic emissions from Canada and Mexico. Its ability to affect those emissions depends on either those sovereign nations deciding to implement all or part of EPA's standard, or States (especially those on the borders) states obtaining external emission reductions where that is cost-effective.

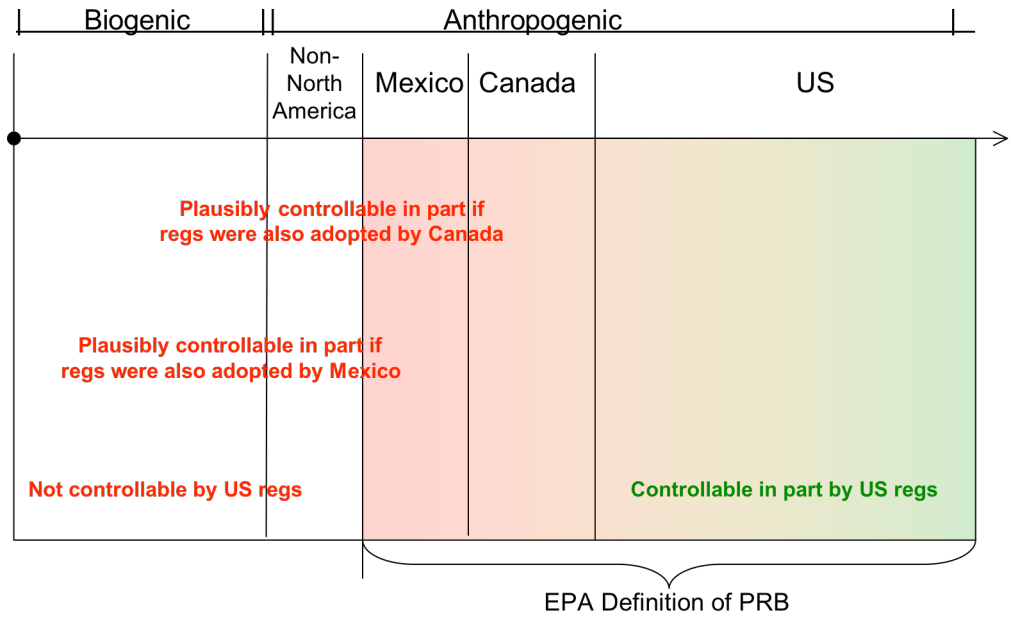
Second, it is technically infeasible to eliminate all U.S. anthropogenic emissions. It is also technically infeasible to achieve the revised ozone standards proposed in the NPRM, even if all known and reasonably anticipated control technologies are applied (U.S. Environmental Protection Agency 2007e).

Only biogenic and non-North American emissions are excluded from the PRB. Long-distance transport of ozone from Asia is rising due to the rapid industrial development now underway (Ortmans et al. 2006; Vingarzan 2004).

EPA's rollback assumption is essentially proportional from the minimum peak reduction needed to achieve the new standard. It supposes that a reduction in ozone peaks will be implemented by a downward shift in all emissions. This assumption is so unrealistic that EPA's Regulatory Impact Analysis characterizes the predicted reductions as "illustrative." Risk reduction estimates based on this

assumption do not adhere to applicable information quality standards because they cannot be objective, even if the underlying risk estimates themselves were

Figure C: Policy Relevant Background and Rollback:Which Emissions Belong? Which Can Be Controlled?



known with certainty and accurately calculated.

VIII. Conclusion

The Preamble to the Notice of Proposed Rulemaking and its supporting documents display a pattern of systematic bias in the analysis and presentation of scientific information. The likelihood that EPA has overestimated ozone health risks is significantly greater than the likelihood that it has underestimated them. This systematic bias unambiguously violates applicable information quality guidelines. This misleads Agency decision-makers and public about both the health risks posed by exposure below the current NAAQS and the amount of risk reduction that would be achieved by setting and implementing a more stringent primary standard.

At each step of its risk assessment, EPA has faced scientific uncertainties that, according to accepted practice in the field for a risk assessment of this magnitude, would be incorporated into the analysis and carried forward. But at each step, EPA has chosen instead to discard uncertainty and incorporate only the assumption, datum, model specification, or parameter value that leads to the highest estimate of health risk. This "envelope" approach to risk assessment provides only a plausible upper bound, which is often called a "screening level risk assessment." Such risk assessments satisfy applicable information quality guidelines only when they are properly characterized as such and are used for the purpose for which they were intended -- to distinguish *de minimis* risks, risks that do not warrant greater analytical effort to improve accuracy (National Research Council 1994).[34] In this case, however, EPA proposes to report a screening level approximation *in lieu of a properly perform risk assessment*. Moreover, in its NPRM EPA relies on screening level risk assessment as the foundation for a major public health decision.

By law, the EPA Administrator has sole discretion to make crucial policy judgments concerning the ozone NAAQS. It is beyond the role and authority of Agency scientists and program managers to exercise this judgment on his behalf. For the Administrator to exercise his statutory authority, the Clean air Act requires that the scientific information presented to him "accurately reflect the latest scientific knowledge useful in indicating the kind and extent of all identifiable effects on public health or welfare which may be expected from the presence of such pollutant in the ambient air, in varying quantities" (§108(a)(2)). These requirements foreshadowed the enactment of the Information Quality Act, which directed the establishment of government-wide criteria for information quality. These criteria are entirely consistent with the directives in Clean Air Act § 108. Nothing in that section, or in § 109, authorizes the Administrator to set air quality standards based on scientific information that is inaccurate.

Yet that is precisely what EPA staff have generated for the Administrator's use. He has been given a risk assessment that systematically violates the information quality standard of objectivity and thus cannot conform to the requirements of Clean Air Act § 108. To exercise his authority in

---

[34] "[S]creening analyses need to incorporate conservative assumptions to preclude the possibility that a pollutant that poses dangers to health or welfare *will not receive full scrutiny*" (p. 246, emphasis added), a very different situation that which applies here, in which greater scrutiny is difficult to imagine.

compliance with the Clean Air Act, these systematic biases must be removed from the ozone risk assessment and supporting documents.

IX. References

Adams WC. 2002. Comparison of Chamber and Face-Mask 6.6-hour Exposures to Ozone on Pulmonary Function and Symptoms Responses. Inhalation Toxicology 14: 745-764.

Adams WC. 2003a. Comparison of Chamber and Face Mask 6.6-hour Exposures to 0.08 ppm Ozone via Square-Wave and Triangular Profiles on Pulmonary Responses. Inhalation Toxicology 15: 265-281.

Adams WC. 2003b. Relation of Pulmonary Responses Induced by 6.6-h Exposures to 0.08 ppm Ozone and 2-h Exposures to 0.30 ppm Ozone via Chamber and Face-Mask Inhalation. Inhalation Toxicology 15(8): 745-759.

Adams WC. 2006a. Comparison of Chamber 6.6-hr Exposures to 0.04-0.08 ppm Ozone via Square-Wave and Triangular Profiles on Pulmonary Responses. Inhalation Toxicology 18: 127-136.

Adams WC. 2006b. Human Pulmonary Responses with 30-minute Time Intervals of Exercise and Rest When Exposed for 8 hours to 0.12 ppm Ozone via Square-Wave and Acute Triangular P rofiles. Inhalation Toxicology 18: 413-422.

Adams WC. 2007. Public Comment to CASAC Ozone Review Panel Teleconference. In: Clean Air Act Scientific Advisory Committee. Washington, D.C., 1-2.

Brown JS. 2007. The effects of ozone on lung function at 0.06 ppm in healthy adults (EPA-HQ-OAR-2005-0172-0175). Research Triangle Park, N.C.: US Environmental Protection Agency 1-8.

Clean Air Scientific Advisory Committee. 2006a. Clean Air Scientific Advisory Committee's (CASAC) Peer Review of the Agency's 2nd Draft Ozone Staff Paper. 112.

Clean Air Scientific Advisory Committee. 2006b. Clean Air Scientific Advisory Committee's (CASAC) Teleconference Meeting to Provide Additional Advice to the Agency Concerning Chapter 8 (Integrative Synthesis) of the Final Ozone Air Quality Criteria Document (AQCD). 1-42.

Clean Air Scientific Advisory Committee. 2007. Clean Air Scientific Advisory Committee's (CASAC) Review of the Agency's Final Ozone Staff Paper.

Gent MH, Triche EW, Holford TR, Belanger K, Bracken WS, Beckett MR, et al. 2003. Association of Low-level Ozone and Fine Particles with Respiratory

Symptoms in Children with Asthma. Journal of the American Medical Association (JAMA) 290: 1859-1867.

Gong H, Wong R, Sarma RJ, Linn WS, Sullivan ED, Shamoo DA, et al. 1998. Cardiovascular Effects of Ozone Exposure in Human Volunteers. American Journal of Critical Care Medicine 170: 1080-1087.

Health Effects Institute. 2003. Revised Analyses of Time-Series Studies of Air Pollution and Health; Revised Analyses of the National Morbidity, Mortality, and Air Pollution Study, Part II; Revised Analyses of Selected Time-Series Studies. Cambridge, Mass.: Health Effects Institute.

Hill AB. 1965. The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine 58: 295-300.

Ioannidis JPA. 2005. Why Most Published Research Findings Are False. PLoS Medicne 2(8): 696-701.

Kamps AW, Roorda RJ, Brand PL. 2001. Peak flow diaries in childhood asthma arew unreliable. American Journal of Epidemiology 46: 750-762.

Kerr NL. 1998. HARKing: Hypothesizing After the Results are Known. Personality and Social Psychology Review 2(3): 22.

Kimata H. 2004. Effect of Viewing a Humorous vs. Nonhumorous Film on Bronchial Responsiveness in Patients with Bronchial Asthma. Physiology and Behavior 81(4): 681-684.

Korrick SA, Neas LM, Dockery DW, Gold DR, Allen GA, Hill LB, et al. 1998. Effects of Ozone and Other Pollutants on the Pulmonary Function of Adult Hikers. Envionmental Health Perspectives 106: 93-99.

Langstaff J. 2007. Analysis of Uncertainty in Ozone Population Exposure Modeling. In: EPA-HQ-OAR-2005-0172-0174 (US Environmental Protection Agency, ed). Research Triangle Park, N.C., 1-87.

Lianges G, Morton JR, Henry RL. 2003. Mirth-triggered Asthma: is Laughter Really the Best Medicine? Pediatric Pulmonology 36(2): 107-112.

Lianges G, Yates DH, Wu D, Henry RL, Thomas PS. 2004. Laughter-associated Asthma. Journal of Asthma 41(2): 217-221.

Lumley T, Sheppard L. 2003. Time Series Analyses of Air Pollution and Health: Straining at Gnats and Swallowing Camels? Epidemiology 14(1): 13-14.

Martin RM. 2002. There Are Two Errors in the Title of This Book: A Sourcebook of Philosophical Puzzles, Paradoxes and Problems: Broadview Press.

Moolgavkar S. 2007. A Critical Review of the Staff Paper on Ozone. US EPA Docket: EPA-HQ-OAR-2005-0172-0493.1, 476-484.

Moolgavkar SH. 2000. Air Pollution and Mortality in Three U.S. Counties. Envionmental Health Perspectives 108: 777-784.

Moolgavkar SH, Luebeck EG, Hall TA, Anderson EL. 1995. Air Pollution and Daily Mortality in Philadelphia. Epidemiology 6: 476-484.

Morgan MG, Henrion M, Small M. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. New York, N.Y.: Cambridge University Press.

Mortimer KM, Neas, L.M.,, Dockery DW, Redline S, Tager IB. 2002. The Effect of Air Pollution on Inner-city Children with Asthma. The European Respiratory Jornal 19(4): 699-705.

National Research Council. 1983. Risk Assessment in the Federal Government: Managing the Process. Washington, D.C.: National Academies Press.

National Research Council. 1994. Science and Judgment in Risk Assessment. Washington, D.C.: National Academies Press.

National Research Council. 2002. Estimating the Public Health Benefits of Proposed Air Pollution Regulations. Washington, D.C.: National Academy Press.

Office of Management and Budget. Standards and Guidelines for Statistical Surveys. Available: http://www.whitehouse.gov/omb/inforeg/statpolicy/standards_stat_surveys.pdf [accessed October 1 2007].

Office of Management and Budget. 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Notice; Republication. Federal Register 67(36): 8452-8460.

Ortmans SJ, Lefohn AS, Harris JM, Galbally I, Scheel HE, Bodeker G, et al. 2006. Long-term Changes in Tropospheric Ozone. Atmospheric Environment 40: 18.

Queen's University Clinical Skills Education Centre. 2007. Peak expiratory flow rate measurement. Available: http://www.qub.ac.uk/cskills/PEFR.htm [accessed October 8 2007].

Sarafino EP, Gates M, DePaulo D. 2001. The role of age at aasthma diagnosis in the development of triggers of asthma episodes. Journal of Psychosomatic Research 51(5): 623-628.

Sarnat JA, Brown KW, Schwartz J, Coull BA, Koutrakis P. 2005. Ambient gas concentrations and personal particulate matter exposures: implications for studying the health effects of particles. Epidemiology 16(3): 385-395.

Sarnat JA, Schwartz J, Catalano PJ, Suh HH. 2001. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? Envionmental Health Perspectives 109(10): 1053-1061.

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. American Journal of Epidemiology 164(6): 505-517.

Section 515, Treasury and General Government Appropriations Act for Fiscal Year 2001. [2000]. In: 44 USC 3516 note.

The National Academies. 2003. Policy on Committee Composition and Balance and Conflict of Interest. The National Academies.

U.S. Environmental Protection Agency. 2002. Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency (EPA/260R-02-008).

U.S. Environmental Protection Agency. 2005a. Guidelines for Carcinogen Risk Assessment (EPA/630/P-03-001F).

U.S. Environmental Protection Agency. 2005b. Plan for Review of the National Ambient Air Quality Standards for Ozone. 1-22.

U.S. Environmental Protection Agency. 2006a. Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume I of III (EPA 600/R-05/004aF). In: AQCD Vol I, 821.

U.S. Environmental Protection Agency. 2006b. Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume II of III (EPA 600/R-05/004bF). In: AQCD Vol II, 873.

U.S. Environmental Protection Agency. 2006c. Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume III of III (EPA 600/R-05/004cF). In: AQCD Vol III, 424.

U.S. Environmental Protection Agency. 2007a. Analysis of Uncertainty in Ozone Population Exposure Modeling. In: ("John Langstaff Memo").

U.S. Environmental Protection Agency. 2007b. National Ambient Air Quality Standards for Ozone: Proposed Rule. Federal Register 72: 37818-37919

U.S. Environmental Protection Agency. 2007c. Ozone Health Risk Assessment for Selected Urban Areas (EPA-452/R-07-009). In: ("Ozone Health Risk Assessment").

U.S. Environmental Protection Agency. 2007d. Ozone Population Exposure Analysis for Selected Urban Areas (EPA 452/R-07-004). In: ("Ozone Population Exposure Analysis").

U.S. Environmental Protection Agency. 2007e. Regulatory Impact Analyses: 2007 Proposed Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone. In: ("Ozone RIA").

U.S. Environmental Protection Agency. 2007f. Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information: OAQPS Staff Paper (EPA-452/R-07-007). In: Staff Paper (Final Updated).

U.S. Environmental Protection Agency Office of the Science Advisor. 2004. An Examination of EPA Risk Assessment Principles and Practices; Staff Paper, EPA/100/B-04/001. Available: http://www.epa.gov/osainter/pdfs/ratf-final.pdf [accessed August 30 2005].

Vingarzan R. 2004. A Review of Surface Ozone Background Levels and Trends. Atmospheric Environment 38: 12.

Whitman v. American Trucking Ass'ns I. 2001. Whitman, Administrator of Environmental Protection Agency et al. v. American Trucking Associations, Inc. et al. In: US: US Supreme Court, 457.