John Engler

President and CEO

October 14, 2008

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

Delivered via email to: mailto:quality.guidelines@epa.gov

Dear Ms. O'Neill:

On behalf of the National Association of Manufacturers (NAM), I am hereby submitting the attached Request for Reconsideration (RFR) in accordance with the procedures set forth in Section 8.6 of EPA's Information Quality Guidelines. The RFR concerns the Request for Correction (RFC) that the NAM submitted on October 9, 2007, logged in by your office as RFC #08001, to which EPA replied in its response to comments on the final revised ozone National Ambient Air Quality Standard (NAAQS). Per EPA's recommendation in Section 8.6 of its Information Quality Guidelines, I am attaching a copy of the RFC.

Among other issues, the RFR shows that many of the epidemiological studies EPA staff find persuasive used research designs that were known at the time to be demonstrably substandard. In some cases, EPA staff have relied on complex statistical methods to coax data into revealing effects from ozone so small that humans cannot even recognize experiencing them. Finally, EPA staff insist that certain studies provide valid and reliable evidence of respiratory health effects from ozone even though they rejected these same studies in their July 2007 draft Integrated Science Assessment for Oxides of Nitrogen -- and for the same reasons we mentioned in the RFC. Through the appeal, I seek more cogent answers than EPA provided in its response to the RFC. The document also identifies a number of process changes that are necessary to ensure that future NAAQS reviews fully and consistently adhere to the Agency's Information Quality Guidelines and the Information Quality Act.

The NAM appreciates the EPA's desire that stakeholders submit an RFR as promptly as possible and acknowledges the complexity of this issue area. The NAM has worked diligently to provide a document in a timely manner that articulates the association's concerns as thoroughly as possible. However, EPA's response to the RFC was scattered throughout both a 210-page Response to Comments document and the preamble of the final rule, which prolonged the analysis of EPA's response to the original petition, and therefore submission of the RFR.

Thank you for your consideration of the NAM's Request for Reconsideration. If you have any questions related to the attached RFR, please contact Bryan Brendle of the NAM staff at bbrendle@nam.org, or (202) 637-3176.

Sincerely,

John Engler

JE/blb

Attachments:
1) Request for Reconsideration
2) Request for Correction (filed with the EPA on October 9, 2007).

Table of Contents

I. Summary 8

A. Procedural Basis for this Request for Reconsideration 8

B. EPA's Response to Comments Offers No Evidence that the Agency Adhered to Its Own Information Quality Principles, Policies and Procedures 11

C. The Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization 13

D. Types of Major Information Quality Errors in the Scientific Record for the Ozone NAAQS Review 14

1. EPA omitted any reference to information quality principles and own Information Quality Guidelines from every document in the ozone NAAQS review, stretching from the 2005 Review Plan to the 2007 NPRM. 14

2. EPA "considers" and "discusses" a phenomenal quantity of scientific information, but only uses information in accordance with the Four EPA Staff Principles. 14

3. EPA makes crucial claims that are easily refutable. 15

4. EPA uses ad hoc statistical analyses devised after the data were obtained. 15

5. EPA disseminates risk characterizations based on epidemiological studies that use unvalidated self-reported data collected in diaries. 16

6. EPA disseminates risk characterizations based on pulmonary function data obtained through a low-resolution clinical diagnostic procedure that cannot reliably or accurately detect subtle effects. 16

7. EPA disseminates risk characterizations based on epidemiological studies of pulmonary function in which the research design discards variability and uncertainty, thus making association with air pollutants appear to be much more certain than they actually are. 16

8. EPA disseminates risk characterizations based on epidemiological studies of pulmonary function in which the research design requires the use of biased estimates. 17

9. EPA disseminates risk characterizations based on epidemiological studies of unrepresentative samples, or samples whose representativeness has not been validated. 18
10. EPA disseminates risk characterizations based on epidemiological studies with unaccounted for or unreported nonresponse bias. 18
11. EPA disseminates risk characterizations based on ambient monitoring data as a proxy for personal exposure despite very low correlation. 18
12. EPA disseminates risk characterizations based on studies using research methods Agency staff have rejected as unreliable and invalid for other air pollutants. 19
13. EPA never informed CASAC about information quality principles or the Agency's Information Quality Guidelines, and information quality played no role in CASAC's review even though it is required by EPA's Peer Review Handbook 20

E. Remedies Requested 20

1. Explicitly, comprehensively, and consistently implement information quality principles and practices throughout the NAAQS review process and within each NAAQS work product 21
2. Establish an external independent body with the limited responsibility of reviewing the quality of scientific and technical information, and advising the Administrator whether information quality principles have been met and applicable information quality policies and practices have been followed 21
3. Remove all policy judgments and similar considerations from the assembly, review and presentation of scientific information in all EPA NAAQS work products. 22
4. Establish new and publicly accountable pre-dissemination review procedures for all EPA NAAQS work products. 22
5. Establish an information quality foundation for all CASAC reviews of NAAQS-related work products whose successful performance is documented and independently validated 23
6. Require that CASAC panel members be recused from the review of their own research or the research of their institutional colleagues, and such individuals should not serve on a CASAC panel in cases

where such research is crucial to risk assessment and characterization. 23
II. Introduction 24
A. Information Subject to this Request for Reconsideration 24
B. Affected Party 25
C. Applicable Error Correction Procedures 26
D. Relevant Information Quality Principles 26
1. Failure to adhere to the objectivity standards 27
2. Failure to adhere to the utility standard 27
III. Information Quality Errors in the Description, Analysis and Presentation of Scientific Information 28
A. EPA Begins with a Structurally Biased Scientific Database 30
1. Effective control by a party with a risk management objective 30
2. Publication bias 33
(a) Outcome-reporting bias 36
(b) Inferential exaggeration 37
(i) The inferences made by Gent et al. (2003) are inconsistent with their own reported results 38
(ii) The inferences made by Mortimer et al. (2002) are inconsistent with their own reported results 42
(c) Methodological error 45
(d) Pulmonary function testing 50
(i) Clinically useful pulmonary function tests have inherent information quality limitations that become defects when used in air pollution epidemiology 50
(ii) Information quality defects associated with investigator bias 53
(iii) Information quality defects associated with the use of diaries 53
(iv) Information quality defects associated with inter-maneuver variability 55
(v) Information quality defects resulting from nonresponse bias 57

(vi) EPA's use of PEFR data depends on whether the results support an inference of pollutant-related health effect 58
3. Peer review practices 60
B. Non-disclosure of critical studies and analyses 61
1. Vingarzan (2004) and Oltmans et al. (2006) 61
2. Brown (2007a) 63
(a) EPA's reanalysis of Adams (2006a) is technically defective. 64
(b) EPA's explanation for why it reanalyzed selected data from Adams (2006a) is materially incomplete and misleading 64
(c) Finding statistically significant effects at 0.06 ppm required EPA to use creative statistical methods 71
(d) The policy-relevant background ozone concentration 76
(e) The policy-relevant ozone exposure wave pattern 77
(f) CASAC's "support" for Brown (2007a) is technically infeasible and contradicted by the recollections of some of the principals 78
3. Gerrymandering the scientific record 80
(a) EPA staff risk assessment methods show a preference for research showing positive effects 81
(b) Some CASAC panel members prefer research showing positive effects 82
(c) Gray literature and "personal communications" 83
C. EPA Interprets and Presents Scientific Information in a Systematically Biased Manner 84
1. The inclusion or exclusion of data or studies based on the extent to which they support stated or unstated risk management objectives 84
2. The inclusion or exclusion of data or studies based on post hoc or non-transparent criteria 85
3. Mischaracterization of results 86
(a) Characterizing a study as "new" since the last ozone NAAQS review when in fact it was part of the last review 87
(b) Characterizing a study as reporting something about which it is silent 89

(c) Characterizing a study as reporting something when it reports the opposite 89
(d) Selective and misleading citation 90
(e) Drawing inferences from a study that are not supported by the data and analysis reported 93
(f) Utilizing for one purpose data that were collected for another purpose 95
(g) Hypothesizing after the results are known 97
4. Study selection bias 97
5. Model selection bias 99
(a) Selecting models based on criteria other than quality of data or analysis 100
(b) Control for known confounders other than air pollution 101
(c) Selecting models known to yield upwardly biased risk estimates, such as single-pollutant models and models that do not control for known confounders 102
(d) In time series models, choosing lags based on statistically convenient but biologically implausible criteria 103
6. Assumption of causality 104
D. EPA's Risk Assessment Is Biased as a Matter of Policy 106
E. EPA Attributes to Ozone Risks That It Has Previously Attributed to other Pollutants 109
IV. Information Quality Errors in the Assessment of Human Health Risk 109
A. EPA Treats Transient and Reversible Effects as Adverse 109
B. EPA Uses Important Scientific Terms and Language in Policy-directed Ways 111
1. The definition of "likely" has nothing to do with "precision." 112
2. If there is no scientific basis for probabilistic language, EPA should not use it 113
3. EPA staff have available to them – and have used in other contexts – clearly defined meanings for probabilistic terms. 114
4. Ad hoc meanings for probabilistic language are not compatible with information quality. 114

C. EPA Confuses Variability and Uncertainty 116
D. EPA Does Not Disclose a Credible Analysis of Uncertainty 118
E. EPA's Particular Use of Default Values Violates Information Quality Principles 119
F. EPA Assumes Confidence Intervals Adequately Describe Variability and Uncertainty 122
G. EPA Assumes that Ambient Monitors Provide Unbiased Estimates of Personal Exposure 124
H. EPA Assumes that Associations Observed in Short-Term Time Series Studies Are Significant and Meaningful, but the Absence of Associations in Long-Term Cohort Studies Is neither Significant nor Meaningful nor Logically Inconsistent 127
I. EPA Assumes Causality 127
J. EPA Does Not Explain the Effects of Ozone with Reference to Any Non-Air Pollution Context 128
K. Double-counting 128
L. EPA's Alternative Risk Estimates 129
V. Information Quality Errors in the Consideration of Reports from CASAC 131
A. CASAC's Scientific Charge 132
B. CASAC's Policy Advice Charge 134
C. EPA Does Not Adequately Distinguish Between Scientific Insight and Policy Advice It Received from CASAC 137
D. EPA's Lack of Pre-Dissemination Review 137
VI. Information Quality Errors in the Rollback Assumption 138
VII. Information Quality Errors in the Description of Policy Relevant Background 139
A. EPA's *Definition* of Policy Relevant Background is Biased 139
B. EPA's *Estimates* of the Magnitude of Policy Relevant Background Are Biased 144
VIII. Conclusion 145
IX. References 146

I. Summary

A. Procedural Basis for this Request for Reconsideration

This Request for Reconsideration (RFR) 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, pp. 47-49, emphasis added).

EPA is publicly committed to the principles of information quality. It is established 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.

On October 9, 2007, the National Association of Manufacturers, in adherence to procedures established by EPA in its 2002 Information Quality Guidelines, submitted a Request for Correction (RFC) contending that EPA's Notice of Proposed Rulemaking (NPRM) announcing the intent to revise the National Ambient Air Quality Standard (NAAQS) for ozone, and several supporting documents, each contained influential scientific information crucial for regulatory decision making under §§ 108 and 109 of the Clean Air Act that materially violated these standards (National Association of Manufacturers 2007). The RFC did not contest the statutory authority of the Administrator to make this decision; to make it promptly; or the nature of the criteria he was

required to take into account in making his decision. Indeed, the RFC makes clear that EPA's adherence to the information quality standards that the Agency committed to uphold was the best and surest way to fulfill this statutory mandate.

EPA responded to our RFC as part of its general response to significant public comments, which the Agency is required to prepare in compliance with Section 307(d)(6)(B) of the Clean Air Act. In its Information Quality Guidelines, EPA committed to integrate its responses to RFCs submitted in the context of regulatory actions within its regular responsibilities under the Administrative Procedure Act (and in this case, the Clean Air Act). We commend EPA for adhering to this important procedural element of its Information Quality Guidelines.

Unfortunately, EPA has not adhered to the substantive elements of its Information Quality Guidelines. Having carefully reviewed EPA's 210-page response, we can discern no instance in which the Agency conceded even the smallest of information quality error. Sixteen times, EPA said it "rejected" our concerns and complaints, often without any presentation of substantive data or argument. Twelve times EPA said it "disagrees" with us regarding the objectivity of a purported statement of fact, knowledge, or scientific inference, as if science can be reduced to a matter of opinion. Based on this review, we have concluded that it is necessary under the Information Quality Act to exercise our statutory right to seek and obtain the correction of error by means of the appeal procedures required by law and prescribed by EPA's Information Quality Guidelines.

EPA responded to our RFC two ways. First, EPA responded procedurally in a letter dated January 3, 2008, stating:

> The Administrator will issue his final decision on the ozone standards by March 12, 2008. At that time, EPA will respond to each of the issues raised in the RFC and other public comments received on the NPRM, either in the preamble to the final rule itself or in the accompanying Response to Comments document which will be placed in the rulemaking docket at the time the final rule is signed (Meyers 2008).

Second, EPA responded substantively to our RFC through the response to comments document that it normally publishes pursuant to Section 307(d)(6)(B) of the Clean Air Act (U.S. Environmental Protection Agency 2008d, hereinafter "Response to Comments") and in the preamble to the final rule. The preamble references NAM's public comment as an information quality RFC and mentions

it four times in the text.[1] The 210-page EPA Response to Comments cites the NAM public comment at least 45 times and twice notes that it was an RFC (pp. 150, 158).[2] Although we were the only public commenter to submit an RFC, EPA's Response to Comments includes responses to significant comments made by dozens of other public commenters.

EPA's decision, consistent with its Information Quality Guidelines, to incorporate its administrative procedures for managing information quality error correction requests within its normal rulemaking procedures, explains why EPA's document does not organizationally track our RFC. Indeed, sorting through the Response to Comments has been challenging. The document is redundant in places, and it includes comments ascribed to our RFC that we do not recognize having made.

This created some difficulty in determining which portions of EPA's Response to Comments are germane to our RFC. We have settled on what we think is a reasonable interpretative strategy:

- Where EPA's Response to Comments mentions a comment that it ascribes to public commenters other than NAM or to unidentified commenters, we interpret this to be not part of EPA's response to our RFC.
- Where EPA's Response to Comments mentions a comment that it ascribes to multiple commenters using the form (ABC, XYZ, NAM) without page numbers, we interpret this to not be part of EPA's response to our RFC unless the issue at hand is strictly scientific.[3]
- Where EPA's Response to Comments mentions a comment by NAM using the format (NAM, p. *x*), we interpret this to be a formal response to the RFC.

---

[1] EPA (2008b, pp. 16454, 16457, 16466, and 16469).

[2] A search of the EPA docket reveals a 72-page document styled as a response to the NAM RFC. However, this document also includes responses to many other public commenters. We infer that this other document is not authoritative but was placed in the docket only because it was shared with the Office of Management and Budget. See (U.S. Environmental Protection Agency 2008c).

[3] Separately from the RFC, which is strictly limited to scientific, statistical and technical matters covered by applicable information quality guidelines, NAM also submitted a traditional public comment that addressed a broad array of issues including policy considerations that are not subject to applicable information quality guidelines.

This RFR is limited to matters within this third category.

Finally, although the text of EPA's January 2008 letter advised us that EPA might respond to our RFC within "the preamble to the final rule itself," we cannot find any text in the preamble that reasonably can be construed as a response to the RFC as opposed to a response to comments more generally.[4] Therefore, in this RFR we focus on EPA's Response to Comments as the authoritative EPA response to our RFC. For everyone's convenience, we follow the structure and organization of our RFC (which contains only information quality-related issues) rather than EPA's Response to Comments (which contains responses to all significant comments, including comments made by others and a large number of comments on policy).

B. EPA's Response to Comments Offers No Evidence that the Agency Adhered to Its Own Information Quality Principles, Policies and Procedures

EPA's Response to Comments proves beyond any reasonable doubt that until we submitted our RFC, EPA staff, management, and policy officials had devoted no attention to information quality in the revision of the ozone NAAQS. In every EPA staff document, beginning with the Review Plan (U.S. Environmental Protection Agency 2005e), proceeding to the Criteria Document (U.S. Environmental Protection Agency 2006a, 2006b, 2006c), the Exposure Assessment and Risk Assessment (U.S. Environmental Protection Agency 2005a, 2005c, 2006g, 2006h, 2006i), and the Staff Paper (U.S. Environmental Protection Agency 2006f, 2007j), there is no mention, discussion, analysis or any other content mentioning, discussing or applying the requirements of the Information Quality Act, the government-wide implementing guidance issued to all agencies by the Office of Management and Budget (Office of Management and Budget 2002), or EPA's own implementing guidelines (U.S. Environmental Protection Agency 2002, 2003, 2006e). EPA's Information Quality Guidelines require that Agency program offices perform sufficient pre-dissemination review to ensure that the quality of information that is disseminated is maximized. However, the Response to Comments shows that EPA staff performed no pre-dissemination review of the information quality aspects any of the scientific information that they transmitted to the Administrator in support of his policy decision.

---

[4] At 73 Fed. Reg. 16510, EPA cites our submission as both a "letter" (i.e., a "Request for Correction") to EPA Assistant Administrator (Environmental Information) Molly A. O'Neill and as "public comment" to Docket No. EPA-HQ-OAR-2005-0172. Some text in the preamble to the final rule is essentially identical to text in the Response to Comments.

By law, EPA's technical staff work products must be peer reviewed by the Clean Air Scientific Advisory Committee (CASAC). EPA peer review guidance, which specifically covers peer reviews such as the one performed by CASAC, commit the Agency to ensure that peer reviews fully address information quality issues (U.S. Environmental Protection Agency 2006e). However, information quality was omitted from the panel's charge. CASAC meetings are dialogues between panel members and EPA managers and staff, yet the transcripts of each in-person meeting shows that neither the principles nor the procedural and substantive requirements of information quality were ever mentioned by any EPA manager or staff member.

The absence of information quality from every aspect of the ozone NAAQS review is complete and comprehensive. Yet in its Response to Comments, EPA "rejects," "disagrees" with, or otherwise denies each and every information quality error claim in our RFC:

> EPA has reviewed NAM's RFC and finds that there is no merit to their objections. EPA disagrees with NAM's allegations that EPA has not complied with the requirements of the Information Quality Act or the Agency's policies for ensuring information quality. EPA has responded to NAM's significant comments in the preamble to the final rule or in this document (U.S. Environmental Protection Agency 2008e, p. 158).

EPA staff devoted thousands of man-hours and millions of dollars to assemble, summarize, analyze, and write thousands of pages of scientific reports for this review. We have found not a single page that concerns information quality.

This RFR responds to EPA's replies in the same fashion that the RFC was written. Where EPA has provided persuasive evidence that it is correct or that our evidentiary case is insufficient, we withdraw our request for correction. Where EPA's reply is problematic, however, we have summarized or restated our claims and put them forward again on appeal. In reviewing EPA's replies, we have discerned certain patterns. Many of EPA's replies fall into one or more of the following categories:

- EPA has mischaracterized our information quality claim, often in the form of a straw man, and responded to its mischaracterization rather than our claim.
- EPA has mischaracterized our information quality claim as a matter of opinion, as if representations of knowledge such as facts and data can be subjectively determined, then asserted that its opinion is superior to ours.

- EPA has mischaracterized a scientific issue as one determined by law or policy judgment.
- EPA has characterized our information quality claim accurately, but responded to an irrelevant or unrelated issue or merely responded with boilerplate.
- EPA has responded to our complaint of information quality error by committing a new information quality error, typically by making new informational statements that fail the substantive and/or presentational objectivity standards.

C. The Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization

A comprehensive review of our information quality error correction claims and EPA's responses has led us to discern a Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization that explains how EPA staff utilize scientific information. The Iron Law is set forth in the nearby text box, and we refer to it frequently in this RFR. The scope, scale and magnitude of risk can be understood as an envelope or a balloon; higher risk means a more expansive envelope or a larger balloon.

Science suggesting the potential for greater risk pushes the risk envelope outward or adds air to the balloon. Science that is equivocal supports the envelope at its current location or maintaining the balloon at its current size. Science suggesting lower risk moves the envelope inward or removes air from the balloon, but EPA staff will use such information only under conditions that are so restrictive as to be nearly impossible to meet. Science that does not meet these conditions is "discussed" or "considered," but ultimately discarded. The principles of information quality play a severely constrained role: they are used only as barriers to the admission of evidence indicating lower risk.

Ironically, the foundation for the Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization was first set forth by EPA staff itself, in a 2004 report titled *An Examination of EPA Risk Assessment Principles and Practices* (U.S. Environmental Protection Agency Office of the Science Advisor 2004b). In that report, EPA staff elucidated publicly for the first time that it is the policy of EPA staff not to understate risk or to grossly overestimate it.

Text Box 1:
The Iron Law of EPA Staff
Ozone Health Risk Assessment and Characterization

1. EPA staff use new scientific information suggesting greater potential or actual health risk as evidence that risk is greater than previously believed. EPA staff make no practical distinction between health risk that is potential (i.e., possible, hypothetical, speculative) or actual (i.e., proved).
2. EPA staff use new scientific information that is equivocal to support their existing assessment and characterization of health risk.
3. EPA staff use scientific information suggesting lesser potential or actual health risk as evidence that risk is lower than previously believed, provided that its quality is flawless in every respect.
4. EPA staff use scientific information suggesting the absence of human health risk only if it proves that risk is biologically infeasible.

D. Types of Major Information Quality Errors in the Scientific Record for the Ozone NAAQS Review

This RFR documents a long list of information quality errors, but some errors clearly are more significant and important than others. We present a Baker's Dozen below:

1. *EPA omitted any reference to information quality principles and own Information Quality Guidelines from every document in the ozone NAAQS review, stretching from the 2005 Review Plan to the 2007 NPRM.*

EPA's Information Quality Guidelines commit the Agency to instill information quality principles and practices throughout its regulatory development process. Yet, information quality is completely missing from the ozone review. It is impossible for EPA to simultaneously have adhered to information quality principles yet have been utterly silent about them.

2. *EPA "considers" and "discusses" a phenomenal quantity of scientific information, but only uses information in accordance with the Four EPA Staff Principles.*

For many of the information quality issues raised in the RFC, EPA says in its Response to Comments that it "considered" or "discussed" it, usually in the

Criteria Document. While it is certainly important for EPA staff to have done these things, they are not the same thing as having examined and evaluated the information quality attributes of scientific information and accounted for information quality throughout the review. The comprehensiveness of EPA's discussion was not the focus of our RFC; we objected to the lack of objectivity in these documents.

We are unable to locate a single scientific study that both pushes the ozone risk envelope outward and was excluded by EPA. Similarly, we cannot identify a single scientific study that EPA puts any weight upon which pushes the ozone risk envelope inward.

3. *EPA makes crucial claims that are easily refutable.*

The quintessential example is EPA's untruthful explanation of the origin for its reanalysis of selected data from the controlled human study by Adams (2006a). EPA claims that its reanalysis was prompted by public comments to CASAC provided by Smith (2007b) in March 2007. We prove that EPA's reanalysis was substantially completed by December 2006. EPA declined to distribute this work for timely public comment, and instead hid it from peer review by CASAC and placed it in the docket the same day that the Administrator signed the proposed rule. EPA also falsely claims that CASAC supported its reanalysis despite the absence of any CASAC review.

EPA's reanalysis (Brown 2007a) was a crucial element of the scientific record on which the Administrator relied to decide what ambient concentration of ozone is requisite to protect public health. EPA's description of its reanalysis of the Adams' data, culminating in the eleventh-hour insertion of Brown (2007a) into the scientific record without CASAC peer review, is deeply defective with respect to presentational objectivity.

4. *EPA uses ad hoc statistical analyses devised after the data were obtained.*

EPA has relied on controlled human exposure studies since at least the 1997 ozone NAAQS review. Indeed, the Agency has made significant investments in facilities and staff to perform controlled human studies. These studies have always been carefully designed (if not flawlessly implemented), and until now EPA has followed recognized and accepted statistical procedures for analyzing data.

For the first time, however, a controlled human study did not reveal statistically significant decrements in pulmonary response. In response, EPA first discarded the author's portrayal of his results, focused on selected observations from individual study subjects, and reanalyzed a gerrymandered subset of the

data to "prove" that the decrements observed are statistically significant after all. Through the vehicle of Brown (2007a), EPA has dispensed with longstanding statistical practice. When challenged, EPA staff defend this by incorrectly claiming that their statistical procedures are more commonly used than are those of the researcher whose analysis they reject.

5. *EPA disseminates risk characterizations based on epidemiological studies that use unvalidated self-reported data collected in diaries.*

Several of the epidemiological studies EPA staff rely upon are based on data obtained from diaries kept by study subjects or their caregivers. It has been shown that this research design results in self-invented data. Such data are unreliable unless significant proactive steps are taken, both in the recruitment phase and during study implementation, to ensure reliability and accuracy. These steps were not taken by the researchers who authored the studies that EPA uses to base its characterization of respiratory risks (Gent et al. 2003; Mortimer et al. 2002; Ross et al. 2002).

6. *EPA disseminates risk characterizations based on pulmonary function data obtained through a low-resolution clinical diagnostic procedure that cannot reliably or accurately detect subtle effects.*

The pulmonary function tests that epidemiologists use were intended for the clinical purpose of diagnosing disease and assigning patients into rough categories. They were never intended for measuring or estimating very small changes within individuals or across populations. Clinicians are trained to coach patients in their performance, a procedure that is reasonable for medical evaluation but improper for research purposes. When investigators are not blind, or multiple investigators with even subtly different coaching techniques are involved, the opportunities for error and bias are legion.

Several of the panel studies that EPA uses to characterize respiratory risk rely on these unreliable methods (Gent et al. 2003; Korrick et al. 1998; Ross et al. 2002). That is the case here, where epidemiologists seek to detect low single-digit percentage differences in pulmonary function and attribute these differences to air pollution.

7. *EPA disseminates risk characterizations based on epidemiological studies of pulmonary function in which the research design discards variability and uncertainty, thus making association with air pollutants appear to be much more certain than they actually are.*

Some of the epidemiological studies EPA staff rely upon are based on pulmonary function tests that measure phenomena that are both uncertain and variable. Researchers obtained data for multiple test maneuvers conducted at the

same time, but discarded data judged to be "unacceptable" and recorded only the average of the values not discarded. This practice violates a crucial assumption underlying the statistical tests that the epidemiologists subsequently performed. Assuming that uncertain phenomena are fixed when they are not reduces estimated standard errors, exaggerates statistical significance, and misleads decision-makers and the public about uncertainty and precision.

In studies of asthmatics and others with compromised pulmonary function, inter-maneuver variability can be very large. Assuming that inter-maneuver variability is zero yields unreliable and artificially narrow standard errors, and inflates statistical significance. It is virtually certain that discarded inter-maneuver variability exceeds in magnitude the small percentage pulmonary function decrements observed in these studies.

The consequence of discarding inter-maneuver variability is least important in controlled human studies of homogeneous subjects at high ozone concentrations. In these studies, inter-maneuver variability is low due to strict selection criteria for study subjects, and relatively large effect sizes are expected. Nevertheless, as the exposure concentration approaches background and the expected effect size approaches zero, the importance of discarded variance increases and almost certainly swamps the effect size. In the case of the only controlled human study to test for pulmonary function under exercise through 6.6 hours of exposures to 0.04, 0.06, and 0.08 ppm (Adams 2006a), study subjects provided pulmonary function data from at least two maneuvers such that the sum of FEV_1 and FVC was ± 200 ml, or about 3% on average. This variation alone is more than half the group mean pulmonary function decrement observed at 0.08 ppm, and is more than twice the group mean pulmonary function decrement observed at 0.06 ppm. Once inter-maneuver variability is accounted for, the minimum effect size that is truly detectable in controlled human studies may well be larger than the difference between ambient and background ozone levels.

8. *EPA disseminates risk characterizations based on epidemiological studies of pulmonary function in which the research design requires the use of biased estimates.*

Some of the epidemiological studies EPA staff rely upon for risk characterization are based on pulmonary function tests in which the fixed value recorded is the largest value obtained from a series of maneuvers performed in close sequence, not the average (Korrick et al. 1998; Mortimer et al. 2002; Ross et al. 2002). As indicated above, in clinical practice subjects are routinely coached to perform so that they achieve maximum results. Maximum performance is especially sensitive to coaching effectiveness (better coaches produce higher

maxima) and the number of maneuvers performed (more maneuvers increase the expected maximum). EPA staff have not acknowledged, much less analyzed, the consequences of bias in the pulmonary function testing performed by epidemiologists.

9. *EPA disseminates risk characterizations based on epidemiological studies of unrepresentative samples, or samples whose representativeness has not been validated.*

Several of the epidemiological studies EPA staff rely upon for risk characterization use convenience samples (Korrick et al. 1998) or cohorts whose representativeness has not been shown (Gent et al. 2003; Mortimer et al. 2002; Romieu et al. 1996). Convenience samples are presumptively unrepresentative. The representativeness of study cohorts that were assembled by non-randomized designs cannot be presumed. EPA staff rely upon coefficient estimates obtained from these studies and simply assume that the underlying samples are representative.

10. *EPA disseminates risk characterizations based on epidemiological studies with unaccounted for or unreported nonresponse bias.*

For at least two decades, federal statistical policy has required that surveys and similar studies achieve response rates that epidemiologists typically find problematic, and this policy has recently been codified in formal government-wide guidance (Office of Management and Budget 2006). Any federally-sponsored study – a term that generally includes EPA-funded epidemiology -- that is not expected or fails to achieve an 80% response rate must include a thorough nonresponse bias analysis. In the epidemiological studies EPA staff rely upon for risk characterization that have individual data, none achieved an 80% response rate and nonresponse bias analyses were either not performed or performed but not disclosed. Publication in a refereed journal confers a rebuttable presumption of "adequate" objectivity, but this presumption is automatically rebutted in any case where response rates do not satisfy applicable federal statistical policy standards and a rigorous nonresponse bias analysis was not performed.

11. *EPA disseminates risk characterizations based on ambient monitoring data as a proxy for personal exposure despite very low correlation.*

Several of the epidemiological studies EPA staff rely upon use ambient ozone data obtained from central monitoring sites as proxies for personal exposure despite overwhelming evidence that ambient and personal exposures are poorly correlated (Bell, McDermott et al. 2004; Gent et al. 2003; Mortimer et al. 2002). Despite this lack of correlation, EPA staff interpret observed weak

positive associations between ambient ozone and various health effects as causal. The best that can be said scientifically about these studies is that ambient ozone monitors are measuring something that might be associated very weakly with morbidity, mortality, and other phenomena such as emergency department visits, hospital admissions, and school absences. EPA staff assume that ambient ozone concentrations are functionally equivalent to personal exposure and that the observed weak associations are causal, but they cannot reach either inference based on science.

EPA staff go even further to assert that the fact personal exposures are tenfold or more lower than ambient levels supports the assumption that the health risks posed by ozone are underestimated. There are two problems with this claim. First, it doesn't matter what their ratio is if they are uncorrelated. Second, it turns algebra on its head. In controlled human studies, subjects were subjected to personal – not ambient – concentrations of 0.06 ppm. EPA acknowledges that ambient ozone concentrations are 2- to 4-fold greater than personal exposure. That means the 1.5% group mean FEV_1 decrement observed by Adams (2002, 2006a) at 0.06 ppm personal exposure corresponds to an ambient ozone concentration of 0.12 to 0.28 ppm.

12. *EPA disseminates risk characterizations based on studies using research methods Agency staff have rejected as unreliable and invalid for other air pollutants.*

Several of the epidemiological studies EPA staff rely upon in the ozone NAAQS use personal expiratory flow rate devices. For ozone, EPA staff say these methods obtain valid and reliable data. However, in their draft Integrated Science Assessment for Oxides of Nitrogen (NO_x), EPA staff say these devices are known to yield unreliable and invalid data. In the ozone review, EPA staff highlight Mortimer et al. (2002) as especially relevant for the estimation of health risks to asthmatics. In the NO_x review, EPA staff highlight Mortimer et al. (2002) as an example of a study whose methods are irredeemably deficient.

The Iron Law of EPA Staff Ozone Health Risk Assessment and Characterization clearly are at work here. Mortimer et al. (2002) reported statistically significant positive associations between ambient ozone and reported symptoms, but no association between those same symptoms and NO_x. Hence, EPA staff use Mortimer et al. (2002) to push out the ozone risk envelope, and they discard it rather than use it to push the NO_x risk envelope inward.

13. EPA never informed CASAC about information quality principles or the Agency's Information Quality Guidelines, and information quality played no role in CASAC's review even though it is required by EPA's Peer Review Handbook.

By law, CASAC has the unenviable task of simultaneously providing both an objective review of the scientific database and policy advice to the Administrator that, by its very nature, cannot be objective. EPA could have made this task easier if it had asked CASAC to clearly distinguish its scientific review from its policy advice. Indeed, information quality review is an explicit component of EPA's Peer Review Handbook. It is essential for peer reviews to fully address information quality principles in order to secure the rebuttable presumption of objectivity that the Information Quality Guidelines provides.

However, EPA didn't inform CASAC about information quality, nor did the Agency's Charge to CASAC even mention the subject. Consequently, the record of the CASAC review shows that the panel wove its scientific review and policy advice into a single fabric and, predictably, paid no attention to information quality. It is infeasible for CASAC's review to have fully addressed information quality principles when it devoted no time at all to information quality issues.

E. Remedies Requested

The evidence for systemic information quality error by EPA staff is overwhelming, both procedurally and substantively. The difficulty EPA staff face is that a candid, accurate and forthright response to this RFR may undermine the Agency's ability to legally defend the Administrator's recent decision. We have tried throughout this process to stay clear of the Administrator's exercise of discretion, as provided for by the Clean Air Act, but we agree that EPA's ability to defend is highly compromised by the existence of systemic information quality error that rendered inaccurate the scientific database on which the Administrator relied. Moreover, the Clean Air Act imposes on EPA the onerous duty to revise each NAAQS every five years despite the fact that it takes about that long to conduct each review. That means EPA is engaged in a never-ending cycle that impedes it from implementing the process reforms necessary to comply with its own information quality policies and guidelines.

Senior EPA officials are obligated under Section 8.6 of the Agency's Information Quality Guidelines to perform an independent review of this RFR and provide a well-documented and comprehensive response to each information quality error that we continue to allege. In some cases, it may not be possible for EPA to repair the error we have identified in the current NAAQS

review cycle. In no case, however, should the inability to make a timely repair justify continued misrepresentation of the information in the scientific database.

In addition, we specifically request that EPA make significant changes to its NAAQS review process so that future information quality errors are rare instead of systematic.

1. *Explicitly, comprehensively, and consistently implement information quality principles and practices throughout the NAAQS review process and within each NAAQS work product.*

The record shows that EPA ignored information quality principles and the Agency's own policies and procedures throughout the ozone NAAQS review. Agency officials must completely overhaul the NAAQS review process to explicitly, comprehensively, and consistently implement information quality. Token efforts, changes made only at the fringe of the process, reforms that shift EPA's burden to the public, and the addition of new legalistic boilerplate are all unacceptable.

2. *Establish an external independent body with the limited responsibility of reviewing the quality of scientific and technical information, and advising the Administrator whether information quality principles have been met and applicable information quality policies and practices have been followed.*

To prevent information quality error, we recommend that the Administrator establish an external and independent Information Quality Review Committee for the express and limited purpose of advising whether information quality principles have been met and information quality policies and practices have been followed. The Committee would supplement, not supplant, the scientific and policy review currently performed by CASAC.

The Committee would not be asked to make policy recommendations or opine on what the science means, both of which are statutory functions currently assigned to CASAC, but to perform information quality reviews that are significantly different from CASAC's current activities. These functions cannot be performed by CASAC because its members generally lack expertise in information quality, and it may be unreasonable to expect them to have both information quality and subject matter expertise. Moreover, they are pressed for time to satisfactorily accomplish their current assignments.

The work of the Information Quality Review Committee should be performed in public subject to the requirements of the Federal Advisory Committee Act. Committee members must be independent of EPA and unaffiliated with the research teams whose scientific work products the Agency

relies on for risk assessment and characterization. It is impossible for any individual or panel of peer reviewers to independently examine the quality attributes of research that they personally performed or that was performed by their institutional colleagues.

3. *Remove all policy judgments and similar considerations from the assembly, review and presentation of scientific information in all EPA NAAQS work products.*

The record shows convincingly that EPA staff routinely attempt (and in the ozone NAAQS review, appear to have succeeded) in restricting the authority delegated to the Administrator by the Clean Air Act through the device of providing the Administrator with a summary of the scientific database that reflects EPA staff views about decisions they believe the Administrator ought to make. This can only be overcome if EPA officials explicitly and forcefully direct the staff to refrain from embedding policy judgments in these work products and instead provide the Administrator with a genuinely objective scientific record. Instructions to CASAC – the "Charge" – also must be modified so that the committee is explicitly and formally directed to clearly distinguish its scientific review from its provision of policy advice.

The Information Quality Review Committee should be tasked with determining whether this directive has been met, and if it hasn't, informing the Administrator where residual policy judgments reside. Relentless effort is needed to ensure that science and policy are clearly distinguished in all EPA NAAQS work products.

4. *Establish new and publicly accountable pre-dissemination review procedures for all EPA NAAQS work products.*

The record shows that despite the existence of an Agency requirement for pre-dissemination review that was established in 2002, no such review ever took place in the current ozone NAAQS review. It is not credible to believe that staff were unaware of information quality principles and Agency guidelines.

EPA officials should rectify this apparent loophole by explicitly establishing a comprehensive program of pre-dissemination review of the information used in NAAQS reviews. These activities can be conducted in parallel with NAAQS regulatory development, and the products of pre-dissemination review can be examined by the Information Quality Review Committee to ensure that they have actually achieved the goals EPA set forth in its Information Quality Guidelines.

5. *Establish an information quality foundation for all CASAC reviews of NAAQS-related work products whose successful performance is documented and independently validated.*

The record shows that EPA did not inform CASAC about information quality principles or the Agency's policies and practices that were established to achieve adherence with these principles. By leaving information quality out of the Charge, never educating CASAC about its meaning and implications, and conducting a multiyear dialogue with CASAC that never broached the subject, EPA staff ensured that CASAC could not and would not ever take account of information quality concerns in its scientific review. The extent to which this failure distorted the committee's policy recommendations cannot be ascertained, but it should be assumed that the committee would have offered different policy advice if information quality had been central to its review of the science.

There is no question that reports prepared by the Information Quality Review Committee would be very useful to CASAC. Committee reviews should be scheduled early and often so that when CASAC convenes to review a NAAQS work product, it has at its disposal a thorough and objective review of the information quality attributes of the scientific information it must examine.

6. *Require that CASAC panel members be recused from the review of their own research or the research of their institutional colleagues, and such individuals should not serve on a CASAC panel in cases where such research is crucial to risk assessment and characterization.*

The CASAC ozone panel, like previous CASAC panels, included a number of members whose primary research interests and activities involve ozone. In some cases, CASAC panel members are the authors or co-authors of studies relevant to the assessment and characterization of health risk, or they are institutional colleagues of such researchers. It is vital that these individuals participate extensively in that portion of the CASAC process which consists of assembling and summarizing scientific data. However, an intellectual conflict of interest arises when these scientists are also asked to review EPA staff work products that interpret their research or judge its quality. It is unreasonable to expect intellectually conflicted CASAC panel members to provide unbiased opinions of EPA staff work products, and the opinions of such panel members cannot reasonably be assumed to be free of self-regard.

In some cases, mere recusal from a portion of CASAC review is not sufficient. The group dynamic of peer review inhibits those panel members who are not conflicted from candidly expressing serious concerns and doubts. Scientists who are authors, co-authors, or institutional colleagues of the handful of scientific studies that are identifiable as crucial should not serve on CASAC at

all. It also would help if EPA did a better job applying its own peer reviewer selection rules to ensure experts selected to serve had open minds. Before the review even began, several members of the CASAC ozone panel were on record supporting major reductions in the ozone NAAQS.

II. Introduction

This RFR is submitted to EPA in accordance with government-wide requirements related to information quality (Information Quality Act 2000; Office of Management and Budget 2002) 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). Pursuant to these Guidelines, a copy of our RFC is attached. EPA Guidelines recommend that RFRs be submitted within 90 days, but we found that the magnitude of the task was too complex to complete in such a short window.[5]

A. Information Subject to this Request for Reconsideration

The RFC set forth a list of documents that constituted the information subject to the petition. This list included the 3-volume Criteria Document (U.S. Environmental Protection Agency 2006a, 2006b, 2006c), the Staff Paper (U.S. Environmental Protection Agency 2007g, 2007j, 2007k, 2007l), exposure and risk assessments (U.S. Environmental Protection Agency 2007d, 2007e), certain internal memoranda (Brown 2007a; Langstaff 2007), and the preamble to the notice of proposed rulemaking (U.S. Environmental Protection Agency 2007h). All of these documents include influential scientific, technical, statistical and economic information that is subject to the Information Quality Act (Information Quality Act 2000), OMB's government-wide guidelines (Office of Management and Budget 2002), and EPA's agency-specific guidance (U.S. Environmental Protection Agency 2002).

[5] EPA's Information Quality Guidelines are subordinate to the government-wide guidelines issued by the Office of Management and Budget (2002). OMB's guidelines authorize agencies to establish "appropriate time limits in which to resolve such requests for reconsideration," taking account of whether "other agencies may have an interest in the resolution of any administrative appeal " (pp. 8458 and 8459). OMB's guidelines do not authorize agencies to impose artificial deadlines on the submission of such appeals.

In our RFC we also listed EPA's Regulatory Impact Analysis (U.S. Environmental Protection Agency 2007i) as a covered document. However, given the very limited amount of time available between the date it was published (August 2, 2007) and EPA's requirement that a timely RFC meet the deadline for public comments on the proposed rule (October 9, 2007), we were compelled to set priorities and defer this matter until a later date. The RIA does not appear to have been distributed for notice and comment – it is not part of the standard-setting process and we could not locate a relevant Federal Register notice requesting public comment -- so it is not covered by this part of EPA's Information Quality Guidelines, which required that we submit our RFC on or before the deadline for public comments (U.S. Environmental Protection Agency 2002, Section 8.5). Moreover, because the RIA was not a factor in the Administrator's final decision, there is no deadline for timely submission of an RFC with respect to the RIA. Nonetheless, the RIA incorporates scientific information from the 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 or derivative information quality errors.

Our RFC concerned influential scientific, technical, and statistical information contained or referenced in these documents. It did 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).[6] Likewise, this RFR also concerns information and not expressions of values or opinion.

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 represents a sector that employs more than 14 million American workers. The NAM's mission is to enhance the competitiveness of manufacturers and

[6] "'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.

improve American living 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. EPA's Information Quality Guidelines (U.S. Environmental Protection Agency 2002) follow the OMB Guidelines in most material respects. We followed EPA's agency-specific procedures for affected parties in submitting our RFC (Section 8); in particular, we simultaneously submitted the RFC as a public comment on the Notice of Proposed Rulemaking (Section 8.5, page 32).

Independent appeal provisions are set forth in Sections 8.6 and 8.7 of EPA's Information Quality Guidelines. A three-member executive panel consisting of the Science Advisor/AA for the Office of Research and Development (ORD), the Chief Information Officer/AA for OEI, and the Economics Advisor/AA for the Office of Policy, Economics and Innovation (OPEI) normally would investigate the claims decide the appeal after presentation of the issues by the "information owner." In this case, there are two "information owners": the Assistant Administrator for Air and Radiation (OAR) and the Science Advisor/AA for ORD. Because information owners must be recused from the appeal process for it to be plausibly independent, EPA's Science Advisor/AA for ORD cannot serve on the executive panel and must be replaced with an Assistant Administrator other than the AA/OAR, or a Regional Administrator. EPA is required to conduct appeals in a timely manner, and the Agency has decided that 90 days meets this requirement.

D. Relevant Information Quality Principles

Each of the documents that was designated a subject of the RFC is *influential*, as that term is defined in both OMB's and EPA's guidelines. 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 our RFC, we claimed that information within the listed documents did not satisfy the information quality principles of objectivity (both subspecies). In particular, information about ozone health risk is neither substantively 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, estimates of costs, risks and benefits also are not substantively objective.

We continue to assert that most of the scientific, statistical, and technical information that we challenged via the RFC does not adhere to the information quality standard of objectivity.

2. Failure to adhere to the utility standard

We claimed that because of these systemic and material defects in objectivity, the documents subject to the RFC did not satisfy the *utility* standard. *Utility* requires that information that is disseminated be useful for the purpose to which it was intended. In the case of the RIA, the purpose of the document was 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 rendered the Agency's risk and 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 do not satisfy the substantive objectivity standard, and by failing that standard they can not have utility for their intended purpose of informing the public about the impacts of a revised ozone NAAQS.

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 Criteria Document and Staff Paper 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 was 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. However, this scientific information was fundamentally flawed because it systematically violated the objectivity standards. For that reason, the Administrator's reasoned basis for decision-making almost certainly relies on inaccurate scientific information. In the RFC, petitioners did 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 Guidelines. Rather, we challenged the scientific and statistical information provided to the Administrator. Nonetheless, it is at least plausible and perhaps highly likely that the Administrator's decision would have been different if he had been provided scientific and statistical information that adhered to applicable information quality principles.

We continue to assert that EPA's failure to adhere to the objectivity standards means that the information still being challenged does not adhere to the utility standard.

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

In our RFC, we noted "[b]ias takes many forms" that "affect the scientific information upon which EPA relies" and "how EPA chooses to utilize this information." We also noted that

> [b]ias 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 (National Association of Manufacturers 2007, pp. 9-10, emphasis in original).

We alleged that EPA began with a database that was structurally biased by (1) control of development and publication by parties with well-defined risk management objectives; (2) multiple forms of publication bias; (3) systemic methodological error; and (4) peer review in which the assurance and maximization of information quality played no part whatsoever. These problems were severely exacerbated by EPA staff's last-minute submission to the scientific record a critical reanalysis that had never been publicly disclosed, peer reviewed, or subjected to any of the normal procedures of scientific review and validation.

Although 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 a persuasive showing that the information is not in fact objective, or that the peer review on which the presumption rests was deficient in a material respect relevant to information quality. For example, to show that peer review did not assure even adequate objectivity, it is sufficient to show that in their charge the peer reviewers were not asked to evaluate whether the information satisfied applicable information quality principles, or if they were so asked that they failed to fulfill their charge. An agency cannot evade its responsibilities under the Information Quality Act by waving peer review as a talisman or wearing it as an institutional phylactery. For peer review to serve its purposes under information quality guidelines it cannot do so accidentally; it must rigorously apply information quality principles.

Federal agencies, not petitioners for correction or independent research scientists, are subject to the strictures of information quality. This system is not designed to set up the perfect as the enemy of the good, but to deter the government from abusing its unique powers and responsibilities. The scheme granting the government a presumption of adequate objectivity for information that has been peer reviewed has three important and desirable features: it establishes a low hurdle that excludes scientific information that has not endured the minimally invasive rigors of professional scrutiny; it creates an incentive for more-objective information to supplant less-objective information at every reasonable opportunity, thus fostering scientific advancement; and it rewards peer review procedures that explicitly and rigorously evaluate adherence to information quality principles, most notably, the principles of presentational and substantive objectivity. An agency cannot justify a preference for less-objective information because it more conveniently conforms to a policy mission or the risk management preferences of staff or management. To succeed in rebuttal, a petitioner need only show that the information cited approvingly by the agency is materially lacking in objectivity. The petitioner need not show that he "knows" the right answer or that he can point to alternative scientific information that is provably unbiased. Agencies, in short, are not allowed to use bad information

just because it is all they have, or to reject better information because it is not perfect.

A. EPA Begins with a Structurally Biased Scientific Database

In our RFC, we stated that influential scientific information provided to EPA may be biased for several reasons. We discussed three such reasons.

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

We noted that health-effects studies have been funded by government, industry, and sometimes jointly. Because industry has well-defined policy interests, it is often suspected or accused of trying to control research so that it yields desired results. EPA is not shy about highlighting such potential conflicts of interest. In its Response to Comments, for example, four times EPA implies that a study or review might be technically unsound because it was industry-funded.[7] In none of these cases, however, does EPA actually provide evidence suggesting how scientific integrity was in fact compromised. Rather, EPA's approach consists of suggestive condemnation by association – by itself, a material breach of applicable information quality standards because it involves the attempted elucidation of quality distinctions based on criteria other than quality.

Nongovernmental organizations and government agencies like EPA also have well-defined policy interests, and thus they have similar incentives to control research to ensure agreeable outcomes. In EPA's Response to Comments, there is no instance in which EPA implies that a study or review might be technically unsound because it was NGO- or government-funded. Indeed, the list of studies EPA heavily relies on that the Agency itself funded is an extensive one.[8] Nowhere in the Response to Comments, however, does EPA ever

[7] See EPA (2008d, p. 45): p. 5 (Brauer et al. 2007), pp. 21 and 97 (Adams 2006a), and p. 22 (Smith 2007b). EPA praises itself for being "a leader" in examining the so-called "GAM problem," having "funded a special workshop and supported the [Health Effects Institute] in a project to reanalyze dozens of studies to fully investigate this issue" (p. 45). EPA does not mention any other sources of funding, nor does it acknowledge the consequences the Agency would have suffered had it refused to participate.

[8] The list of critical studies funded by EPA but not identified as such in EPA's Response to Comments includes mortality epidemiology (Bell et al. 2005; Bell, McDemott et al. 2004; Bell et al. 2006; Levy et al. 2005); morbidity epidemiology (Korrick et al. 1998; Mortimer et al. 2002; Ross et al. 2002); school absence epidemiology (Chen et al. 2000; Gilliland et al. 2001); and controlled human exposure (McDonnell 1996).

acknowledge the Agency's role or imply that EPA funding might have infiltrated the studies' designs, implementation, results or reporting.

It is because of this asymmetry in EPA's treatment of scientific information that industry routinely funds research through arm's length grants and contracts that insulate researchers from sponsor interference. The extent to which NGOs and government agencies do so is not well documented.[9]

Federal information quality guidelines deal with the problem of sponsor bias two ways. First, they place a high value on full disclosure sufficient to ensure reproducibility. Reproducibility is widely believed to be the best procedural tool for determining whether interference occurred. When a research sponsor declines to make its data available, that which it does disclose may become 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 because they would be detectable.

Many times in EPA's Response to Comments, the Agency attempts to deflect scientific questions raised by many commenters based on the policy preferences of the commenter rather that the scientific merit of the comment.[10] This tactic creates the false perception that the warring interests on both sides are motivated solely by policy disputes and only EPA is motivated by the pursuit of science. This practice is false – EPA staff, managers, and officials all have policy views – and it is anathema to good government because it unfairly stigmatizes the scientific integrity of public commenters generally.

Information quality guidelines require that influential information that an agency proposes to disseminate be capable of being reproduced by a competent and independent third party. This transparency requirement is a precursor step

---

[9] When scientists perform unquestionably independent research funded by industry, it is often then alleged that they skew their work to ensure a steady stream of future research grants. Such allegations can never be disproved because they are not testable. In any case, the identical claim can be made about NGO and government-funded research programs, such as for example EPA's STAR grant program.

[10] For examples in which EPA deflected critical scientific (not policy) comments on the ground that the commenter opposed revising the NAAQS, see, e.g., pp. 12, 14, 15, 19, 30, 37, 40, 58, 75, 77, 104, and 128. For examples in which EPA deflected critical scientific (not policy) comments on the ground that the commenter supported a revised NAAQS below the value selected by the Administrator, see, e.g., pp. 10, 11, 12, 14, 15, 55, 56, 104, and 105.

in the assurance of presentational and substantive objectivity, for it is by checking the government's work that departures from objectivity are most readily detected. If the government could withhold information necessary to enable reproducibility, it could obstruct the public's ability to exercise its legal right to objective information.

In the ozone case, EPA asked for and promptly obtained from Prof. William C. Adams data from several of his controlled human studies.[11] EPA was then able to reproduce a subset of Adams' results, and even to perform a reanalysis of his data to partially test his work for objectivity.[12]

In contrast, EPA or an allied federal agency funded virtually all of the epidemiological studies that Agency staff consider highly influential for estimating human health risks. Moreover, EPA has by law the right to obtain data from researchers who perform Agency-funded research (Office of Management and Budget 1999, Sec. 36). However, we can discern no instance in the ozone review in which EPA has exercised this right. For government-funded research, the Agency's staff is generally satisfied that the information provided in published papers is full and complete.[13]

---

[11] Note that EPA has a pattern of requesting and obtaining data from industry-funded studies: "As in the 1997 risk assessment, EPA obtained individual data from several 6.6-hour O_3 controlled human exposure studies from [Adams]. API, the funding sponsor of the Adams studies, urged EPA to use the data from these studies, particularly the most recent study by Dr. Adams in its health risk assessment in its comments on the draft Staff Paper and draft health risk assessment in January 2006. EPA obtained the individual data used in the health risk assessment directly from the author and explained that the data would be combined with other individual data from the Horstman, Folinsbee, and McDonnell 6.6-hour O_3 studies" (U.S. Environmental Protection Agency 2008d, pp. 97-98, emphasis added, internal references omitted).

[12] EPA requested and obtained only selected data, and proceeded to analyze only this subset. It did not seek to reproduce Adams' analysis. It is noteworthy that in this reanalysis, EPA does not claim that Adams' data or his statistical analysis departed in any manner from the information quality standard of objectivity.

[13] The record shows several cases in which, when questions arose concerning details not reported in refereed articles resulting from EPA-sponsored research, EPA staff simply requested analytic results not included in the published papers from these researchers and cited them as "personal communications" (U.S. Environmental Protection Agency 2006a, pp. 7-179, 177-185, 178-183; 2007g, p. 3-93). Additional data or analytic results provided to EPA via "personal communications" cannot be reproduced by independent third parties, and thus are inherent violations of applicable information quality standards.

We said in our RFC that EPA staff analyzed the scientific record with a policy-driven bias in favor of discovering risk. EPA denies this, but the fact that it regarded its own funded research as inherently trustworthy and industry-funded research as presumptively biased is prima facie evidence that our allegation is in fact correct. A necessary condition for the absence of interpretative bias is persuasive evidence that EPA had in place, and actually followed, a plan for pre-dissemination review that applied the same information quality standards for the review of all scientific information irrespective of how it was funded. Not only did EPA fail to follow such a plan, its Response to Comments reveals that it didn't even have a plan to follow.[14]

A practical consequence of EPA's managerial control over both the scientific record and the development of policy alternatives is the Agency staff appears to have been unable to prevent its policy preferences from influencing its presentation and review of the scientific record. The evidence is overwhelming that these conflicting missions resulted in a systemically biased characterization of the human health effects of ozone.

2. *Publication bias*

In our RFC, we identified three subspecies of publication bias that we believe are present in EPA's scientific record: positive-results bias, outcome-reporting bias, and inferential exaggeration (National Association of Manufacturers 2007, pp. 11-13). Positive-results bias occurs because studies that do not show positive associations are published less frequently, if at all. Outcome-reporting bias occurs when researchers report results with the highest apparent association, a widely observed phenomenon. Sometimes, dozens of models will have been examined but only the handful with the strongest association will be reported (Lumley and Sheppard 2003). Inferential exaggeration occurs when scientists draw (and editors accept) conclusions that are not supported by the data and analysis actually performed.

We noted in our RFC (and EPA did not dispute in its Response to Comments) that positive results bias and outcome-reporting bias are endemic in epidemiological literature. EPA's Response to Comments (U.S. Environmental Protection Agency 2008e, pp. 31-32) says Agency staff "recognized the potential impact of publication bias" in Section 7.1.3.6 of the Criteria Document. EPA staff

[14] See EPA (2008e, p. 150): "EPA's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated does not require the Agency to discuss, separately, whether the pre-dissemination review actually occurred."

acknowledged the problem of publication bias,[15] set forth a consistent method for addressing the portion of publication bias caused by multiple hypothesis testing,[16] then abandoned this method because it was infeasible and retreated to the default assumption that the problem did not exist.[17] EPA staff also said they would emphasize zero and 1-day lags[18] in tine-series studies and give more weight to primary over secondary analyses.[19] In the Staff Paper, however, EPA staff did exactly the opposite. They relied on the distributed lag models of Mortimer et al. (2002) and Bell, McDermott et al. (2004), defended this reliance in the Response to Comments (U.S. Environmental Protection Agency 2008e, e.g., pp. 32-33, 43-44, 45-48), and made little or no distinction between primary and sensitivity analyses.

Outcome-reporting bias can be documented by comparing the protocols that researchers established prior to beginning work against the analyses that they actually published. Inferential exaggeration can be detected by carefully comparing study results against researchers' inferences to ensure that inferences do not reach beyond what the data and analyses support. To test for outcome-reporting bias, however, EPA staff would have to obtain researchers' data, made the effort to reproduce their work, and explored what alternative results could have been reported using the same data and alternative methods. EPA obtained research data only from industry-sponsored researchers, so it was able to examine outcome-reporting bias only for this subset of research projects. One interpretation of EPA's reanalysis of the Adams (2006) data – Brown (2006, 2007a, 2007b) -- is that Agency staff sought to uncover outcome-reporting bias in the form of statistically significant associations not reported by Adams. There is no

---

[15] "The summary of health effects in this chapter is vulnerable to the errors of publication bias and multiple testing."

[16] "To address multiple hypothesis testing, emphasis will be placed in this chapter on a priori hypotheses."

[17] "As identifying a priori hypotheses is difficult in the majority of the studies, the most common hypotheses will be considered."

[18] "For example, although many studies examined multiple single-day lag models, priority would be given to effects observed at 0- or 1-day lags rather than at longer lags."

[19] "Analyses of multiple model specifications for adjustment of temporal or meteorological trends will be considered sensitivity analyses. Sensitivity analyses shall not be granted the same inferential weight as the original hypothesis-driven analysis…"

public evidence that EPA made any effort at all to uncover outcome-reporting bias among those researchers whose work the Agency itself sponsored.

In our RFC, we had no all-purpose remedy for the problem of positive-results bias because we recognized that it was so hard to detect. With respect to outcome-reporting bias, however, we said

> [f]or each critical study, EPA should determine the extent to which nonpositive outcomes were not reported and include that information in its presentation (National Association of Manufacturers 2007, pp. 12-13).

In its response, EPA gives three reasons for denying our remedy request:

> EPA rejects NAM's contention that it should determine the extent to which nonpositive outcomes were not reported and include that information in its presentation of each critical study. First, there is no evidence to show that researchers are not reporting all results. Second, EPA can not include in its assessment results that were not reported. Third, EPA uses a weight of evidence approach to evaluate evidence that does not depend on a few critical studies (U.S. Environmental Protection Agency 2008e, p. 32).

Each of these reasons is logically irrelevant but revealing.

Addressing these arguments in order, it is generally agreed that the absence of evidence is not the evidence of absence. In this case, the most obvious reason why "there is no evidence to show that researchers are not reporting all results" is that EPA has not looked – except in the case of industry-funded researchers, where it has found none.

Second, EPA is not prohibited from supplementing the scientific record with data and analyses beyond that which was published by original authors, provided that the Agency makes the new data and analyses transparent and reproducible, and subjects them to effective peer review. Indeed, the EPA staff practice of constraining its review to published results encourages and intensifies outcome–reporting bias.

Moreover, if it is true that "EPA can not include in its assessment results that were not reported" in published studies, then it must discard several of its internal memoranda because they rely on unpublished data or nontransparent syntheses of published data.[20] EPA also must discard the unpublished results it

[20] There are many prominent internal memoranda that meet this definition (Cox and Camalier 2006; Langstaff 2006a, 2007; McCluney 2007; McCluney et al. 2006; Rizzo 2005, 2006).

has obtained from Agency-funded researchers.[21] EPA cannot simultaneously rely on unpublished results that it has specifically requested and claim that it is not allowed to rely on unpublished results.

Third, EPA's "weight-of-evidence approach" is inherently noncompliant with EPA's own Information Quality Guidelines. Without researchers "showing their work," these results are not "capable of being substantially reproduced" (U.S. Environmental Protection Agency 2002, p. 47).

In its response, EPA mischaracterizes our complaint as saying that "EPA has not considered publication bias," then proceeds to rebut its mischaracterization rather than our complaint (U.S. Environmental Protection Agency 2008d, p. 31, emphasis added). We made no such claim; our concern was that while EPA may have "considered" it, the Agency's efforts to estimate its magnitude consisted of meta-analyses primarily intended to address uncertainty about the magnitude of effect estimates across studies and strengthen statistical significance.[22]

(a) Outcome-reporting bias

We specifically identified several panel studies on which EPA relies as displaying evidence suggestive of outcome-reporting bias (Gent et al. 2003;

---

[21] Expanding on footnote 13, in the Criteria Document and Staff Paper EPA references unpublished mortality estimates obtained from Bell ("Bell, M. L. (2006) Community-specific maximum likelihood estimates of O_3-related excess risk in mortality for the NMMAPS U.S. 95 communities study [personal communication with attachments to Jee Young Kim]. New Haven, CT: Yale University School of Forestry and Environmental Studies; January 6." (U.S. Environmental Protection Agency 2006a, p. 7-179; 2007m, p. 3-93) and Ito ("Ito, K. (2004) Revised ozone risk estimates for daily mortality and hospitalizations in Detroit, Michigan [personal communication with attachments to Jee Young Kim]. New York, NY: New York University School of Medicine, Nelson Institute of Environmental Medicine; October 31.") (U.S. Environmental Protection Agency 2006a, p. 7-185). EPA reprints the unpublished estimates from Bell in Figure 7-17 of the Criteria Document but not in the Staff Paper. EPA cites the unpublished results obtained from Ito on pages 7-76 and 7-82 of the Criteria Document.

Outputs of statistical analyses delivered by researchers to EPA via personal communication inherently violate EPA's Information Quality Guidelines because they are not reproducible.

[22] In Section I.D.1 on page 14, this phenomenon is first on our list of major types of information quality error. Whether EPA has "discussed" or "considered" something is immaterial; what matters is what the Agency disseminates as authoritative.

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.

In its Response to Comments, EPA says "there is no evidence to show that researchers are not reporting all results" (U.S. Environmental Protection Agency 2008e, p. 32). As we noted earlier, EPA's lack of evidence is assured by the staff's having not inquired. Gent et al. (2003) is very tightly written to accommodate the journal's notoriously severe space constraints, but nevertheless it discloses that that a variety of models were examined and not every result was reported. Korrick et al. (1998) acknowledge reporting results only from reduced-form models in which variables that *a priori* they considered important were dropped due to lack of statistical significance. Mortimer et al. (2002) acknowledge examining a wide array of lagged exposure models, and they imply that they analyzed morning PEF values because evening values, which they expected to be diminished from the day's ozone exposure, were not. More generally, it is a ubiquitous practice for journals to publish a (usually small) subset of the analytic work actually performed. EPA's position that "there is no evidence" that what's reported is incomplete is impossible to credit.

(b) Inferential exaggeration

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.

In our RFC, we cited Gent et al. (2003) and Mortimer et al. (2002) as examples of refereed papers in which notable inferential exaggeration was present (National Association of Manufacturers 2007, p. 13). 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. We said that this conclusion went beyond what could be inferred from the reported data and analysis. The language alone is laden with policy judgment; the phrase has no objective scientific meaning.

Mortimer et al. (2002) also concluded that ozone below current standards has adverse effects on asthmatic children (p. 704). In our RFC, however, we said

their conclusion was based on a selective reporting of model results (Id.). The authors performed analyses using an exhaustive set of lag models[23] and reported results from a subset of these analyses, then discussed as representative and meaningful only those results that yielded statistically significant positive effects.

In its response, EPA does not actually rebut our claim that Gent et al. (2003) and Mortimer et al. (2002) engage in inferential exaggeration. With regard to Gent et al. (2003), EPA says only that "there is no reason to believe that the peak O_3 concentrations drive the findings of this study" (U.S. Environmental Protection Agency 2008d, p. 32). EPA ignores both the fact that the study area was in nonattainment and the authors' own statements about their work. The authors say it is a strength of their study design that it "use[d] both the maximum 1-hour average (sensitive to spikes in concentration) and 8-hour average (a measure of short-term, cumulative exposure) to assess daily ambient ozone levels" (p. 1865, emphasis added). In other words, the Gent et al. (2003) study was designed to capture peaks, and to the extent that the study uncovered a few weak concentration-response trends it was peaks that gave these trends away.[24]

(i) The inferences made by Gent et al. (2003) are inconsistent with their own reported results

Examples of significant inferential exaggeration in Gent et al. (2003) are not hard to find:

- "A 4% increase in bronchodilator use was … associated with same-day levels of ozone (51.6-58.8 ppb) (Table 4, model 1)" (p. 1862). The authors do not mention that this odds ratio is barely statistically significant; the 95% confidence interval is 1.00 – 1.09. Also, they do not mention that it is the only statistically significant positive odds ratio among 20 odds ratios reported for five ozone concentration quintiles compared across two exposure scenarios (previous day, same day) and two averaging times (1-hour, 8-hour). There is no statistically significant positive trend for either exposure scenario or either averaging time (range of *p* values: 0.13 to 0.64), a fact they also do not mention. Finally, these results pertain to single-pollutant models,

[23] "Lagged air pollution effects were evaluated using moving averages, unrestricted distributed lags, and polynomial distributed lags." See Mortimer et al. (2002, p. 700).

[24] EPA has said that triangular exposures are more representative of real-world conditions (U.S. Environmental Protection Agency 2007g, p. 6-10).

which would be expected to overstate the effect of ozone if other pollutants (such as fine PM) are positively correlated.

- "In logistic regression models of both ozone and fine particles for children taking maintenance medication, an increased likelihood of respiratory symptoms was associated with levels of ozone on the same day, previous day, or both; and increased bronchodilator use was associated with the highest level of same-day ozone" (p. 1863). Consistently positive trends in same-day 1-hour effects are shown for wheeze and shortness of breath; for chest tightness only in the highest two quintiles, and not at all for persistent cough. Previous day 1-hour effects are statistically significant only for chest tightness. A statistically significant increase in bronchodilator use was reported only for the highest exposure quintile; this increase was reported for only one of ten odds ratios. The authors glossed over these details, focusing instead on the handful of statistically significant effects.
- "In logistic regression models of both ozone and fine particles for children taking maintenance medication, an increased likelihood of respiratory symptoms was associated with levels of ozone on the same day, previous day, or both…" (p. 1864). Their Table 5, however, shows significant positive same-day effects for only two of five symptoms for the 1-hour averaging time, and significantly positive previous-day effects for only one of five symptoms for the 1-hour averaging time. Two of five endpoints displayed positive trends for previous-day 8-hour averaging times; none of the five displayed positive trends for same day 8-hour exposures.
- "In models controlling for ambient fine particle concentration and typically at levels below EPA air quality standards, daily ambient ozone was found to be significantly associated with increased risk of respiratory symptoms and increased use of rescue medication among children with asthma severe enough to require maintenance medication" (p. 1865). Their Table 5, however, shows significant positive trends only for two of five symptoms for same-day 1-hour averaging times, and significant positive trends only for two of five symptoms for previous-day 1-hour averaging times.

Single-pollutant models yield different results than co-pollutant models where exposure to co-pollutants is highly correlated. In the single-pollutant models, nine of the 80 ozone-related odds ratios reported in Table 4 for asthmatic children on medication are statistically significant. In all cases, it is the previous day's ozone for which the positive trend is statistically significant. No same-day

trends were observed. In the co-pollutant models (Table 5), there are 80 odds-ratios reported and only seven are statistically significant. Two of the three consistently positive trends appear in same-day ozone exposures. This pattern might make sense, but Gent et al. (2003) do not try to explain why. Instead, they reach conclusions identical to their premises but not supported by their own data and analysis.

In our RFC (p. 14), we faulted the EPA staff's reliance on Gent et al. (2003) in part because the authors relied heavily on self-reported symptoms. In its response, EPA says our claim is factually "incorrect." The Agency writes, "In addition to respiratory symptoms, Gent et al. also observed an association between O_3 and rescue medication use, which is an objective measure" ((U.S. Environmental Protection Agency 2008d, p. 48).

EPA is correct that bronchodilator use is a more objective indication that symptoms have led subjects to take action, and that a decision to take action is a better threshold indicator of potential adversity than are changes in FEV_1 or FVC too small for subjects to even notice. However, the results reported by Gent et al. (2003) are inconsistent with their reported (and oft-repeated) claim that they actually found an association between O_3 concentration and bronchodilator use. Their Table 4 shows odds ratios for each quintile (except the baseline) for 1- and 8-hour exposures to previous- and same-day concentrations: four single-pollutant analyses containing 16 odds ratios. Of these 16 odds ratios, one is weak and marginally statistically significant (OR, 1.04; CI, 1.00-1.09). It is the middle exposure quintile for same-day 1-hour exposures; there is no theory we know that would predict this result. We can even discard statistical significance for the individual odds ratios and look for positive trends. Tests performed by Gent et al. (2003) are show a range of *p* values for positive trends ranging from 0.13 to 0.64.

The claim by Gent et al. (2003) that they found an association between ozone concentration and bronchodilator use is an excellent example of inferential exaggeration, even if it's one that EPA has pointed out to us rather than *vice versa*. Indeed, EPA staff seem to agree that the association claimed by Gent and coworkers does not actually exist, for they have abandoned epidemiology in favor of a policy-driven constraint that defines any increased symptoms, whether dose-related or not, as evidence of an ozone effect unless they can be proved to have another cause:

> Regardless [of what Gent et al. (2003) actually found], EPA deems respiratory symptoms to be a valuable health outcome, which considered in conjunction with various other more "objective" measures, allows a

> more complete depiction of the potential respiratory health effects of pollutants (U.S. Environmental Protection Agency 2008d, p. 48):

Inferential exaggeration is a tool for disguising researchers' policy views under a cloak of science. Thus, when Gent et al. (2003) say that "current standards do not protect these more vulnerable members of the population" (p. 1859), they are expressing their policy view that the primary ozone NAAQS ought to be lowered because science cannot define either *vulnerability* or *adequate protection*. They are, of course, entitled to that opinion. Scientific evidence might help inform their opinion, but it also might not, as policy views are often driven by values, and values influence how much evidence is considered sufficient to either act or stay put.

Scientific peer reviewers often detect and excise inferential exaggeration, but it is entirely plausible that it went undetected in this case. Gent et al. (2003) was accompanied by an editorial that repeated the authors' inferential exaggerations as if they were the same as the authors' research results:

> In the group using maintenance medication, the level of ozone exposure was significantly associated with worsening of symptoms and an increase in the use of rescue medication (Thurston and Bates 2004, p. 1915).

Despite its scientific limitations, Thurston and Bates interpret the evidence reported by Gent et al. (2003) to mean "there is no reason to doubt that ozone exposure is a cause of asthma exacerbations"(p. 1916). When scientists express certainty, it is a powerful hint that they are not talking about science. Their willingness to throw scientific caution to the winds is entirely consistent with holding the unshakeable policy conviction that the ozone standard ought to be lowered irrespective of the scientific evidence. The EPA staff decision to "deem" respiratory symptoms as adverse irrespective of severity or reversibility has a foundation in the Thurston and Bates editorial.[25]

---

[25] "But regardless of the role of air pollution as a contributing factor to the prevalence of asthma, the study by Gent et al. and others like it indicate that the increasing numbers of children with asthma represent an expanding pool of children at risk for respiratory symptoms caused by air pollution, and by ozone in particular")…. Of the many triggers of asthma in the environment, air pollution is one of the few that can be legislated and regulated" (Thurston and Bates 2004, p. 1916, emphasis added). EPA relies heavily on scientific studies performed by Thurston and Bates (there are at least a dozen references in the Staff Paper in which one or both are co-authors), thus raising an obvious question: Where does their scientific analysis end and their policy advocacy begin?

EPA has an obligation under information quality guidelines to conduct sufficient pre-dissemination review to ensure that the information it relies on satisfies information quality standards. That includes detecting inferential exaggeration in scientific papers, limiting the inferences it draws from scientific papers that engage in inferential exaggeration, and subjecting those papers to more rigorous review. In the Staff Paper, EPA includes Gent et al. (2003) within a list of studies which it says "have reported fairly robust associations between ambient O_3 concentrations and daily symptoms/asthma medication use, even after adjustment for co-pollutants" (U.S. Environmental Protection Agency 2007m, p. 3-11). This represents the conclusions Gent et al. (2003) reached, but it does not accurately describe the actual study. It is difficult to honestly infer robustness from a set of exploratory analyses in which all of the odds ratios reported are low and only a handful of them are statistically significant.[26]

(ii) The inferences made by Mortimer et al. (2002) are inconsistent with their own reported results

A similar story can be told with respect to Mortimer et al. (2002). The authors relied on a substandard research design consisting of unvalidated opinion about symptoms obtained from caregivers,[27] limited researcher contact,[28] and the use of diaries to record caregiver-reported data.[29] Response rates were well below the 80% minimum normally required to assume that the presence of

[26] We point out in Section III.A.2(d)(i) that unvalidated self- and caregiver reported data recorded in diaries have been shown to be unreliable. In Section III.C.5(c) beginning on page 102, we note that EPA frequently uses the term "robust" and its variants to describe the consistency of the scientific evidence, but never defines the term.

[27] "Study children had either: 1) parental report of physician-diagnosed asthma and symptoms in the past 12 months or 2) respiratory symptoms consistent with asthma, such as cough, wheezing or shortness of breath, that lasted w6 weeks during the previous year, together with increased symptoms with exercise or cold air exposure or a family history of asthma" (Mortimer et al. 2002, p. 700).

[28] Researcher contact consisted of "an in-person baseline interview, a home survey, three brief telephone follow-up interviews at three-month intervals, and two-week peak expiratory flow rate (PEFR) and symptom diaries after the baseline interview and prior to each follow-up interview" (Mortimer et al. 2002, p. 700).

[29] Diaries were filled out only for the two-week intervals prior to a scheduled "brief telephone follow-up interview." See (Mortimer et al. 2002, p. 700). In Section **Error! Reference source not found.** beginning on page **Error! Bookmark not defined.**, we remind EPA that data from self-administered PEFR meters recorded in diaries have been shown to be unreliable and include manufactured data.

fatal nonresponse bias is not a material defect,[30] and the standard errors calculated by the authors assume without evidence that nonresponse bias is not present.

The authors found no evening effects from ozone that could be measured using self-administered PEF devices. Given the toxicological evidence and chamber exposure studies, one would have expected that evening PEF values exceed morning readings if ozone exposure during the day was causing respiratory effects. But Mortimer et al. did not find elevated evening PEF values. They speculate about how ozone exposure might cause morning but not evening decrements in PEF, and EPA has obligingly amplified their ruminations in the Staff Paper (U.S. Environmental Protection Agency 2007g, p. 3-10).

At this point, all further analysis in Mortimer et al. (2002) should have been understood (both by the researchers and by EPA) to be entirely exploratory. That is especially so given that the claimed statistically significant decrements in morning PEF are all less than 1%. This is a small fraction of the variance in inter-maneuver performance.[31]

EPA admits that Mortimer et al. (2002) chose this multiday lag structure only after it became clear that single-day lags were not going to yield statistically significant positive effects:

> Examination of these single lag day effects led to the consideration of a multiday lag period of 1 to 5 days in the case of PEF and 1 to 4 days in the case of respiratory symptoms to estimate the cumulative effect of O_3 (U.S. Environmental Protection Agency 2008d, p. 32, emphasis added).

With enough statistical effort, Mortimer et al. (2002) were able to unearth a few lags that displayed small statistically significant positive effects, but only buried amidst a huge number of others that were not. For example, none of the

---

[30] Response rates for federally-sponsored surveys with response rates below 80% require an analysis of nonresponse bias simply to be approved under the Paperwork Reduction Act. See (Office of Management and Budget 2006, p. 16). Mortimer et al. (2002, p. 700) report that "approximately 60% of the children returned a diary for each of the four visits." However, they do not reveal the response rate for all four visits, and they do not report the results of any analysis of nonresponse bias. For more about the problem of nonresponse bias as a fatal information quality defect, see section III.A.2(d)(v).

[31] See section III.A.2(d)(i) for a discussion of inter-maneuverinter-maneuver variation in spirometry that is consistently discarded by researchers using these technologies.

single day lags reached statistical significance. The incidence of ≥ 10% declines in PEF was statistically significant for a 4-day lag and a 5-day moving average, but not for lags of zero, 1, 2, 3, or 5 days. The increased incidence of self-reported symptoms was statistically significant for lags of 2 and 4 days, but not lags of 1, 3, 5 or 6 days. See Mortimer et al. (2002, p. 702, Table 2).

Reviewing EPA's Response to Comments led us to take a second look at another vitally important epidemiological study, the time-series analysis by (Bell, McDermott et al. 2004). This time we noticed how frequently the authors described their work as *exploratory*.[32] Despite this cautious adherence to scientific standards in the description of their analyses, for some reason in their conclusions they said their results provided "strong evidence of an association" and implied that this association was both causal and large.[33] They did this even though the used exploratory methods, found weal effects, and could only speculate that the mechanism for ozone exposure causing death "may differ" from the mechanism whereby it caused minor respiratory effects.[34]

---

[32] "Distributed-lag models are appropriate for estimating relative rates of mortality associated with exposure to pollution levels during several previous days, thus allowing more flexibility for exploring the lag between exposure and death than single-lag models. At the second stage, we use hierarchical models to combine the relative rate estimates obtained from the community-specific distributed-lag models to produce a national average estimate. With this 2-stage model, variation across communities in the short-term effects of ozone can be explored and an effect estimated for the nation." (Bell, McDermott et al. 2004, p. 2373, emphasis added and internal citations omitted); "We explored whether the association between ozone and mortality was modified by the long-term average of $PM_{2.5}$ (PM with an aerodynamic diameter less than 2.5 μm) by performing a weighted second-stage linear regression with the community-specific estimate of ozone's effect on mortality as the dependent variable and the long-term $PM_{2.5}$ average as the independent variable. No association was observed" (p. 2376, emphasis added).

[33] "This multisite time-series study of 95 large US urban communities throughout a 14-year period provides strong evidence of an association between mortality and short-term exposure to ozone"; "The results indicate a substantial health burden from ozone pollution" (Bell, McDermott et al. 2004, p. 2376, emphasis added).

[34] "Although the temporal dynamics of the underlying processes linking ozone exposure to increased mortality may differ from those of the inflammatory response, inflammation has been postulated as having a central role in the increased mortality and morbidity associated with ozone "(Bell, McDermott et al. 2004, p. 2377, emphasis added).

(c) Methodological error

In our RFC, we said examples of methodological error can be found in several studies on which EPA heavily relies (National Association of Manufacturers 2007, pp. 14-16). As examples, we mentioned several studies of respiratory symptoms; here, we repeat our examples in the order in which we presented them and discuss EPA's response.

We clearly defined what we meant by a "material effect" – one large enough that it impeded the Administrator's ability from making his decision based on an accurate scientific record (National Association of Manufacturers 2007, p. 7). EPA says it "does not agree that methodological errors exist in these studies that are 'so severe that they have a material effect on utility, particularly for regulatory decision-making'" (U.S. Environmental Protection Agency 2008e, p. 47). However, EPA does not provide much more in the way of rebuttal, as if this is a dispute about policy rather than the application of information quality principles that EPA itself has enunciated and claims to uphold.

- 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).

EPA replies generally that it "conducted a rigorous assessment of potential methodological error in epidemiologic analyses" in section 7.1.3 of the Criteria Document (U.S. Environmental Protection Agency 2008d, p. 47). Yet, EPA's presentation in the Criteria Document discusses, but does not actually address, any of the issues we raised. Neither Korrick et al. (1998) nor Mortimer et al. (2002) come up in this discussion despite the heavy weight EPA places on them in its weight of evidence review.

EPA acknowledges that Korrick et al. (1998) performed repeated statistical tests, but says

> these hypotheses can be divided into confirmatory vs exploratory hypotheses. The main confirmatory hypothesis is whether O_3 concentrations are associated with pulmonary function (U.S. Environmental Protection Agency 2008d, p. 47).

EPA incorrectly equates exploratory data analysis with sensitivity analysis, a tool for evaluating the extent to which the results of a confirmatory data analysis are robust with respect to model specification and other assumptions. Properly understood, exploratory analysis is "detective work" undertaken to get clues

about what a data set might have to say (Tukey 1977).[35] In contrast, Korrick et al. (1998) subjected their data set to numerous statistical techniques in search for the "best" (i.e., strongest) evidence of concentration-response. EPA staff, in turn, rely on "best" statistical results to push the ozone risk envelope outward, and the Administrator relied on the staff's opinion as if were objective (U.S. Environmental Protection Agency 2007h). EPA has insisted that a Bonferroni correction is appropriate for accounting for multiple comparisons in controlled human studies without resulting in excessive Type II error (Brown 2007a). If that correction were applied to the results reported by Korrick et al. (1998), none of the associations reported would have been statistically significant.

In EPA's rebuttal, the Agency staff say that if they cannot find statistical significance they will look for "patterns":

> EPA notes that while statistical significance (i.e., confidence intervals) is considered in the evaluation of the scientific evidence, EPA has emphasized the importance of examining the pattern of results across various studies and not focusing solely on statistical significance as a criterion (U.S. Environmental Protection Agency 2008e, p. 33).

This is more evidence of the Iron Law of EPA Staff Ozone Risk Assessment and Characterization. A dispassionate review of EPA's scientific record shows that in every case where an association is statistically significant, EPA staff interpret that association as evidence of a causal relationship between ozone and health effects. In every case in which an association is not statistically significant, EPA concludes that it nevertheless supports a causal inference. In every case in which a "pattern" of effects is found, that "pattern" also is evidence supporting an inference that a causal relationship exists. Where none of these conditions occur, the study was poorly designed or insufficiently powerful to detect an effect.

- Statistically significant but biologically implausible lags are reported (Mortimer et al. 2002).

In its response, EPA correctly says that it is a reasonable technique in exploratory data analysis to compute and report a wide variety of lags when a biological basis for predicting a specific lag is lacking. However, Mortimer et al. (2002) do not describe their work as exploratory. Rather, they imply that their (statistically significant) moving-average distributed lag model yields results that

[35] In the Criteria Document, EPA confuses sensitivity analysis (performed to illustrate the significance of uncertainty) with exploratory data analysis (the use of statistical tests for which there is no underlying theory to generate new hypotheses) (U.S. Environmental Protection Agency 2006a, p. 7-19).

more accurately describe the pattern of effects than do estimates obtained from (nonsignificant, mixed and oftentimes nonpositive) individual day lag models.[36] They make sweeping, unsupported generalizations about the practical importance of their results.[37]

The EPA staff is clearly smitten by Mortimer et al. (2002). The Criteria Document glosses over the absence of a nonresponse bias analysis despite the low response rate and the absence of evening effects; credits the authors for having "discussed biological mechanisms for delayed effects on pulmonary function" that might support their preferred model, but without characterizing these explanations as speculative; and reports the results of an extensive but nonreproducible additional analysis based on data not reported in the published paper. The Criteria Document repeats the most statistically significant findings reported in Mortimer et al. (2002), but without even the authors' own understated caveats (U.S. Environmental Protection Agency 2006a, pp. 7-43 to 47-46).

EPA's Response to Comments also repeats the post hoc rationalization for moving average distributed lag models using morning-only effects: it is "consistent with the understanding that the development of asthma exacerbation through an inflammatory mechanism would occur over time, with symptoms manifested hours after the exposure period" (U.S. Environmental Protection Agency 2008e, p. 48). But the stated purpose of Mortimer et al. (2002) was to "estimat[e] individual mean effects and individual change over time as well as population mean effects over the entire study period" using "methods [that] require no assumptions about stability of population characteristics over time…" (p. 699). The absence of evening effects was unexpected and the published article offers no explanation why the presumed inflammatory mechanism responsible for morning effects would shut itself off in the evening.

- Single rather than multipollutant models are emphasized (Korrick et al. 1998; Mortimer et al. 2002).

---

[36] "Findings in these USA inner-city asthmatic children are comparable to findings reported elsewhere, suggesting the magnitude of the air pollution-related effect on asthma morbidity is not substantially greater in this population in relation to more socioeconomically diverse groups of asthmatic children" (Mortimer et al. 2002, p. 704).

[37] "In conclusion, summer-time air pollution is associated with increased asthma morbidity and decreased pulmonary function among inner-city children with asthma in the USA. These findings from generalized estimating equations and mixed models support previously published reports from time-series analysis, and those reported from less urban populations" (Mortimer et al. 2002, p. 705).

Our complaint was that EPA relied on single- rather than multi-pollutant models, thus overstating the effect of ozone even if all other considerations were ignored. EPA does not directly respond to this specific error claim (U.S. Environmental Protection Agency 2008d, pp. 47-48), but it does respond to related questions elsewhere in the Response to Comments. EPA's first defense is that Agency staff "include[d] and discuss[ed] results from both single- and multipollutant models when available" and "rigorously and thoroughly evaluated the potential for confounding," but concluded that "the inclusion of copollutants into the models did not substantially affect O_3 risk estimates" and "that effects of O_3 on various health outcomes were robust and independent of the effects of other copollutants" (U.S. Environmental Protection Agency 2008e, pp. 44-45, emphasis added). EPA's second defense is opposite the first: multi-pollutant models were troublesome because they resulted in "reduced stability of the O_3 coefficient estimate in such models (p. 88). This is another example of the Iron Law of EPA Staff Ozone Risk Assessment and Characterization in action: multipollutant models with equivocal results support the current location of the ozone risk envelope; multipollutant models that conflict with that risk envelope are discarded.

- Known confounders are inadequately controlled (Gent et al. 2003; Korrick et al. 1998)

EPA does not directly respond to this specific error claim (U.S. Environmental Protection Agency 2008d, pp. 47-48), but it does respond to related questions elsewhere in the Response to Comments. EPA says it is satisfied with its review of the problems of confounding in epidemiological studies (p. 40), but this satisfaction appears to be focused on co-pollutants rather than non-air pollution confounders. EPA also says that factors such as cockroach and dust mite allergens are unlikely to be genuine confounders in time-series studies because they are not "temporally correlated with O_3" and that they "do not vary from day to day as do ambient O_3 concentrations" (p. 42), though it does not mention that ambient ozone concentrations are not found indoors.

We identified Gent et al. (2003) and Korrick et al. (1998) as studies where control for confounding was inadequate. The most obvious problem is that both did not control for relative humidity, and humidity is clearly associated with pulmonary function changes (Ross et al. 2002). It also seems plausible that the lowest ozone levels occurred on days when humidity was relatively low. The authors also did not obtain data from subjects concerning bronchodilator use, yet there should be no question that the use of such medication will significantly affect respiratory function indicators – that's what they are supposed to do. Pollen is a known cause of allergic rhinitis and asthmatic symptoms, and different pollens have been associated positively -- or inversely -- with

pulmonary function, with magnitude much greater than those for ozone (Ross et al. 2002).[38] The Criteria Document discusses pollen as a confounder only in the context of school absenteeism (U.S. Environmental Protection Agency 2006a, p. 7-58), even though the primary author of Ross et al. (2002) was an EPA employee. Gent et al. (2003), Korrick et al. (1998), and Mortimer et al. (2002) do not attempt to control for pollen.

Controlling for confounding tends to reduce estimated effects. Thus, it is generally incompatible with the Iron Law of EPA Staff Ozone Risk Assessment and Characterization. Ineffective control for confounding results in little change in effect estimates and yields equivocal results that, under thee Principles, support the prevailing location of the ozone risk envelope. Control for inappropriate confounders, such as PM_{10} rather than $PM_{2.5}$ (Bell, McDermott et al. 2004), also will not materially change effect estimates and give the false impression that confounding is not a problem.

If EPA took substantive and presentational objectivity seriously, it would compile a balanced portfolio of causes and risk factors for each major health effect of interest instead of trying to force a causal relationship to air pollution in every case. If EPA had done this with respect to asthma, for example, it would have noticed that its prevalence is rising at the same time that air pollution is falling. Any contribution air pollution might be making thus must be declining, and something else of much greater public health significance is going on. Several recent papers have estimated much larger associations between asthma prevalence and colonization by the gastric bacterium *Heliobacter pylori* in the human gut. Using data from NHANES III, a well-known representative population sample, Chen and Blaser (2007) found that the presence of cagA+ *H. pylori* strains was inversely related to ever having asthma (OR, 0.79; 95% CI, 0.63-0.99), supporting the hypothesis that the lack of normal acquisition ore retention

[38] "O_3 was correlated positively with temperature, as has been frequently observed. It was also correlated positively with grass pollen count, *Curvularia*, and *Drechslera*, and correlated negatively with ragweed pollen count and several mold genera (*Alternaria, Cladosporium, Epicoccurn*). (These correlations were based primarily on seasonal variability and not physical relationships, as with O_3 and temperature.)" Associations with pulmonary function are subject to the limitations and caveats mentioned elsewhere with respect to spirometry, but ozone, temperature and pollen count correlations are not. Reported effects per 20 ppb ozone were 2.3% PEFR decrements in the morning and 2.6% PER decrements in the evening, both before adjustment for aeroallergens. The reported effects from pollen and spores ranged from a PEFR decrement of 6.8% to a PEFR improvement of 31%. See Table 3.

of *H. pylori* is associated with childhood asthma and allergy.[39] A follow-up paper examining very young children obtained even stronger results: the odds ratio for onset of asthma among children under five years of age who had acquired *H. pylori* was 0.58; 95% CI, 0.38–0.88) (Chen and Blaser 2008). As long as EPA persists in trying to link every conceivable health effect to air pollution, it cannot break free of the Iron Law of EPA Staff Ozone Risk Assessment and Characterization.

(d) Pulmonary function testing

Pulmonary function testing is crucial to many of the studies on which EPA relies. In our RFC, we highlighted the information quality problem that these techniques have for non-clinical purposes:

> 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 (National Association of Manufacturers 2007, p. 14).

In its response, EPA summarizes our complaint accurately but answers one we didn't raise concerning reproducibility across devices. We followed that thread, however, and discovered that information quality defects in this body of research are much worse than we originally thought.

(i) Clinically useful pulmonary function tests have inherent information quality limitations that become defects when used in air pollution epidemiology

Pulmonary function tests used to estimate the effects of ozone rely on guidelines published by the American Thoracic Society (ATS) (Miller, Hankinson et al. 2005).[40] These guidelines show that successful testing in clinical settings depends on a combination of factors including the skill of the technician administering the test, the environmental conditions under which the test is

---

[39] For inverse effects, odds ratios are expected to be less than 1. Effects reported to be statistically significant must have upper confidence intervals less than 1.

[40] The most recent editions of the ATS guidance were published in 2005 (MacIntyre et al. 2005; Miller, Crapo et al. 2005; Miller, Hankinson et al. 2005; Pellegrino et al. 2005; Wanger et al. 2005), but on the margins relevant to this discussion earlier ATS guidance is not materially different.

performed, and the level of training and coaching subjects receive. The definition of an "acceptable" test is complex and subject to technician judgment.[41]

After three "acceptable" maneuvers, the two largest FVC and FEV_1 values each must be within 150 mL, or about 3%. Additional maneuvers up to eight can be performed to achieve this error bound, and the technician must "[s]ave, at a minimum, the three satisfactory maneuvers" (p. 325, Table 5). Data should not be discarded solely on the basis for poor repeatability, and the largest values of FVC and FEV_1 should be recorded. If a mid-expiratory flow is taken, it must be measured with an accuracy of ± 0.5% (p. 326).Children present special complications, so ATS recommends that technicians administering tests to children be specially trained.[42]

The ATS guidelines were intended for use in clinical settings where the purpose is to diagnose disease. Reflecting the guidelines' complexity, there is some evidence suggesting that physicians and nurses who administer pulmonary function tests in primary health care settings do not do so very

---

[41] Miller, Hankinson, et al. (2005, p. 325): "The acceptability criteria are a satisfactory start of test and a satisfactory [end of test], i.e. a plateau in the volume–time curve. In addition, the technician should observe that the subject understood the instructions and performed the manoeuvre with a maximum inspiration, a good start, a smooth continuous exhalation and maximal effort. The following conditions must also be met: 1) without an unsatisfactory start of expiration, characterised by excessive hesitation or false start extrapolated volume or EV .5% of FVC or 0.150 L, whichever is greater; 2) without coughing during the first second of the manoeuvre, thereby affecting the measured FEV_1 value, or any other cough that, in the technician's judgment, interferes with the measurement of accurate results; 3) without early termination of expiration; 4) without a Valsalva manoeuvre (glottis closure) or hesitation during the manoeuvre that causes a cessation of airflow, which precludes accurate measurement of FEV_1 or FVC; 5) without a leak; 6) without an obstructed mouthpiece (e.g. obstruction due to the tongue being placed in front of the mouthpiece, or teeth in front of the mouthpiece, or mouthpiece deformation due to biting); and 7) without evidence of an extra breath being taken during the manoeuvre."

[42] "A bright, pleasant atmosphere, including age-appropriate toys, reading material and art, is important in making children feel at ease. Encouragement, detailed but simple instructions, lack of intimidation and visual feedback in the teaching are important in helping children to perform the manoeuvre. Even if unsuccessful at the first session, children will learn to be less intimidated and may perform far better in a subsequent session. Testing children in ''adult'' laboratories, where no effort is made to cater for the specific needs of the younger subjects, is to be discouraged" (Miller, Hankinson et al. 2005, pp. 323-324).

competently. In a randomized prospective study performed in New Zealand, Eaton et al. (1999) found that even after training only 13.5% of patients produced spirometric data that met ATS standards. In the field epidemiology studies on which EPA relies, investigators do not disclose what they did to achieve "acceptable" data.

ATS guidelines also cover the interpretation of spirometric abnormalities (Pellegrino et al. 2005, p. 957, Table 6). The least severe category is "mild" and encompasses all FEV_1 values greater than 70% of predicted; there is no "normal" category.[43] ATS counsels against overinterpreting small changes because intra-personal variability is high:

> It is more likely that a real change has occurred when more than two measurements are performed over time… [S]ignificant changes, whether statistical or biological, vary by parameter, time period and the type of patient. When there are only two tests available to evaluate change, the large variability necessitates relatively large changes to be confident that a significant change has in fact occurred. Thus, in subjects with relatively ''normal'' lung function, year-to-year changes in FEV_1 over 1 yr should exceed 15% before confidence can be given to the opinion that a clinically meaningful change has occurred.[44]

Overinterpreting small changes is a consistent feature of the panel studies the EPA staff relies on to support its inference that ozone concentrations below the 1997 NAAQS cause morbidity. To give just one prominent example, Korrick et al. (1998) reported (and EPA gave considerable weight to) group mean FEV_1 decrements of 2.6% per 50 ppb ozone. Assuming linearity, that's a decrement of 0.4% over the 15 ppb difference between the 1997 and 2007 primary ozone standards. This is less than 10% of the variation in FEV_1 that ATS judges to be clinically meaningful, and is after discarding inter-maneuver variance.

---

[43] None of the studies EPA relies on obtained FEV_1 decrements outside of this category.

[44] See Pellegrino et al. (2005, p. 961, Table 12). Also judged by ATS to be not "clinically meaningful" in normal subjects: (1) within-day changes less than 5% in FVC or FEV_1, and (2) weekly changes less than 11% and 12%, respectively.

(ii) Information quality defects associated with investigator bias

The ATS guidelines specifically directs technicians performing spirometric measurements to coach their patients to ensure best performance.[45] In a clinical setting where the purpose is diagnosis, this is a manageable concern. In a research setting, however, coaching imparts unknown bias to the data. Epidemiologists are not blind to either the hypotheses they are testing or the identity of their subjects, and even if they are scrupulous in their efforts to be unbiased in test administration, test results can be expected to vary across technicians with different coaching skill.

(iii) Information quality defects associated with the use of diaries

EPA's account of the PEFR measurements in Mortimer et al. (2002) is very positive, specifically noting that the National Cooperative Inner-City Asthma Study "used standard protocols that included instructing caretakers of the subjects to record symptoms in the daily diary by observing or asking the child" (U.S. Environmental Protection Agency 2007m, p. 3-11).

In our RFC, we said information quality problems associated with diaries such as these posed special problems. We noted that at least one very high quality study had been performed to ascertain the reliability of data recorded by children's caregivers. Serendipitously, the information that caregivers were supposed to record were the results of spirometric monitoring.

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. These declines were statistically significant using repeated measures ANOVA. Data were correctly recorded only about half

[45] "The subject should be prompted to ''blast,' not just 'blow,' the air from their lungs, and then he/she should be encouraged to fully exhale. Throughout the manoeuvre, enthusiastic coaching of the subject using appropriate body language and phrases, such as 'keep going', is required" (Miller, Hankinson et al. 2005, p. 323).

of the time; incorrectly recorded about 30% of the time; missing about 6% of the time; and invented between one–eighth and one-fourth of the time. Self-reported data were biased toward understating electronically measured respiratory performance.[46]

Kamps et al. (2001) concluded that self-reported PEFR data were unreliable, and that electronic meters should be used instead of diaries. Similar data obtained by ozone researchers who relied on diaries might have been much more reliable than what Kamps and coworkers found, but the information they report does not offer much comfort. Children enrolled in the cohort examined by Mortimer et al. (2002) were "4–9 yrs old and resided in inner-city neighbourhoods in which the income of ≥ 30% of residents was below the federal poverty level" (p. 700). Caregivers are not well described; the authors only say "children and their parents were recruited." Data were supposed to be recorded in diaries, but the response rate was no greater than 60% for a single round -- worse than the response rate reported by Kamps et al. for their entire study.[47]

A more recent study of adult asthmatics being actively treated yielded similar results (Chowienczyk et al. 1994)618. Diaries contained only 70% of the expected number of records, and 26% of recorded entries were invented or mistimed:

> The rationale behind inventing data or entering data retrospectively may be patients' reluctance to admit poor record keeping. The most striking example to support this is the patient who performed 54 forced expirations in three hours on one day and entered these data retrospectively for the previous six days.

---

[46] 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 under the Paperwork Reduction Act. See Section III.A.2(d)(v) beginning on page 57.

The problem of systematic understating actual values presents additional problems. It means some subjects wanted to be perceived as worse off than they actually were, and that means they cannot be trusted to produce valid data even if they could be persuaded to record it correctly.

[47] Mortimer et al. (2002, p. 700) say "[a]pproximately 60% of the children returned a diary for each of the four visits," but they do not report how many children returned diaries for <u>all</u> of the study period.

The authors experienced a 38% loss of sample size among those patients using manual diaries. Accounting for nonresponse bias almost certainly would have made these figures worse.[48]

In its response, EPA says only that it "recognizes that PEF measurements have been shown to be more variable than FEV_1 in some studies" (U.S. Environmental Protection Agency 2008e, p. 48).[49] The relative variability of PEF to FEV_1 measurements is a *non sequitur*, but it turns out to be a revealing one nevertheless. We deal with this in the following subsection.

In their study of asthmatic adults, Ross et al. (2002) acknowledge that they had problems with data quality – problems that are inherent to the research design:

> Our study also had shortcomings that are shared by most panel studies, such as the possibility of incorrect data recording by study participants. Previous surveys have reported that diary cards with self-reported PEFR and symptom data may contain a high number of invented or retrospective entries.[50]

EPA appears to have been well aware of the problems posed by diary recordation of pulmonary function data. The lead author of Ross et al. (2002) is an employee of EPA's Office of Air Quality Planning and Standards.

(iv) Information quality defects associated with inter-maneuver variability

One of the two studies EPA cites for the observation that PEF measurements are more variable than FEV_1 is the study comparing alternative devices by Vaughan et al. (1989) – a study with which we previously had been

---

[48] Electronic data collection assures that the data collected are accurate, but it does not assure that data will be collected. Medical researchers have concluded that both electronic data collection and sufficient motivation to adhere to the prescribed data collection regimen are essential. See Reddel et al. (2002).

[49] See (U.S. Environmental Protection Agency 2006a, pp. 7-27 to 27-47). EPA also tries to rebut Kamps with a paper by Lippmann and Spektor; part of the appeal may be that Lippmann is a longtime CASAC member. The rebuttal paper is off target; it is a comparison of the performance of alternative devices and has nothing to do with the reliability of data recorded in diaries.

[50] See Ross et al. (2002, p. 577, internal citations omitted). They authors add: "We would, however, expect these limitations to bias the study results in the direction of nonsignificance." They do explain why this would be so.

unfamiliar. They did more than just compare inter-instrument variability, however. They also estimated inter-maneuver standard deviations for three maneuver trials. The mean inter-maneuver standard deviations across 102 patients was 3.01% (FEV_1), 5.12% (mini-Wright peak flow meter), and 7.2% (Jones Pulmonar Spirometer). Mean inter-maneuver standard deviations were higher for patients with COPD (3.11%, 5.50%, and 7.43%) than for patients with no disease (2.82%, 4.29%, and 7.03%) (p. 560).

Several of the epidemiological studies on which EPA staff use for risk assessment rely on pulmonary function measurements. Some cite the ATS guidelines as the basis for their procedures, but at least one crucial study relied on caregivers to do this (Mortimer et al. 2002). If the ATS guidelines are followed, then researchers will have obtained between three and eight separate measurements and they will have made crucial decisions concerning which single value is most representative of the subject's contemporaneous pulmonary function. Choosing any single value, taking an average or performing some other calculation, and discarding the remaining data all create a false sense of precision. Typically, epidemiologists use the difference between subjects' pre- and post-exposure performance as their measure of effect due to exposure. Subtracting a pair of fixed values, when each is actually uncertain, exacerbates the excess precision problem.

We took a closer look at these studies and discovered that inter-maneuver variability is never accounted for. In every case, a single value is recorded as representative for each test, often with a very high degree of implied precision. Similarly, differences between pre- and post-exposure pulmonary function are calculated to retain this precision. Instead of taking account of uncertainty and variability inherent to the pulmonary function test, both are discarded. Thus, all reported standard errors in these observational studies are significantly underestimated. Odds ratios and relative risk measures that are reported to be statistically significant almost certainly are not.

Korrick et al. (1998) is representative. They obtained expiratory flow measures from hikers at Mt. Washington, New Hampshire, following the ATS guidelines issued in 1987:

> Each participant performed a <u>minimum of three and a maximum of eight</u> forced expiratory maneuvers before the day's hike and again after returning to the base. For each hiker, mean values for forced expiratory volume in 1 sec (FEV_1) and forced vital capacity (FVC) were the <u>means of the two or three best acceptable and reproducible (± 5%) values</u>.

The adjusted mean percentage changes in FEV_1 and FVC reported by Korrick et al. (1998) were 5.1% and 4.3%, respectively. These figures are about the half the

magnitude as the inter-maneuver standard deviations reported by Vaughan et al. (1989). Had Korrick et al. (1998) taken account of inter-maneuver variability (e.g., by recording the value for each "acceptable" maneuver), it is very unlikely that the effects they reported would have been statistically significant. The PEFR decrements reported by Mortimer et al. (2002) – less than 1% -- are one-ninth to one-twelfth of the mean inter-maneuver standard deviation for PEFR tests reported by Vaughan et al. (1989).

Vaughan et al. (1989) is 19 years old. EPA staff have long been aware that pulmonary function test measurements are not fixed, but highly variable. They have chosen not to include this important information in their discussion and analysis of the short-term epidemiological studies that show weak but barely statistically significant evidence of respiratory effects from ozone exposure below the 1997 NAAQS. Although EPA cites Vaughan et al. (1989) in its Response to Comments, it does not list the paper as a reference; the paper is discussed in the Criteria Document only with respect to variance in PEFR measurements (U.S. Environmental Protection Agency 2006a, p. 7-29); and it is missing entirely from the Staff Paper. In the Criteria Document, EPA staff summarize and discuss many short-term epidemiological studies in which small differences in pulmonary function are estimated and determined to be statistically significant. Not once does EPA staff mention that inter-maneuver variance exists, much less than it had been routinely discarded.

When epidemiologists try to use crude clinical diagnostic tools for sophisticated research purposes like estimating very weak associations, the consequences of discarded inter-maneuver uncertainty and variability become quite serious. Epidemiologists have achieved marginal statistical significance by employing innovative techniques (e.g., distributed lag models) and made expansive claims about the policy relevance of their work. Had they accounted for inter-maneuver variance instead of discarding it, however, the statistical significance of these weak associations would have vanished.

(v) Information quality defects resulting from nonresponse bias

We have already noted that some of the studies on which EPA relies have response rates too low to reasonably assume that nonresponse bias is not a problem. The response rate in Mortimer et al. (2002) was no greater than 60%. Korrick et al. (1998) used a convenience sample and achieved a 78% response rate.

Recent guidance issued by the Office of Management and Budget codified in writing the longstanding but informal government-wide statistical policy which requires that surveys with response rates below 80% include a rigorous

nonresponse bias analysis in order to qualify for approval under the Paperwork Reduction Act (Office of Management and Budget 2006). Both Mortimer et al. (2002) and Korrick et al. (1998) would have failed this test, and possibly also because they obtained a convenience sample. In the study by Gent et al. (2003), 357 children were determined to be eligible, 75 (21%) refused or were lost to follow-up, and 14 (4%) withdrew, leaving a response rate of 76%. The authors say nothing about any efforts they made to estimate nonresponse bias, and it is assumed but not shown that their original sample was representative. They simply assume representativeness sufficient to justify the statistical tests they perform and assume away nonresponse bias.[51]

(vi) EPA's use of PEFR data depends on whether the results support an inference of pollutant-related health effect

In the scientific record for the ozone NAAQS review, EPA considers pulmonary function test data to be valid and reliable despite the problems discussed in the previous five sections. In its Response to Comments, EPA persists in defending the use of "small inexpensive flow meters" apparently because a longstanding CASAC member likes them.[52] In its discussions in the Criteria Document, Staff Paper, and elsewhere, there is no hint of doubt that pulmonary function measurements are anything but reliable.

On August 30, 2007, about six weeks after finalizing these documents and publishing the proposed rule, the Agency separately distributed for public comment and CASAC review its draft Integrated Science Assessment for nitrogen oxides (U.S. Environmental Protection Agency 2007b). Unsurprisingly, some of the same studies that are relevant to estimating human health risks from ozone also are relevant to estimating analogous risks from NOx. Very surprisingly, however, in the NO_x ISA EPA says that pulmonary function test data are "notoriously" unreliable:

---

[51] Korrick et al. (1998) claim that their convenience sample was representative because "[t]he study researcher and hikers were unaware of the ambient O_3 or other pollutant levels." Lack of awareness of ambient ozone levels helps avoid strategic behavior but it cannot achieve sample representatives. Moreover, the purpose of the study was communicated to prospective subjects, and it would not be a surprise if some hikers tried to "help" the researchers prove their point.

[52] See EPA (2008d, p. 33, citing a paper co-authored by Lippmann). Lippmann proposed the citation in his comments on the draft Criteria Document (Henderson 2005a, p. C-66), and EPA obliged. However, the issue at hand was not diary reliability but the relationship between FEV_1 and PEF.

> Reliable data are notoriously difficult to come by using portable peak flow measuring devices (p. 3-16).

EPA summarizes – and dismisses – several studies in which pulmonary function data were collected. Among them: the study by Mortimer et al. (2002), the same study of asthmatic children that, in the ozone Staff Paper, EPA said "suggest[s] that O_3 exposure may be associated with clinically significant changes in PEF in asthmatic children" and identified "plausible biological mechanisms that would explain delayed effects consistent with the distributed lag models that yielded that only statistically significant results."

In the ozone Staff Paper, EPA considers the use of PEFR monitors by Mortimer et al. (2002) to be state of the art and their results persuasive:

> The multicities study by Mortimer et al. (2002), which provides an asthmatic population most representative of the United States, and several single-city studies indicate a robust association of O_3 concentrations with respiratory symptoms and increased medication use in asthmatics (U.S. Environmental Protection Agency 2007m, p. 3-11)

In the NO_x Integrated Science Analysis, however, their work was no good at all.

These differences may be extreme but they are not random. The difference in EPA staff treatment of Mortimer et al. (2002) in the ozone and NO_x cases cannot be the result of a change of heart about pulmonary function tests. The only material difference is that Mortimer and coworkers found statistically significant effects for ozone but no effects for NO_x. Consistent with the Envelope Theory we enunciated in Section I.C, Mortimer et al. (2002) pushes the ozone risk envelope outward (and thus it is valid and reliable) but pushes the NOx risk envelope inward (and thus it must be discarded).[53]

This phenomenon is not an isolated occurrence. In its Response to Comments, EPA is dismissive of the randomized panel study of asthmatic children by Schildcrout et al. (2006) (U.S. Environmental Protection Agency

---

[53] Discarding Mortimer et al. (2002) did not pose much of a barrier to the NO_x health risk characterization: EPA staff found other studies to support its predictable conclusion that NO_2 posed a health risk to asthmatic children:

> Taken together, these studies indicate that short-term exposure to NO2 is associated with respiratory symptoms in children…. For children, the results of new multicity studies provide substantial support for associations with respiratory symptoms, particularly in asthmatic children (U.S. Environmental Protection Agency 2007b, p. 3-31).

2008e, p. A-3 to A-5). EPA faulted it for having just 990 subjects. "As a result," EPA writes, "the total number of children observed by Schildcrout et al. is not comparable to other large multi-city studies that examined the effect of O_3 concentrations on asthma exacerbation, such as Mortimer et al. (2002)." This is an especially odd complaint, inasmuch as the study by Mortimer et al. (2002) included 846 children.[54]

EPA's low opinion of Schildcrout et al. (2006) is limited to ozone, however. In EPA's final Integrated Science Assessment for SO_2, EPA says "the strongest epidemiological evidence for an association between respiratory symptoms and exposure to ambient and SO_2 comes from two large multi-city studies" -- Mortimer et al. (2002) and Schildcrout et al. (2006). The difference is that Schildcrout et al. (2006) reported a statistically significant positive association between SO_2 and respiratory symptoms, but no association with ozone. EPA likes Mortimer et al. (2002) for both ozone and SO_2; Mortimer et al. (2002) found positive associations for both.

3. Peer review practices

In our RFC, we raised questions about the peer review practices of scholarly journals and noted how they differed from government peer review. Most importantly in this context, it is EPA policy to fully incorporate information quality into its peer review practice (U.S. Environmental Protection Agency 2006e). Few, if any, scholarly journals have followed suit. Thus, there is no reason to assume that information quality principles play any significant role in journal peer review.

We also raised questions about EPA's Clean Air Scientific Advisory Committee (CASAC) as a peer review body. We noted that its statutory charge included 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.

---

[54] EPA then resorted to a double negative to reinterpret the authors' no-effect finding, and demand that evidence of no-effect be accompanied by proof, just as we have hypothesized in our Envelope Theory of EPA Staff Ozone Risk Characterization: "Although Schildcrout et al. did not find an association between O_3 concentrations and asthma exacerbation, Shildcrout does not imply the results are inconsistent with those previously found because a thorough evaluation of study populations, uncertainty in parameter estimates, precise scientific questions, and additional comparisons between studies that examined the effect of O_3 exposure on asthma exacerbations has not been conducted." See EPA (2008e, p. A-5, emphasis added).

This problem could have been significantly reduced if EPA had included information quality principles within its charge to CASAC. It did not. EPA's Information Quality Guidelines are not even mentioned in the charge, and unsurprisingly, none of CASAC's reports has anything to say about the subject. EPA may have established a policy whereby information quality is incorporated in peer review, but at least with respect to the CASAC process, that policy has yet to be implemented.

In EPA's Response to Comments, the Agency is silent with respect to these issues.

B. Non-disclosure of critical studies and analyses

In our RFC we said that EPA "excluded scientific information for reasons other than defects in information quality" relevant to determining Policy Relevant Background (PRB) (p. 62.). We highlighted the data reported by Vingarzan (2004), Oltmans et al. (2006), and Brown (2007a).

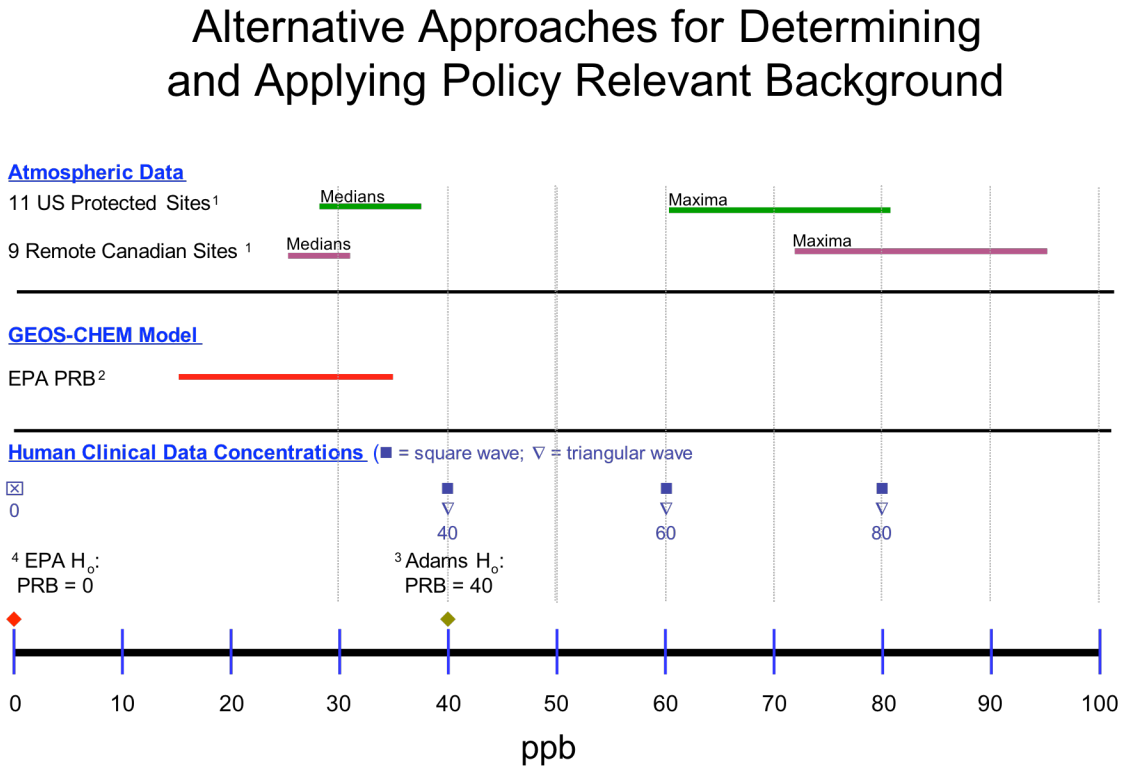
1. Vingarzan (2004) and Oltmans et al. (2006)

In its response, EPA says that the study by Vingarzan (2004) "was considered by EPA" (U.S. Environmental Protection Agency 2008d, p. 96), but as we have noted in many earlier contexts, the nature of that consideration is not at all clear. Vingarzan's data are summarized in the Criteria Document (U.S. Environmental Protection Agency 2006a, pp. 10-49 to 10-50), but that appears to be the sum total of EPA's "consideration." Subsequently, these data disappeared into a black hole: they are not referenced, much less "considered," in the Staff Paper.

Comparing Vingarzan's data as reported in the Criteria Document against the range EPA selected (15 to 35 ppb) shows the bias we allege exists in EPA's determination of PRB:

- At background stations in protected areas of the U.S. (Table AX3-11), the means (standard errors) of 11 lower- and upper-bound medians are 28 (2.8) ppb and 37 (4.6) ppb, respectively. EPA's lower-bound PRB (15 ppb) is 4.6 standard errors below the mean of the lower bounds. EPA's upper-bound PRB is 0.4 standard errors below the mean of the upper bounds.
- Concentrations at background stations in Canada (Table AX3-12), the means (standard errors) of nine lower- and upper-bound medians are 26 (0.5) ppb and 31 (0.7) ppb, respectively. EPA's lower-bound PRB (15 ppb) is 22 standard errors below the mean of the lower bounds. EPA's upper-bound PRB is 5.7 standard errors above the mean of the upper bounds.

The figure below illustrates visually the gap between observational data, EPA's GEOS-CHEM model, and the controlled human data on personal exposure.[55]



Sources: [1]Vingarzan (2004), [2]EPA (2008), [3]Adams (2006), [4]Brown (2007)

EPA acknowledges that it did not "consider" the paper by Oltmans et al. (2006). This EPA attributes to the study having been "published after completion of the Criteria Document" (p. 96). The paper was accepted for publication on

---

[55] The ranges labeled "medians" are arithmetic averages of the lower- and upper-bound annual medians reported by Vingarzan. Similarly, the ranges labeled "maxima" are the arithmetic averages of the reported lower- and upper-bound annual maxima. Averaging reduces the influence of individual annual values.

The concentrations used in Adams also are reported, with the PRBs that Adams and EPA, respectively, assumed in hypothesis tests. Elsewhere we deal with the problem of EPA's transparent effort of ex post data mining. The relevant point here is that the extremism inherent in EPA's use of zero ppb as the "background" level in its hypothesis is visually obvious.

January 18, 2006 (before the "completion" of the Criteria Document) and posted online on March 22, 2006. Oltmans et al., (2006) shows that background ozone at 33 remote locations varies significantly within and across years, and by location. Values similar to EPA's upper-bound for PRB (0.035 ppm) have been frequently observed, but no example is reported in which background levels ever approached EPA's lower-bound (0.015 ppm), even though some values include ozone from anthropogenic sources.

With respect to the substance of this scientific information, EPA's response is just dismissive: "EPA has already discussed the fact that there is spatial and seasonal variability in PRB in the Criteria Document and Staff Paper and the GEOS-CHEM model runs also show this spatial and seasonal variability" (U.S. Environmental Protection Agency 2008e, p. 93). Consistent with the Envelope Theory, scientific studies that push the ozone risk envelope inward -- or, in this case, reduce the potential risk reduction that a more stringent NAAQS might achieve -- will be "considered" and "discussed" before they are discarded.

2. Brown (2007a)

The most obvious example of EPA's gerrymandered scientific database is EPA's own reanalysis of the Adams (2006a) clinical data. EPA placed this document into the docket six days before the Administrator signed the proposed rule (Brown 2007a). This reanalysis is the lynchpin to EPA's scientific database, EPA's denials notwithstanding.[56] It is the only putative scientific basis EPA has for claiming that there is clinical evidence that 0.06 ppm ozone causes any decrement in pulmonary function, adverse or otherwise, in healthy adults. Yet it appeared in the docket at the eleventh hour – without public notice and comment (unlike the Criteria Document) and without peer review (unlike Adams (2006a)).

In our RFC, we characterized EPA's action as a clear violation of applicable information quality guidelines:

> This reanalysis is fully subject to information quality standards and does not benefit from the weak rebuttable presumption of objectivity

[56] "[T]he Brown Memorandum [i.e., Brown (2007a)] is not a crucial element of the staff's policy recommendations, as it was prepared after completion of the Staff Paper, or the Administrator's final decision" (U.S. Environmental Protection Agency 2008d, p. 20). This statement is counterfactual. EPA relied specifically on Brown (2007a) in the preamble to the proposed rule (U.S. Environmental Protection Agency 2007h, p. 37828) and even more so in the preamble to the final rule (U.S. Environmental Protection Agency 2008b, pp. 16454-16455).

> 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 (National Association of Manufacturers 2007, p. 17).

(a) EPA's reanalysis of Adams (2006a) is technically defective.

In our RFC, we objected on information quality grounds to the technical merits of EPA's reanalysis of the Adams (2006a) data (National Association of Manufacturers 2007, p. 18). This reanalysis (Brown 2007a) consists of post hoc statistical tests of selected data originating in an admittedly low-quality analytic review ("visual comparison" and "cursory evaluation," p. 3) in which EPA staff homed in on pulmonary function responses from two of the 30 subjects. The reanalysis was structured for the purpose of minimizing Type II error (failing to reject the no-effect hypothesis when in fact it is false, pp. 5ff). The reanalysis compares changes in pre- to post-exposure responses for square- and triangular-wave exposures as if the intermediate effects during the 6.6 hour test period are unimportant. In the analysis, EPA provided no bona fide external technical defense for the statistical methods it used, and after being challenged, it has failed to provide a technical defense in its Response to Comments. In both instances, EPA cites itself as its technical authority.

(b) EPA's explanation for why it reanalyzed selected data from Adams (2006a) is materially incomplete and misleading

In its Response to Comments, EPA says that its reanalysis was merely "a logical progression" from Adams (2006a) that was somewhat delayed only because Adams' papers were "not published until 2006" (U.S. Environmental Protection Agency 2008e, p. 20). The logical progression we see is less scientific than policy-driven. With the help of Adams,[57] we have reconstructed the timeline of events to prove that EPA's explanation is highly misleading, and thus in its Response to Comments EPA has committed a new violation of the Agency's information quality guidelines.

---

[57] William C. Adams, personal communications during July and August 2008.

Contrary to EPA's claims, EPA staff began to reanalyze Adams (2006a) in search of statistically significant effects at 0.06 ppm about 18 months before Brown (2007a) was placed in the docket. The key event appears to be a request from CASAC that Adams (2006a) be included in the Criteria Document. That occurred in December 2005. From that date onward, EPA staff obtained portions of Adams' data, reanalyzed them, presented their results informally at an EPA-sponsored symposium, and tried to persuade Adams to join them in supporting their statistical reanalysis.

To show how misleading and self-serving is EPA's version of the story, the **facts reported by EPA in its Response to Comments are presented in bold green font** and ***EPA's significant omissions are presented in bold italic red font***:

- September 13, 2002: EPA hires Adams as a consultant to co-author Chapter 7 of the Criteria Document. EPA asks Adams to update the summary of human ozone exposure research and instructs him to ignore all studies not accepted for publication; his review excludes the recent research that became Adams (2006a, 2006b) which had not been submitted.
- August 2005: Adams submits his updates of human ozone exposure research following CASAC review of the first draft Criteria Document. It excludes any discussion of the research that became Adams (2006a, 2006b). These studies are subsequently included in the Criteria Document, but EPA never asks Adams to revise this update, nor does EPA ask Adams to review the summary they write.
- July 28, 2005: Adams (2006a) is accepted for publication in the refereed journal *Inhalation Toxicology*. Adams confirms results for 0.08 ppm obtained in Adams (2003) but reports no statistically significant effects at the previously untested concentrations of 0.06 ppm and 0.04 ppm.
- November 2, 2005: Adams (2006b) is accepted for publication in the refereed journal *Inhalation Toxicology*. Adams largely confirmed results obtained by (Hazucha et al. 1992).
- December 9, 2005: ***EPA's James Brown notifies Adams that CASAC panel member Henry Gong has requested a copy of Adams (2006a). Adams sends Brown a copy of the corrected galley proofs of Adams (2006a). The galleys contain only a few minor handwritten corrections.***
- December 15, 2005: ***Brown notifies Adams that CASAC has asked EPA to include Adams (2006a) in the Criteria Document.***

- December 21, 2005: *EPA's Harvey Richmond requests from the American Petroleum Institute (API), at a minimum, the pre- and post-exposure (but not hourly) data from Adams (2006a).*
- January 10, 2006: *Richmond requests from Adams the pre- and post-exposure (but not hourly) FEV_1 data in Adams (2002, 2003, 2006a). This is the first and only time Adams received any request for data from EPA.*
- January 17, 2006: *API public comment on the initial draft Staff Paper asks EPA to include Adams (2006a) in the final risk assessment and notes that EPA staff have asked Adams for a portion of the data from Adams (2006a).* **EPA accedes to API's request that Adams (2006a) be included,** *but does not acknowledge CASAC's identical prior oral request.*
- January 20, 2006: *Adams sends pre- and post-exposure FEV_1 (but not hourly) data in Adams (2006a) to Richmond.*
- January 23, 2006: *API responds to Richmond noting their earlier request to Adams to consider providing the data; Richmond confirms Adams' provision of the requested data.*
- January 23, 2006: *Adams sends pre- and post-exposure FEV_1 (but not hourly) data in Adams (2003) to Richmond.*
- January 25, 2006: *Adams sends pre- and post-exposure FEV_1 (but not hourly) data in Adams (Adams 2002) to Richmond.*
- February 10, 2006: *In its letter to the Administrator on the second external review draft of the Criteria Document (Henderson 2006a, p. 5), CASAC specifically recommends that Adams (2006a) be included.*
- December 14, 2006: *Brown presents summary results of his reanalysis of Adams (2006a) at "EPA Workshop on Interpretation of Epidemiologic Studies of Multipollutant Exposures and Health Effects," Chapel Hill, N.C. (Brown 2006).*
- January 3, 2007: *Brown provides Adams with his draft reanalysis (Brown 2007b) and seeks Adams' collaboration on a final version.* The draft text includes a courtesy acknowledgement that Adams provided the data that EPA had requested.
- January 6, 2007: *EPA adds to the draft final Staff Paper a summary of Adams (2006a), including material "not mentioned in the CD," and a summary of Brown's reanalysis (U.S. Environmental Protection Agency 2007k, pp. 3-5 to 3-9). Brown (2007b) is neither cited nor*

disclosed, thereby limiting public comment and CASAC review to the nonreproducible summary presented in the draft final Staff Paper.

- January 9, 2007: Adams declines Brown's offer to collaborate on Brown's interpretation of the FEV_1 results for 0.06 ppm from Adams (2006a). Adams tells Brown that he disagrees with Brown's focus on pre- and post-exposure FEV_1 data only and Brown's choice of statistical methods.
- January 24, 2007: Richmond requests from API a copy of Adams' 1998 draft final report to API cited in Adams (2006a, p. 133). The 0.12 ppm exposure results, but not the 0.06 ppm results, are published in Adams (Adams 2000). API makes no response to Richmond's request; this is reported by EPA as "**API refus[es] to provide Dr. Adams technical report describing that data.**"
- March 4, 2007: **During the CASAC teleconference to review the final Staff Paper, presentations are made by Adams (2007) and Richard Smith (University of North Carolina--Chapel Hill) (Smith 2007b)**:
  - ***Adams objects to several aspects of EPA's summary of his work in the draft final Staff Paper, which he believes is not fairly or accurately presented.***[58]
  - ***Smith "used the same statistical approach" that Brown used in his reanalysis of the Adams data.***[59] ***He "also utilized t tests to evaluate the statistical significance of the Adams data…"***[60] ***Smith "specifically indicated that the FEV_1 responses, in the Adams (2006[a]) study, following the two 0.06 ppm O_3 exposures were statistically different from the FEV_1 responses following filtered***

[58] See Adams (2007). Adams raised three specific objections: (1) EPA's use of standard errors instead of standard deviations, which he says reduces subject variability by a factor of about 5.5, thereby making apparent statistical significance much easier to observe; (2) EPA's statement that exposure to 0.06 ppm causes small group mean FEV1 decrements, which Adams says were not statistically significant; and (3) EPA's claim that the fraction of Adams' subjects who experienced greater than 15% FEV1 decrements was lower than in EPA chamber studies because of adaptation to higher ozone levels in Davis CA than in Chapel Hill NC, which Adams says is simply factually incorrect based on EPA's own ambient monitoring data.

[59] See EPA (2008e, p. 22).

[60] See EPA (2008e, p. 27).

air exposures using a paired t test."[61] Smith's purpose in using the "same statistical approach" was to reproduce results reported by EPA in the draft final Staff Paper (U.S. Environmental Protection Agency 2007k, pp. 3-5 to 3-9), not to provide guidance on correct statistical procedures.

- ***Smith says the draft final Staff Report understates the confidence interval on its tests of the proportion of individuals who showed an FEV_1 decrement greater than 10%. Smith calculates the confidence interval as 0.8% to 22%, not 6% to 16%.***
- ***Smith objects to EPA's use of filtered air as the baseline from which to measure the effects of 0.06 ppm. "In making policy-relevant comparisons, those with 0.04 ppm ozone level are more relevant than those with filtered air, which does not represent a realistic background level."***
- ***Smith objects to EPA's use of a logistic response curve because it "assumes that the response curve fitted to higher ozone levels can be extrapolated downwards to 0.06 ppm. Given the large uncertainty in the probability of response at 0.06 ppm ozone, I do not believe the staff paper's conclusions on this point are justified."***
- ***Smith says, "[W]hen all possible comparisons are taken into account, there is insufficient evidence to conclude that there is any well-defined response to ozone exposure below the 0.080 ppm level."***
- ***Smith recommends against using the statistical method employed by EPA.***

- March 8, 2007: Richmond makes second request to API for a copy of Adams' 1998 draft final report to API.
- March 15, 2007: API declines Richmond's March 8th request and encourages EPA to rely on published 0.06 ppm exposure results (Adams 2002, p. 741).
- March 16, 2007: Richmond asks API for additional details about the 0.06 ppm exposure cohort described in Adams (Adams 2002, p. 741; 2006a, p. 133).

[61] See EPA (2008e, p. 21).

- June 20, 2007: **EPA places its reanalysis of the Adams data (Brown 2007a) in the rulemaking docket. The text says EPA "obtained" the data from Adams (2006a)** ***but does not say that Adams' provided these data in January 2006. The text implies that EPA performed its reanalysis in response to March 2007 comments to CASAC by Smith (2007b), and says nothing about the December 2006 presentation in which preliminary results of EPA's reanalysis were disclosed (Brown 2006) or the January 2007 draft of the reanalysis (Brown 2007b), both of which predate Smith (2007b).***[62]
- March 15, 2008: In its Response to Comments:
  - **EPA says API asked that Adams (2006a) be included in the Criteria Document**, ***but does not acknowledge that CASAC had already made the same request.***
  - **EPA says Smith's public comment to CASAC initiated EPA's reanalysis of the data in Adams (2006a)**, ***but does not acknowledge that:***
    - ***EPA began its reanalysis in early 2006***
    - ***Summary results from EPA's reanalysis were first disclosed by EPA in December 2006***
    - ***Brown shared a draft of his reanalysis with Adams in January 2006 seeking Adams' collaboration.***

The most reasonable interpretation of this history is that EPA staff fully intended not to include Adams (2006a) in the Criteria Document because it did not show statistically significant effects at 0.06 ppm. However, once CASAC asked that it be included, EPA staff had to find a way to discredit Adams' conclusions without challenging Adams' professional reputation). EPA staff resolved to obtain and utilize only selected parts of Adams' dataset, thus making the statistical challenge of "finding" significant effects less daunting. Brown

---

[62] In its Response to Comments (U.S. Environmental Protection Agency 2008e), EPA several times refers to Smith's March 2007 comments to CASAC, which are referenced herein as RL Smith (2007b). However, the list of references cites only Smith's October 2007 public comment to EPA, which is referenced herein as RL Smith (2007a). In addition to selectively presenting RL Smith (2007b) and materially misrepresenting them as supporting EPA's statistical methods, EPA never responds to any part of RL Smith (2007a) even though Smith says that Brown materially misrepresented his work in Brown (2007a). This new information quality error persists throughout EPA's Response to Comments.

telegraphed their preliminary success to the public via his December 2006 presentation at the EPA-sponsored conference in North Carolina. Subsequently, Brown offered Adams an opportunity to assist in the reanalysis -- the most conventional of ways used to co-opt academic researchers – but Adams refused.

The prospect that EPA would reinterpret a published, peer-reviewed study to reach an opposite conclusion alarmed the study's sponsor, the American Petroleum Institute. API hired Smith to replicate EPA's alternative analysis. Smith was mostly able to do so, but discovered significant technical errors in Brown's work and further opined that Brown's entire approach was fundamentally flawed because it failed to account at all for multiple comparisons. However, in the process of describing fully why EPA ought not to proceed along the course it set forth in the Staff Paper, Smith unwittingly gave EPA staff a road map for how to correctly perform the calculations for its preferred (but fundamentally flawed) statistical method. EPA staff seized the opportunity to avoid technical error and recast the reanalysis as a "confirmation" of Smith's work. EPA placed the final reanalysis in the rulemaking docket the same day the Administrator signed the proposed rule.

There is no record that any CASAC member actually focused on the issues raised by Adams and Smith during the March 2007 teleconference. Brown's work was never subjected to review by CASAC because it was placed in the docket after CASAC review was completed. Brown asserted that unnamed CASAC members "supported" his statistical approach,[63] but this cannot be documented because EPA's Science Advisory Board does not make transcripts of teleconferences.

We have looked elsewhere for evidence of CASAC support for the EPA staff's statistical methods. In comments prepared in August 2006 – four months before Brown disclosed results from his preliminary reanalysis (Brown 2006) and 10 months before EPA placed his work in the rulemaking docket (Brown 2007a) -- CASAC panel member Frederick Miller supported highlighting selected

[63] Brown (2007a, p. 5): "On the March 5, 2007 teleconference, members of the CASAC O_3 Panel noted the very conservative nature of the statistical test used by Adams to evaluate the research questions posed by the author. These same CASAC Panel members also supported the approach adopted in the OAQPS Staff Paper to evaluate the statistical significance of O_3-related lung function responses associated with pre- versus postexposure responses. The CASAC Panel members also supported the use of the paired t test approach as the preferred method for analyzing the pre- minus postexposure lung function responses."

subjects from the Adams' dataset,[64] but this is the only published suggestion of support we can find.[65] EPA repeats this undocumented claim as fact in its Final Rule[66] and in its Response to Comments.[67] There is no public evidence of any CASAC debate about the propriety of paired *t* tests for analyzing controlled human exposure data in any of the CASAC reports or in the transcripts of CASAC meetings. EPA's assertion is not supported by any factual record, without which it must be inferred that the Agency has responded to our allegations of information quality error by committing additional information quality errors that it hoped would not be detected.

(c) Finding statistically significant effects at 0.06 ppm required EPA to use creative statistical methods

Brown does not say that Adams' choice of hypothesis tests was incorrect, nor does Brown claim that Adams' concern about controlling for multiple comparisons was misplaced. Rather, Brown says he used Adams' data for a completely different purpose than the one for which the study was intended, and therefore multiple comparisons adjustments are not necessary for the reanalysis:

---

[64] "While the discussion of the low level exposures used in the controlled human studies by Adams and colleagues is technically correct that no statistically significant changes were found in FEV_1 compared to filtered air, the fact that a reasonable percent of the subjects had large decrements is glossed over" (Henderson 2006c, p. D-39)

[65] Fellow CASAC panel member Svere Vedal seems to have strongly opposed EPA's cherry-picking of the data. See subsection III.B.2(f) below.

[66] EPA (2008b, p. 16456): "[M]embers of the CASAC Panel noted on the March 5, 2007 teleconference the very conservative nature of the approach used by Adams to evaluate the research questions posed by the author. These same CASAC Panel members also supported the use of the statistical approach (i.e., paired-t test) used in the analysis prepared by the public commenter, which was the same approach later used in EPA's reanalysis, as the preferred method for analyzing the pre-minus post-exposure lung function responses reported in this study."

[67] EPA (2008e, p. 21): "[I]n the Staff Paper, it was noted that a statistically significant difference in FEV1 responses was suggested by a lack of overlap in the standard error of the responses following 6.6 hours of exposure to 0.06 ppm ozone versus filtered air. That interpretation of the data was supported by CASAC review." Elsewhere in the Response to Comments, EPA offers the much weaker defense that its work "was reviewed by the CASAC O_3 Panel and there were no objections expressed" (p. 98). In short, EPA's position is that panel member Vedal's concerns objections (see subsection III.B.2(f) below) do not constitute "objections," and the absence of strenuous peer reviewer opposition is equivalent to peer reviewer endorsement.

> [A]lthough appropriate for the design and intent of the Adams' studies, the multiple comparison correction is overly conservative (increased Type II error and decreased power) for the evaluation of pre- to postexposure changes in FEV_1 between an air and an O_3 exposure and we adopted the standard approach used by other researchers (e.g., Hazucha et al., 1992; Horstman et al., 1995; McDonnell et al., 1991).[68]

The "standard approach used by other researchers" is an example of the logical fallacy known as *argumentum ad verecundiam* -- an appeal to external authority without regard for the truthfulness of the claim itself. EPA's reliance on this logical fallacy constitutes indisputable information quality error.[69]

It also turns out that the "other researchers" cited by Brown do not in fact support his peculiar statistical methods. Horstman et al. (1990, p. 1160) used paired *t* tests to "determine[] the time point at which significant decrements in FEV_1 were observed" during intermediate points of a protocol involving five hours' exposure to 0.00, 0.08, 0.10, and 0.12 ppm ozone. However, they acknowledged that multivariate analysis of variance (MANOVA) for repeated measures would have been "more appropriate." They did not use MANOVA because "this analysis revealed no significant differences ($p = 0.6$) among the four concentrations" – including 0.12 ppm.[70] McDonnell et al. (1991) used paired *t* tests, but a single concentration for a single time period was tested in the study, making the multiple comparisons question irrelevant. Both Hazucha et al. (1992) and Horstman et al. (1995) conducted studies in which multiple comparisons were being made. Hazucha et al. (1992) used two-way ANOVA followed by an

---

[68] See Brown (2007a, p. 5, emphasis added). In our RFC, we cited this statement when we said "Agency staff used the Adams data for purposes that were never intended by the study design" (National Association of Manufacturers 2007, p. 18). In its Response to Comments, EPA recites our information quality objection but provides a reply that it unresponsive (U.S. Environmental Protection Agency 2008e, pp. 21-22).

[69] In fact, the "other researchers" cited in Brown (2007a, p. 5) are not independent. Six of eight co-authors of McDonnell et al. (1991) were at the time EPA employees; one of the non-EPA employees subsequently joined EPA. Of the three co-authors of Hazucha et al. (1992), two were EPA employees. Of the five co-authors of Horstman et al. (1995), the lead and one other co-author was an EPA employee; one of the non-EPA co-authors subsequently joined EPA. The identity of the non-EPA co-author who subsequently joined EPA is James Brown.

[70] The authors say MANOVA was "strongly biased toward a negative outcome" because of limited degrees of freedom, but they do not mean "bias" in a statistical sense. Rather, they mean MANOVA was too demanding as a statistical tool.

unspecified multiple comparisons procedure. Horstman et al. (1995) also used ANOVA, but did not follow with adjustments for multiple comparisons. Brown was a co-author of this study. Thus, Brown's *argumentum ad verecundiam* is worse than merely an appeal to external authority; it's a circular reference to his own prior work.[71]

EPA staff assert that it is acceptable practice to perform simple paired *t* tests on selected results and discard the other data. In our RFC, we asked EPA to disclose an external, independent authority for this statistical method:

> 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 (National Association of Manufacturers 2007, p. 18, emphasis added).

In its response, EPA does not provide a supporting external statistical authority. Smith is a recognized external statistical authority, and he submitted public comments on the proposed rule (Smith 2007a) that are highly critical of EPA's statistical practice, including EPA's mischaracterization and misuse of his public comment to CASAC (Smith 2007b) by Brown (2007a). Instead of responding to the substance of Smith's objections, EPA implicitly suggests that his review is biased because it was funded by API.[72] Where Brown (2007a) engages in the fallacy of *argumentum ad verecundiam* -- appealing to authority instead of logic or fact -- in its Response to Comments EPA commits the highly analogous fallacy known as *argumentum ad hominem circumstantiae* – rejecting claims based on unrelated circumstantial aspects of the opponent – in this case, the opposing

---

[71] The literature that Brown doesn't cite also isn't helpful to his cause. For example, the first controlled human exposure study that tested prolonged exposures – Folinsbee et al. (1988) – used "[m]ultivariate analysis of variance methods appropriate for designs with repeated measurements." Unlike Brown, Folinsbee and his two EPA colleagues refrained from drawing confirmatory inferences based on statistical tests of exploratory hypotheses: "All other tests of hypotheses were of secondary importance and were done only to describe other potential ozone effects and clarify patterns in the data" (p. 30).

[72] "The Brown Memorandum confirms analyses completed by Dr. Smith who was funded by API to perform his analyses and to provide comments to CASAC" (U.S. Environmental Protection Agency 2008e, p. 22, emphasis added). In his public comments, Smith notes that in addition to funding from API, he has received funding from EPA and NIH – neither of which EPA mentions.

authority's source of funding -- rather than the merits of the opponent's argument.

In our RFC we said that Brown (2007a) was not reproducible (National Association of Manufacturers 2007, p. 17) and in its Response to Comments EPA counters that in fact it is (U.S. Environmental Protection Agency 2008e, p. 20)(p. 20). We failed to make our concerns as clear as we should have. The issue at hand is not the algebraic calculation of paired *t* tests. Rather, what is missing from Brown (2007a) is any cogent rationale justifying EPA's analytic approach. The simplest explanation is that EPA staff determined that, to support their policy goals, it was necessary to have statistically significant group mean effects at 0.06 ppm, and when Adams (2006a) came up dry EPA staff needed to find a statistical test that would produce the desired results. The task was challenging in part because EPA staff had never before questioned the statistical methods of Adams or any other researcher performing controlled human experiments. The only way to be able to avoid the burden of making multiple comparisons adjustments was to discard all of Adams' intermediate time period data.

We said in our RFC that EPA staff was so wedded to the policy conviction that the primary NAAQS should be set at 0.06 that they did not merely blur line between science and policy, but they obliterated it. EPA's Response to Comments does nothing to contradict us. In its response, EPA simply "rejects" our arguments without offering an iota of evidence supporting its position (U.S. Environmental Protection Agency 2008e, pp. 21-22). EPA's defense consists of (1) noting that the American Petroleum Institute had asked that Adams (2006a) be included in the scientific record, though conveniently neglecting to mention that CASAC had previously made the same request; (2) cherry-picking Adams' data and statistics textbooks to "discover" statistically significant effects; (3) pretending that these methods were in commonplace use by "other researchers"; and (4) misleading the public to believe that the purpose of EPA's reanalysis of these selected data was only to confirm what Agency staff first learned from Smith's public comment to CASAC (Smith 2007b).

EPA's attempt in the Response to Comments Document to hide Brown (2007a) behind Smith (2007b) is obvious:

> Consistent with common practice for comparing pre-and postexposure [sic] responses to test for whether or not an O_3-related effect is significant, Dr. Smith used a conventional paired t test (U.S. Environmental Protection Agency 2008e, p. 3).

Unfortunately for EPA, Smith is unwilling to serve as the Agency's intellectual shield. He says EPA's statistical procedure is invalid:

> The use of paired t tests to determine significant effects, as originally performed in the EPA Staff Paper and subsequently defended in Brown (2007[a]), is invalid without taking account of the "multiple comparisons" issue (Smith 2007a, p. 1).

Brown (2007a) purports to take account of the multiple comparisons problem but struggles to discover a procedure sufficiently weak that statistical significance – the EPA staff's essential public policy goal -- is still achieved. Having first rejected Scheffé as "conservative," he then rejects the less "conservative" Bonferroni correction[73] because it, too, is too demanding. Needing a threshold no smaller that $p < 0.001$, Brown stumbles upon a solution, though its improvisational reverse-engineering nature cannot be disguised:

> By contrast, a critical p-value might more appropriately be 0.05/5 or 0.01 for assessing pre- to postexposure changes in FEV_1 between an air and an O_3 exposure in the Adams (2006) study.[74]

Brown had many other multiple-comparisons adjustment procedures to choose from, but apparently none fit the bill. In his public comment to EPA on the proposed rule, Smith examined several such procedures, including those devised by Scheffé, Tukey, and Dunnett. All yielded the same result:

> Although the Scheffé procedure used by Adams (2006[a]) is arguably too conservative, alternative options are available through the Tukey and Dunnett procedures. These yield similar results to the Scheffé procedure when performed as part of an analysis of variance, and imply that there is no clear evidence of a decrease in lung function at a mean ozone concentration of 0.06 parts per million (ppm), compared with filtered air.[75]

What Smith makes painfully clear is that the EPA staff's choice of post hoc multiple comparisons adjustment was driven by its need to discard data so that they could dispense with analysis of variance, the standard statistical technique used by scholars, who publish in peer reviewed journals. For its part, in its Response to Comments EPA has nothing to say about Smith's analysis and observations; the Agency is obligated by the Clean Air Act only to respond to

---

[73] In this case, the Bonferroni correction for 90 comparisons yields a p threshold of 0.05/90 = 0.000556.

[74] Brown (2007a, p. 5, emphasis added).

[75] Smith (2007a, p. 1).

those public comments it alone judges to be "significant,"[76] a threshold that Smith (2007a) apparently failed to meet.

(d) The policy-relevant background ozone concentration

Adams estimated whether effects at 0.06 ppm were statistically significant when compared to both filtered air and 0.04 ppm, the same level EPA used for background in its 1997 revision of the ozone NAAQS primary standard (U.S. Environmental Protection Agency 1996b, p. 65726). In both cases, Adams found no statistically significant effects when using statistical methods that account for multiple comparisons, as discussed in the previous subsections. However, Adams did report statistically significant "net" responses for 0.080 ppm whether 0.06, 0.04, or filtered air (i.e., zero ppb) was used as the presumptive background. Adams also tested the difference in FEV_1 response between 0.08 and 0.06 ppm and found that the difference was statistically significant. It is for these reasons Adams concluded that 0.04 and 0.06 pm behaved more like background than like 0.08 ppm.[77]

In his March 2007 comment to CASAC, Smith (2007b) also opined that the tested 0.04 ppm concentration was "more relevant than filtered air, which does not represent a realistic background level" (p.1). EPA staff insist that background is well below 0.04 ppm – so much lower, in fact, that zero ppm is a better proxy for background than 0.04 ppm (Brown 2007a, p. 4, footnote 4). Brown dismisses Smith's objection, once again relying not on any independent authority but on a combination of EPA staff wisdom and an EPA staff policy decision (disguised as "science") to push the policy relevant background below 0.04 ppm:

> As discussed below, we and most authors of the controlled human exposure studies believe that the appropriate approach for testing for an O_3-related response is to compare with filtered air to correct for the effect of exertion in clean air. Additionally, as discussed in the O_3 AQCD (EPA, 2006, AX3-131) and in Chapter 2 of the OAQPS Staff Paper, the scientific evidence supports estimates of policy-relevant background that are in the

---

[76] "The promulgated rule shall also be accompanied by a response to each of the *significant* comments, criticisms, and new data submitted in written or oral presentations during the comment period." See Clean Air Act § 307(d)(6)(B), emphasis added.

[77] The peak exposure in Adams' 0.080 ppm triangular exposure was 0.15 ppm – significantly above the current 1-hour NAAQS. Thus, it is highly inappropriate to construe this exposure protocol as approximating actual ambient conditions at the existing NAAQS.

> 0.015 to 0.035 ppm range in the afternoon during the O_3 warm season, rather than the 0.040 ppm level cited by Dr. Smith (Brown 2007a, p. 4[fn 4], emphasis added).

Brown is the sole named author of the memorandum, so his use of first person plural is stilted at best, and none of the "most authors" he has in mind are identified. Presumably, Brown is referring to the coterie of researchers located in the Research Triangle Park area who perform controlled human exposure studies.[78] These researchers either are EPA employees or are funded by EPA grants. Thus, it hardly would be surprising that, if forced to take a position, "most authors" of the controlled human exposure studies would agree with Brown. They are, after all, his EPA colleagues and answer to the same master. Nevertheless, the EPA staff commitment to using zero ppb as a proxy for ambient background is a policy-driven constraint not supported by scientific evidence.[79]

Apparently unwittingly, Adams stated clearly the conundrum that studies of ambient background levels and his results posed for EPA staff: "[H]ealth effects well may be overestimated in the U.S. Environmental Protection Agency (EPA) risk assessment if [filtered air] is used as the background control" (Adams 2006a, p. 135).

(e) The policy-relevant ozone exposure wave pattern

Previous research has shown that triangular-wave exposures cause earlier respiratory effects than square-wave exposures of the same time-weighted average concentration, and EPA agrees that triangular-wave exposures are more realistic.[80] A principal purpose of the study design in Adams (2006a) was to compare effects under both wave forms to determine whether differences across

---

[78] In the reports published from EPA-sponsored controlled human exposure studies, filtered air (i.e., zero ppb) is used as the baseline for comparisons (e.g., Hazucha et al. 1992; Horstman et al. 1995; McDonnell et al. 1991). None of these studies, however, tested concentrations lower than 0.08 ppm.

[79] In its Response to Comments, "EPA rejects NAM's contention that the Brown Memorandum exemplifies any violation of the information quality standard of objectivity" (U.S. Environmental Protection Agency 2008e, p. 21). In the remainder of the text expounding on this "rejection," EPA argues by *non sequitur*: Brown (2007a) is objective because the sponsor of Adams (2006a) asked that Adams (2006a) be included in the Criteria Document.

[80] EPA agrees that triangular exposures "more closely mimic typical ambient O_3 exposure patterns." See EPA (2007m, p. 3-81).

wave forms that had been observed at concentrations 0.08 ppm and higher also would be present at lower concentrations. For three square-wave concentrations (0.08, 0.06, and 0.04 ppm), Adams devised triangular-wave exposure patterns with the same total exposures, but with higher peaks [0.15, 0.09, and 0.05 ppm) and no exposures less than 0.03 ppm. If the pattern of exposure mattered at these lower concentrations, then stronger effects would be observed with triangular-wave than square-wave exposures.

Adams found that FEV_1 decrements and total symptom scores were significantly greater for triangular-wave exposures after 4.6 and 5.6 hours at 0.08 ppm, but not for 6.6 hours.[81] These results were consistent with results Adams had previously obtained at 0.08 ppm (Adams 2003) and Hazucha et al. (1992) had previously obtained at 0.12 ppm (which Adams also confirmed [(Adams 2006b)]. However, Adams did not observe statistically significant differences between wave patterns at 0.06 or 0.04 ppm. In short, the differential effect of wave-pattern that is detectable at 0.08 ppm and greater concentrations is not apparent at 0.06 ppm and below.

(f) CASAC's "support" for Brown (2007a) is technically infeasible and contradicted by the recollections of some of the principals

Brown (2007a) suggests that the issue is moot because CASAC endorsed his statistical approach:

> On the March 5, 2007 teleconference, members of the CASAC O_3 Panel noted the very conservative nature of the statistical test used by Adams to evaluate the research questions posed by the author. These same CASAC Panel members also supported the approach adopted in the OAQPS Staff Paper to evaluate the statistical significance of O_3-related lung function responses associated with pre- versus postexposure responses. The CASAC Panel members also supported the use of the paired t test approach as the preferred method for analyzing the pre- minus postexposure lung function responses (Brown 2007a, p. 5).

The basis for Brown's claim is hardly self-evident. First, CASAC never reviewed Brown's January 2007 draft memorandum (Brown 2007b). It was not placed in the docket (a requirement for transmittal to CASAC), nor was it on the agenda for CASAC's March 5, 2007, teleconference scheduled to review the draft

[81] Adams used the same statistical methods to adjust for multiple comparisons that Brown (2007a) discarded as "too conservative" with respect to Type I error.

final Staff Paper.[82] Indeed, Brown's January 2007 draft is dated after the release of the draft final Staff Paper (December 2006) even though the draft final Staff Paper includes results from Brown's unpublished January 2007 draft. The earliest CASAC could have seen Brown's reanalysis is June 20, 2007, the day Administrator Johnson signed the proposed rule and EPA placed Brown's finished work product into the rulemaking docket. By this date, CASAC's review was over.

Second, the documentary record indicates that CASAC devoted very little time to this statistical controversy. It appears that CASAC was completely unfamiliar with it until about 15 minutes before the March 5, 2007 conference call. That's when they were provided copies of Adams' and Smith's public comments.[83] Indeed, it appears that EPA worked hard to limit CASAC's exposure to the controversy. Adams and Smith were two of 10 public commenters shoehorned into a 30-minute slot.[84] Under these extraordinary conditions, it would have been quite a remarkable feat for CASAC to digest a pair of oral presentations supplemented by written versions supplied 15 minutes before the conference call began, again review EPA's limited and nontransparent presentation in the final Staff Paper, debate the merits of the competing position, and reach a conclusion – all in the space of maybe an hour -- knowing that the Agency's deadline for disseminating the final Staff Paper was only a couple weeks away.

Third, EPA claims in both its Response to Comments and the preamble to the final rule that it was Smith's public comment to CASAC that created the impetus for EPA's reanalysis of Adams' data.[85] If that were so – and the existence

---

[82] The draft final Staff Paper was made available for public comment on December 27, 2006 (71 Fed. Reg. 77742). The CASAC teleconference to review it was announced on February 5, 2007 (72 Fed. Reg. 25289-5290).

[83] Adams' and Smith's written comments were provide to CASAC by email at 12:45 pm. See the transmittal email from Fred Butterworth to CASAC panel members, Docket No. EPA-HQ-OAR-2005-0172-0075.

[84] The meeting agenda is found at Docket No. EPA-HQ-OAR-2005-0172-0084.1. Public commenters were scheduled from 1:30 pm until 2:00 pm.

[85] From the Response to Comments (U.S. Environmental Protection Agency 2008e, p. 22) "[I]t was a public commenter [i.e., Smith (2007b)] that first placed the analysis of FEV1 responses following exposure to 0.06 ppm O_3 versus filtered air in the public rulemaking docket." From the preamble to the final rule (U.S. Environmental Protection Agency 2008b, p. 16455): "EPA notes that its reanalysis of the Adams (2006)

of Brown's January 2007 draft proves beyond any doubt that the claim is false – then CASAC could not have reviewed the matter carefully enough to "support" the EPA staff position.

Fourth, in their public comments to EPA, Adams and Smith both objected to Brown's claim that CASAC had endorsed his work during the March 5, 2007, teleconference. Neither of them recalled any such expression of support, and no expressions of support can be found in CASAC's March 26, 2007, letter review of the final Staff Paper (Henderson 2007b). The only relevant statement in this letter is a comment from panel member Sverre Vedal objecting to the statistical methods in the draft final EPA Staff Paper:

> [EPA's] 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... (Henderson (2007b, p. C-30)).

The EPA staff is undeterred, however. Brown wraps his work in an imaginary CASAC endorsement. EPA staff then recycle Brown's unsupportable claim in the staff's Response to Comments (U.S. Environmental Protection Agency 2008e, p. 21) and, in the "voice" of Administrator Johnson, in the preamble to the final rule (U.S. Environmental Protection Agency 2008b, p. 16455/16451). Neither of these official Agency documents provides evidence that CASAC actually reviewed the matter beyond hearing a pair of three-minute presentations during its March 2007 conference call. Now that EPA has been challenged via our RFC, EPA has a very strong incentive to publicize such evidence if it exists, but in its Response to Comments EPA does not do so. EPA's response to our claim of information quality error is to attempt to cover it up by committing new information quality error.

3. Gerrymandering the scientific record

In response to EPA's reply, we've noticed that other public commenters expressed similar concerns about the possibility of systemic bias in the inclusion and exclusion of scientific studies. In a comment prepared on behalf of the Utility Air Regulatory Group (UARG), scientists at the Gradient Corporation identified 30 epidemiological studies published between 2000 and 2007 that EPA did not include in its scientific database (Gradient Corporation 2007, pp. A-1 to A-21). Consistent with the Iron Law, we have been unable to locate a single study that

---

study was prepared in response to the issues and analysis raised by a public commenter who made a presentation to the CASAC Panel at its March 5, 2007 teleconference."

arguably pushes the ozone risk envelope outward and was excluded from the scientific record.

There is ample evidence from both EPA documents and CASAC reports that both EPA staff and CASAC were primarily interested in research papers purporting to show a positive association between ozone and health effects. Thus, in addition to the problem of the "market supply" problem of publication bias (covered in Section IIIA.2 above), EPA's scientific record is contaminated by a matching "market demand" problem: only scientific evidence supporting EPA staff and predominant CASAC members' opinions about what policy the Administrator ought to choose were relevant to EPA's ozone review.

(a) EPA staff risk assessment methods show a preference for research showing positive effects

In December 2005, separate from the ozone review, EPA Deputy Administrator Marcus Peacock ordered a "top-down review" of the NAAQS standard-setting process. It appears that senior EPA officials had concluded that the existing process was not serving their needs. In the language of information quality, the process lacked adequate utility. Peacock's Memorandum does not reference EPA's Information Quality Guidelines, but nonetheless it refers to important information quality principles. For example, the Memorandum established as a presumptive norm that the EPA staff scientific record must be unbiased:

> The current NAAQS process has been in place for over 20 years, with some aspects required by law, and therefore not amenable to changes except through new legislation. Other important aspects of the NAAQS process, however, are discretionary -- the agency has established practices that set parameters for how science supports decision making. The Administrator is interested in determining whether those practices reflect the most rigorous, up-to-date, and unbiased scientific standards and methods (Peacock 2005, p. 1, emphasis added).

The Memorandum also reinforces the Administrator's desire that science be distinguished from policy in risk assessment, and in doing so strongly implies that EPA's Offices of Research and Development (ORD) and Air and Radiation (OAR) had persistently failed to make such distinctions. The assistant administrators for ORD and OAR were directed to establish a senior-level staff working group to solve this problem:

> In addition, the working group should focus on the nexus between scientific analysis and standard setting, including the degree to which we are successful in separating the exposition of scientific information from

> the development of risk management strategies and policy judgments (Peacock 2005, p. 2, emphasis added).

If senior EPA officials had been satisfied with the objectivity of the scientific information they were getting from Agency staff, there would have been no need to describe the initiative in these terms. In addition, they would not have encountered the strident opposition of current and former CASAC members, some of whom saw in the initiative a diminution of their ability to indirectly make policy decisions through their ostensibly scientific review function (Henderson 2008b; Vu 2005a, 2005b).

(b) Some CASAC panel members prefer research showing positive effects

CASAC members have not been shy about sharing strong policy preferences for more stringent NAAQS standards, and these views were known when they were recruited to serve on the panel. One CASAC member publicly opined that the ozone NAAQS ought to be more stringent and that the Administrator's most recent decision revising the particulate matter standard was illegal.[86] Another CASAC panel member participated in a process that in 2000 recommended an Air Quality Guideline for Europe of 120 $\mu g/m^3$ (~0.06 ppm) averaged over 8 hours,[87] and which in 2005 recommended that the AQG be lowered to 100 $\mu g/m_3$ (~0.05 ppm).[88] While these policy preferences often are

---

[86] Pinkerton et al. (2007): "To protect the nation's health, it is imperative that the EPA take action to issue a more stringent standard for ozone pollution." "We find the EPA posturing over scientific uncertainty to be disingenuous, uncompelling, and, ultimately, in violation of the Clean Air Act." The editorial acknowledges that co-author John Balmes was at the time a CASAC member who had been paid $52.80 per hour for approximately 25 hours of work over two years serving on the committee.

[87] See World Health Organization (2000, p. 33). A WHO Air Quality Guideline is similar to a primary NAAQS standard. It is a value that "provides a concentration below which no adverse effects or … nuisance or indirect health significance are expected, although it does not guarantee the absolute exclusion of effects at concentrations below the given value" (p. 42).

[88] See World Health Organization (2006, pp. 14-15). Note that the precision in these recommendations appears to be ± 10 $\mu g/m^3$ (~ 0.005 ppm). CASAC recommended that the Administrator set the primary NAAQS with precision ± 0.0005 ppm. See, e.g., the comments of Michael Kleinman on the 2nd draft Staff Paper (Henderson 2006c, p. D-33), and the CASAC letter opposing the Administrator's final decision (Henderson 2008a).

couched in scientific language, it is impossible to miss their policy content.[89] With regard to the scientific database, some CASAC members have openly called on EPA to include in the Criteria Document only those studies supporting the conclusion that ozone exposure below the current NAAQS poses significant human health risks[90] even though the panel as a whole advised EPA that "both positive and negative studies be given the same careful consideration."[91]

(c) Gray literature and "personal communications"

Several times in the Criteria Document, EPA cites as a scientific reference a conference or symposium presentation that was never published in a refereed journal.[92] It is especially noteworthy that EPA cites "Bell et al. (2006)" as the source for the strong claim that "if a population threshold existed for mortality, it would likely fall below a 24-h avg O_3 concentration of 15 ppb" (U.S. Environmental Protection Agency 2006a, p. 8-43). There is no scientific reference in the Criteria Document; "Bell et al. (2006)" is a personal communication between EPA staff and the lead author of an EPA-funded study who informally transmitted unpublished results in response to a staff query.[93]

---

[89] Kleinman's recommendation, cited in footnote 88, is clearly a mix of science and policy: "It would be appropriate to restate the current standard to 3 significant figures which is consistent with the precision of current monitoring devices and which will improve the margin of safety by eliminating 'rounding up' to 0.084." See also, e.g., comments by panel member Cowling, Lippmann, and Russell in Henderson (2005a). Cowling seems to have understood his job was to assist EPA staff in persuading the Administrator to endorse the staff's policy views; see Henderson (2005b, p. D-3). After the final rule was promulgated, CASAC sent what it called "unsolicited advice" stating that the Administrator's decision to set the NAAQS at 0.075 ppm was not "sufficiently protective" and characterizing their collective policy judgment as a "consensus scientific opinion" (Henderson 2008c, p. 2, emphasis added).

[90] See, e.g., individual comments by panel members Balmes, Lippman, and Miller, and the joint comment by panel members Legge, Hanson, Poirot, and Cowling in Henderson (2005a),

[91] See Henderson (2006a, p. 1; more detail on pp. 3 and 6).

[92] Gray literature in the Criteria Document includes Linn et al., 1983b; Lattimer et al., 1984; Selwin et al., 1985) Hogsett et al., 1989; Folinsbee and Hanucha, 1989; Spektor and Lippman, 1991; Tingey et al., 1991; Lebowitz et al., 1991; Linn et al., 1992; Laskin et al., 1996; and Sarwar et al., 2001.

[93] See (U.S. Environmental Protection Agency 2006a, pp. 7-179 and 178-183), citing "Bell, M. L. (2006) Community-specific maximum likelihood estimates of O_3-related excess risk in mortality for the NMMAPS U.S. 95 communities study [personal

In its Response to Comments, EPA claims that the Agency "can not include in its assessment results that were not reported" (U.S. Environmental Protection Agency 2008e, p. 33). EPA's reliance on unpublished data and results obtained through personal communications with Agency-funded researchers is inconsistent with that claim.

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 bias appears to intensify in the progression from Criteria Document to Staff Paper to NPRM.

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

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

In our RFC, we said the inclusion of EPA's reanalysis of the Adams (2006a) data was evidence of purposeful bias because the reanalysis extracts selected data to jury rig support for Agency staff policy recommendations. It is a violation of information quality principles to choose a conclusion first, then fill in behind with selected data and contrived analysis to "support" it. A risk assessment performed this way cannot be unbiased, either in substance or in presentation.

In its Response to Comments, EPA "rejects" our contention that this is what Agency staff actually did (U.S. Environmental Protection Agency 2008e, pp. 21-22). We have already documented in Section III.B.2 the extraordinary efforts that EPA staff expended to rebut the statistical analysis in Adams (2006a). We also have shown that EPA's explanation for why it performed the reanalysis is false, that the Agency's recitation of the facts is both highly selective and self-serving, and that it has claimed CASAC's endorsement for its analysis despite CASAC review concluding before the reanalysis was completed. In short, EPA's

---

communication with attachments to Jee Young Kim]. New Haven, CT: Yale University School of Forestry and Environmental Studies; January 6." The Criteria Document also cites a personal communication with EPA-funded NYU assistant professor and EPA Science Advisory Board staff member Kazuhiko Ito (p. 7-185).

response to our claim of information quality error consists of disseminating new information quality errors.

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

In our RFC, we objected to the EPA staff practice of drawing inferences from individual subjects in controlled human exposure studies when group mean effects are statistically nonsignificant (National Association of Manufacturers 2007, p. 19). The EPA staff's stated justification in the Staff Paper for cherry-picking Adams' data was a fishing expedition: EPA said "responses during the 0.06 ppm O_3 exposures appear to diverge from responses for filtered-air and 0.04 ppm O_3" in a manner that "is suggestive of a significant effect on FEV_1." EPA staff inferred that high interindividual variability combined with a "cursory evaluation" of Adams' newest data "strongly suggested that exposure to 0.06 ppm O_3 causes small group mean FEV_1 decrements in healthy adults with some individuals having notable effects," in this case FEV_1 decrements exceeding 10% (U.S. Environmental Protection Agency 2007f, pp. 3-8 to 3-9). EPA staff could not convert this "strong suggestion" into putative evidence until they discarded most of Adams data and applied their short-cut statistical procedure to the remnant.

In its Response to Comments, EPA repeats the fiction begun in the Criteria Document that its concern about the results in Adams (2006a) occurred because of apparently surprising interindividual variability in FEV_1 responses after 6.6 hours' exposure under exercise at 0.06 ppm (U.S. Environmental Protection Agency 2008e, p. 21). EPA has known about high interindividual variability in FEV_1 responses in controlled human studies for more than three decades. In the 1996 Criteria Document, EPA staff spent two pages summarizing interindividual variability observed in studies dating from 1972. In one EPA-funded study, ozone had accounted for only 31% of the of variance in FEV_1, "clearly demonstrating the importance of as yet undefined individual characteristics that determine responsiveness to O_3" (U.S. Environmental Protection Agency 1996a, p. 7-13). In the 2006 Criteria Document, no new research is cited attempting to explain this phenomenon; interindividual variability is simply characterized as "wide" and "considerable" (U.S. Environmental Protection Agency 2006a, pp. 8-16 to 18-18).[94]

[94] Intra-individual variation also appears to be large. Brown (2007a) analyzed a small subset of the data in Adams (2006a) in order to maximize the likelihood that differences in response could be interpreted as statistically significant. However, the

EPA categorizes individual FEV_1 decrements as "large" (≥ 20%), "moderate" (>10 but < 20%), "small" (3 to 10%), and "none" (± 3%) (U.S. Environmental Protection Agency 2006a, pp. 8-67 to 68-68). The group mean FEV_1 decrements reported by Adams (2006a) (and, incidentally, reproduced by EPA staff (Brown 2007a), fall within the "none" category irrespective of whether background is assumed to be 0.04 ppm (1.5%) or 0 ppm (2.8%). In a policy paper issued in 2000, the American Thoracic Society (ATS) noted the existence of a graded classification scheme for FEV_1 decrements issued by EPA in 1989, but commented that "[t]his classification has not been validated for acceptability or against other measures" (American Thoracic Society 2000, p. 671). ATS says nothing about EPA's 1996 graded scheme, which the Agency recycles in its 2006 Criteria Document.[95] Even if EPA's current graded scheme is assumed to be valid, only three of 60 subject-exposure pairs in the Adams' cohort experienced an FEV_1 decrement exceeding 10% after 6.6 hours of personal (not ambient) exposure to an average ozone concentration of 0.06 ppm -- one of 30 subjects for the square-wave test, and two of 30 subjects for the triangular-wave test (Brown (2007a, Attachment 1). These individual subjects – the only ones with responses in the "moderate" category -- drive the EPA staff reanalysis and provide the foundation for their reinterpretation of Adams (2006a) as displaying statistically significant decrements in FEV_1.[96]

3. Mischaracterization of results

In our RFC, we noted that scientific results can be misrepresented many ways, and we said several of these ways were evident in EPA's risk assessment documents.

---

correlation in responses after 6.6 hours exposure to 0.06 ppm ozone under the square- and triangular-wave protocols was only 0.48.

[95] EPA ignores the ATS caveat that EPA's 1989 scheme was not validated, then asserts that the Agency's latest graded scheme "appears to be valid and reasonable even in the context of the new ATS statement" (U.S. Environmental Protection Agency 2006a, p. 8-66).

[96] We are aware that CASAC encouraged EPA to cherry-pick data from the Adams cohort. The record shows that CASAC's motive was to advance its members' policy views. See Henderson (2006c, pp. 3-4, emphasis in original): "Adverse lung function effects were also observed in some individuals at 0.06 ppm (Adams, 2006[a]). These results indicate that the current ozone standard of 0.08 ppm is not sufficiently health-protective with an adequate margin of safety."

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

In our RFC, we said that EPA cited in this review many of the same studies the Agency had cited in its 1997 decision, but portrayed them as representing "new" scientific information.[97] For example, many of the controlled human studies EPA cites in the Criteria Document EPA also cited in the 1996 Criteria Document. These pre-1997 studies could be used now to say that the health risks posed by ozone are unchanged; after all, the scientific content of these studies cannot have changed. However, these studies cannot be used to support a claim that the health risks posed by ozone are more serious. Before using them to support a different risk characterization, EPA must show either that these studies contained previously unrecognized errors or that EPA had misinterpreted them. EPA did not do this; the EPA staff's "integrated synthesis" approach allows it subtly and nontransparently to reinterpret the scientific content of pre-1997 studies. Without such transparency, the public cannot test whether the EPA staff's portrayal of the science is substantively or presentationally objective.

We said in our RFC that it was misleading for EPA to confuse "old" and "new" scientific information in this manner. We said "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 (National Association of Manufacturers 2007, p. 20). We noted that a reanalysis of an "old" study" constituted "new" science (p. 20, footnote 10). We also invited EPA to identify any pre-1997 study if "the Agency has learned about a material error" or discovered an error in its interpretation (p. 21).

In its Response to Comments, EPA says it "disagrees" with us "on both legal and scientific grounds" (U.S. Environmental Protection Agency 2008e, p. 157). EPA's legal ground for disagreement is facially suspect; the text consists solely of a restatement of the relevant law, including the very provision that most contradicts the Agency's position:

> Section 108 calls for the air quality criteria to 'accurately reflect the latest scientific knowledge useful in indicating the kinds and extent of all identifiable effects on public health or welfare' (U.S. Environmental Protection Agency 2008e, p. 157, emphasis added).

---

[97] For convenience, we use the term "pre-1997" as shorthand for those studies EPA included in the scientific database for the 1997 decision. We do not intend it to mean a literal dated demarcation.

Under EPA's interpretation of the law, the adjective "latest" is superfluous.

EPA's scientific ground for disagreement is that the staff chose an analytic framework that expressly permits it to reinterpret pre-1997 science differently from how they interpreted it in 1997:

> EPA implements this charge by reviewing the newest scientific information, and conducting this review not in isolation but by synthesizing and integrating the newest information with the prior scientific knowledge. An integrated synthesis of the entire body of evidence allows all of the evidence to be evaluated in context, without artificially segregating new from old information. It allows EPA to draw the most appropriate implications and conclusions from the evidence when seen as a whole (U.S. Environmental Protection Agency 2008e, p. 157).

This ignores the statutory context for decision-making under the Clean Air Act. The Administrator's task is to decide whether to revise the existing ozone standard, not to promulgate a brand new one. Thus, if the scientific record is going to have utility for that decision, it must segregate new from old information. Indeed, the rationale for the Administrator's proposed decision is segregated precisely this way: the Administrator first considered whether the existing standard was requisite to protect public health with an adequate margin of safety, and second the Administrator considered how much to lower it.[98]

EPA's Response to Comments further mischaracterizes the implications of distinguishing old from new science, and in doing so, makes surprisingly transparent the staff's desire to be able to reinterpret old science to meet new needs. Such a distinction would "call for freezing our understanding of the information gained from the 'old' studies," which is true only in part. Maintaining a clear distinction between "old" and "new" studies would in no way impede EPA staff from highlighting errors they have discovered errors in these 'old' studies, or identifying errors in their prior interpretation. What such a distinction would so, and which is highly desirable for substantive and presentational objectivity, is deter EPA staff from reinterpreting "old" studies in nontransparent ways.

Contrary to EPA's protests, such an approach is clearly "grounded in scientific principles" for it mimics almost exactly how scientists use and build

---

[98] In the preamble to the NPRM, the first step is set forth in Section II.C., with a conclusion in Section II.C.4. The second step is set forth in Section II.D., with a conclusion in Section II.E. See EPA (2007h).

upon prior literature. Never does the editor of a scholarly journal ask scientists to perform a de novo review of everything that precedes their submitted manuscript. Rather, they are required to summarize that literature briefly to provide a foundation for their work, and they are expected to clearly highlight any instance in which they believe the literature contains error or it has been incorrectly interpreted. EPA says our model is "neither required nor appropriate." Clearly, it is both.

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

In our RFC, we noted that in the NPRM EPA stated that results from "numerous" multi-city and single-city studies show that the associations 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 2007h, p. 37839). We noted that these "numerous" studies consist of the NMMAPS studies, and that the associations in these studies "do not appear to be changed" primarily because they do not measure $PM_{2.5}$.

We also noted two other examples of this form of bias: EPA's reanalysis of Adams (2006a), which we have already covered quite extensively, and EPA's misinterpretation of studies by Moolgavkar and coworkers (Moolgavkar 2000; Moolgavkar et al. 1995). We noted that Moolgavkar had disagreed with how EPA staff used his work (Moolgavkar 2007, pp. 4-5), and EPA staff has ignored these disagreements.

In its Response to Comments, EPA continues to deny that it has incorrectly interpreted Moolgavkar 's work (U.S. Environmental Protection Agency 2008e, pp. 54-55). EPA does acknowledge, however, having not included negative results reported by Moolgavkar – which is precisely the point.

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

In our RFC, we said EPA staff interpreted the literature as showing ambient ozone monitoring provided a satisfactory proxy for personal exposure. This is expressed most succinctly in the Staff Paper, which claimed that

> 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, p. 3-41).

This is a strange construction. Whether average ambient ozone is correlated with average personal exposure matters only if the risks posed by ozone are the result of averages. Yet, the most peculiar aspect of the EPA staff claim is the authority cited to provide the scientific foundation. Neither of the studies referenced as the basis for this conclusion (Sarnat et al. 2005; Sarnat et al. 2001) actually make any such claim. Both studies say that ambient $PM_{2.5}$ but not ambient ozone is correlated with personal ozone exposure, and the researchers believe this is true because ambient ozone is a surrogate for personal $PM_{2.5}$. EPA asserts, but never explains, how these studies show that ambient ozone concentrations may serve as valid surrogates for personal ozone exposure.

(d) Selective and misleading citation

In our RFC, we provided numerous examples in which EPA cited CASAC selectively in the NPRM such that the result was a biased presentation of the panel's scientific review. We listed examples from the NPRM and added the text from the relevant CASAC document that EPA left out (National Association of Manufacturers 2007, pp. 22-27).[99]

In its Response to Comments, EPA "strongly denies" our claims and says that each of these issues was "thoroughly discussed in the NPRM" (U.S. Environmental Protection Agency 2008e, p. 152). However, each of the selective citations we listed came from the NPRM, making EPA's rebuttal technically infeasible. We agree wholeheartedly with EPA that it is "not required to quote verbatim all of an important comment made by the CASAC O_3 Panel," and that "[d]oing so in the Staff Paper or NPRM could have the effect of obstructing clear communication of the concepts involved rather than facilitating communication." However, the issue we raised was the lack of presentational objectivity in EPA's NPRM. Our complaint, which EPA does not rebut, is that the NPRM provides the public a severely biased and self-serving picture because it

[99] The task of discriminating between CASAC's scientific review and its policy advice is admittedly challenging. As we note in Section V beginning on page 131, this task was made immeasurably more difficult by EPA's decision not to ask CASAC to clearly distinguish between its scientific review and its policy advice, and CASAC's own decision not to be transparent about such distinctions.

quotes only the underlined text from a CASAC comment and excludes the context:[100]

- 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).
  - **EPA's defense is that it "addressed" CASAC's concerns in the Staff Paper and in Section II.D.4.a of the NPRM (U.S. Environmental Protection Agency 2008e, pp. 152-153). In the Agency's lexicon, to "address" a concern means to "discuss" or "consider" it, not to "resolve" or "reconcile" it.**
  - **The Staff Paper is irrelevant to our information quality complaint about the presentational objectivity of the NPRM, and the subsection of the NPRM EPA says is responsive appears 39 to 41 pages of dense Federal Register text later. A review of that text shows that EPA did not in fact "address" CASAC's concerns.**
  - **Whereas CASAC said the mortality studies "do not seem to allow us to confidently attribute observed [mortality]" to ozone, that is exactly what EPA did: "A standard set at [0.074 ppm] is estimated to reduce nonaccidental mortality [from ozone exposure] by**

[100] EPA also shifts the burden back to CASAC, with the added twist of insisting that it obey a warp in the space-time continuum: "If these issues had not been fully addressed, the CASAC O_3 Panel would have noted that in its final review of the Staff Paper, but it did not" (U.S. Environmental Protection Agency 2008e, p. 152). In short, CASAC's failure to propagate its every unresolved concern throughout each report in the series lets EPA off the hook. CASAC further failed to anticipate how EPA would cite its comments selectively in the preamble to the NPRM, which of course was published after CASAC's review was completed.

about 10 to 40 percent"(U.S. Environmental Protection Agency 2007h, p. 37877).

- 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, p. 3).
    - **EPA's defense is that it "addressed" CASAC's concerns in several sections of the Staff Paper and in Section II.D.1 of the NPRM (U.S. Environmental Protection Agency 2008e, pp. 153-154). As indicated above, in EPA-speak "address" means "discuss" or "consider," not "resolve" or "reconcile."**
    - **The NPRM section referenced by EPA (U.S. Environmental Protection Agency 2007h, p. 37872), which appears 36 dense Federal register pages later than the selective citation from CASAC, is irrelevant. It consists of the Administrator's policy decision to retain ozone as the indicator for photochemical oxidant air pollution; it has nothing to do with CASAC's expressed scientific concerns about EPA's inferences about ozone-induced mortality.**

In our RFC, we said EPA had excluded from the NPRM crucial scientific comments from CASAC that did not support EPA's exposition of the data (National Association of Manufacturers 2007, pp. 25-27). We identified four such examples:

- The lack of correlation between ambient ozone levels (upon which all estimates of health risk depend) and personal exposures (upon which actual health risk must depend), especially among the elderly and infirm in which the alleged mortality effects from ozone are assumed concentrated.

- The inability to detect a threshold in concentration-response because of measurement error implied by the use of ambient ozone levels instead of personal exposures.
- The need for sensitivity analysis in the estimate of effects at different values for background, rather than the imposition of a policy-charged PRB.
- The possibility that ambient ozone serves as a surrogate for other pollutants, most notably $PM_{2.5}$.

For each issue we identified both the critical element of the CASAC comment and (unlike EPA) provided its context.

In its Response to Comments, EPA "rejects" our examples, saying that in each case the Agency 's exposition in the NPRM is complete and unbiased (U.S. Environmental Protection Agency 2008e, pp. 155-156). In each case, however, EPA actually cites irrelevant text from the Staff Paper, as if presentational objectivity in the NPRM is achieved as long as the Agency can point to text somewhere else in a subordinate document. EPA also cites pages in the NPRM where it says this material "can be found" or is "highlighted" (e.g., 72 *Federal Register* 37878), but EPA's reference concerns the Administrator's policy determinations, not an exposition of science.

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

In our RFC, we noted that EPA claims controlled human exposure studies provide compelling evidence that ozone exposure below the current ozone NAAQS causes lung function decrements, inflammation, and respiratory infection (National Association of Manufacturers 2007, p. 27). We also noted that the vast majority of the studies that EPA cites involve exposures at or above the current standard. EPA provided only a quasi-policy rationale for its ostensibly scientific inference, but that is impermissible under information quality principles. Policy officials have discretion over policy statements, but scientific statements must be supported by science.

In its Response to Comments, EPA denies that it drew inferences unsupported by the Adams (2006a) data (U.S. Environmental Protection Agency 2008e, p. 22). However, EPA does not deny that the group mean decrements in FEV_1 of 1.5% (compared to 0.04 ppm) and 2.8% (compared to 0 ppm) that Adams observed after 6.6 hours of exposure to 0.06 ppm is elsewhere characterized by the Agency as "within normal range (+3%)" (U.S. Environmental Protection Agency 2006a, p. 8-76, Table 8-2). EPA also does not deny that the sample standard deviation in FEV_1 responses after 6.6 hours' exposure to filtered air (i.e.,

zero ppm ozone) also was 3%.[101] EPA relies exclusively and completely on its eleventh-hour, never peer reviewed reanalysis of a selected fraction of Adams' data, having first tortured it to reveal statistical significance by discarding the experimental protocol in which the data were collected (Brown 2007a). EPA staff then characterize zero ozone exposure as "background" in order to try to nudge the FEV_1 decrement into its "small" effect size category (> 3%) – a threshold it still could not achieve without rounding 2.8% upward above the nearest integer.

Despite the obvious relevance and criticality of the EPA staff reanalysis of Adams (2006a) to the Administrator's policy determinations, the entire discussion of the reanalysis in the NPRM consists of a portion of a single paragraph, found at 72 *Federal Register* 37828, column 2). This discussion is peculiarly supplemented by footnotes (numbered 14-16) that are highly revealing. First, EPA uses passive voice to say that these results "were not included" in Adams (2006a). Second, EPA's reanalysis was truly an eleventh-hour work product (the memorandum for the docket is dated June 14, 2007, just six days before the Administrator signed the NPRM, and actually placed in the docket the same day as the NPRM). Third, what attracted EPA staff attention was that "7 percent" of Adams' subjects (i.e., two out of 30) experienced FEV_1 reductions after 6.6 hours exposure to 0.06 ppm ozone that they describe as "notable" (an undefined term), and then only when 0 ppm is used the presumptive.

Even this limited degree of transparency EPA provided grudgingly. The interagency review draft of the NPRM, dated May 22, 2007, contains no reference to EPA's reanalysis, which EPA had kept hidden since at least December 2006 when Brown alluded to results at a public meeting sponsored by EPA (Brown 2006). EPA even misleadingly tried to claim that its reanalysis was motivated by the need to "confirm" a public comment submitted to CASAC in early March 2007 (Smith 2007b). Smith tried had tried in vain to persuade CASAC to investigate more carefully the fundamentally flawed statistical analysis summarized in the Staff Paper.[102]

In our RFC, we said that EPA's analysis of clinical data on cardiac effects

---

[101] These data can be found in Brown (2007a, Attachment 1).

[102] The flicker of candor found in the NPRM appears to have been the product of interagency review. The story of EPA's secret reanalysis, and its unrelenting effort to mislead the Administrator and the public about its origin, is documented in Section III.B.2 beginning on page 63. The CASAC teleconference call at which Smith made his appeal occurred just three weeks before the court-ordered deadline for publication of the Staff Paper, so it may well be the case that by this time CASAC was helpless to act.

was 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[103] 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 the NPRM, EPA emphasized the increase in AaPO2 and interpreted this 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 2007h, p. 38734). EPA also was silent about the relevance of the exposure level (0.3 ppm, or 3.75 times greater than the current 8-hour NAAQS), or the uncertainties implied by extrapolating to the population clinical data obtained from a sample of 16.

In its Response to Comments, EPA does not "reject" our concern as it does so many times elsewhere (U.S. Environmental Protection Agency 2008e, p. 23). Instead, EPA says we should be mollified by other language it also used in which the Agency describes the cardiac data as "a very limited body of evidence" with "evidence for some potential plausible mechanisms." Re-reading the NPRM, we see that EPA characterized the cardiac epidemiology as providing "limited evidence suggestive of a potential association," which seems to us to be so qualified by caveats as to be meaningless if taken literally. The problem is that in the Administrator's statement of conclusions on the elements of the primary standard, these caveats are almost completely abandoned and "possibl[e] cardiovascular effects" are cited as evidence (U.S. Environmental Protection Agency 2007h, p. 37870). If scientific evidence this weak is considered "supportive" of a lower primary standard, it is difficult to imagine how weak evidence must be before EPA declines not to rely on it.

(f) Utilizing for one purpose data that were collected for another purpose

In our RFC, we objected to EPA staff's use of the Adams (2006a) data for purposes different that those which were intended by the study design (National Association of Manufacturers 2007, pp. 29-30). EPA staff first focused on the two of Adams' 30 subjects who had with the largest FEV_1 decrements after 6.6 hours of exposure to 0.06 ppm ozone under moderate exercise, and extrapolated their

[103] 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."

responses to the population. EPA staff then asked Adams for a highly restrictive subset of his data, and they proceeded to analyze these data without regard for Adams' study design. EPA staff never sought independent expert review of their analytic procedures, nor did they ever ask CASAC to review their work.[104]

In its Response to Comments, EPA defends this practice several ways (U.S. Environmental Protection Agency 2008e, pp. 97-98). First, EPA says that it performed a similar statistical analysis in support of its 1997 revised ozone standard. However, we have examined both the Criteria Document and the Staff Paper for the 1997 standard (U.S. Environmental Protection Agency 1996a, 1997), and we have found no evidence of an analogous statistical analysis, much less one utilizing paired *t* tests without adjustment for multiple comparisons.

Second, EPA says that it included data from Adams (2006a) because it was "urged" to do so by the American Petroleum Institute (API), which sponsored Adams' study. As we recounted earlier (see Section IIIB.2 beginning on page 63), the record shows that CASAC was first to ask EPA to include Adams (2006a); EPA neglects to mention this vital fact.[105]

Third, and most misleadingly, EPA asserts:

> The health risk assessment for lung function responses was reviewed by the CASAC O_3 Panel and there were no objections expressed by CASAC panel members or by Dr. Adams in either his oral or written comments to EPA concerning EPA's use of the Adams data as part of the basis for estimating the exposure-response relationships used in the health risk assessment (U.S. Environmental Protection Agency 2008e, p. 98, emphasis added).

---

[104] Only a summary is presented in the draft final Staff Paper.

[105] EPA's in-text reference in the Response to Comments for API's "urging" is "(API, 2006)," a reference not included in the bibliography. We infer that this reference is API's public comment dated September 18, 2006, on the second draft Staff Paper (Docket No. EPA-HQ-OAR-2005-0172-0057.1). This public comment says EPA "reasonably incorporates data" from Adams' studies, but notes with obvious concern that "rather than relying on these group mean results, the draft Staff Paper chooses to rely on data from individual subjects," a practice API correctly describes as statistically "invalid" (American Petroleum Institute 2006, pp. 19-20). EPA incorrectly states that Agency staff "obtained the individual data used in the health risk assessment directly from the author" [i.e., Adams] when in fact they sought only a very limited subset of the data set sufficient to perform its constrained statistical test.

Read carefully, is sentence refers only to the unobjectionable part of what EPA staff actually did. It should go without saying that the purpose of performing the research that became Adams (2006a) was precisely to help EPA "estimat[e] the exposure-response relationships used in the health risk assessment." That is a fundamentally different purpose, however, than the purpose for which EPA staff ultimately used it. The purposes of Adams' research were to (1) determine whether there were group mean decrements in pulmonary function at 0.06 and 0.04 ppm ozone, and if so, (2) determine whether these decrements differed by wave pattern. Adams (2006a) shows that pulmonary function decrements were not statistically significant at these lower concentrations, and that there was no difference in effect by wave pattern.

Had EPA staff allowed Adams' research to speak for itself, that would have been the end of the story. They didn't, and it wasn't. EPA staff cherry-picked from Adams' dataset, applied inappropriate statistical methods to make the selected data appear to show a statistically significant effect, and interpreted these results as compelling evidence of ozone health risk at 0.06 ppm. Adams publicly objected to this, and EPA attempts to cover up that fact.

(g) Hypothesizing after the results are known

In our RFC, we suggested that one of the information quality defects in the EPA staff approach is that it was hypothesizing after the results were known – a practice sometimes called "data mining" (National Association of Manufacturers 2007, p. 30). We are unable to locate any reply from EPA in its Response to Comments. EPA staff do not ever examine a health effect and attempt to discern its likely causes to estimate the fraction, if any, attributable to ozone exposure. No other factors matter, for ozone is the sole culprit of interest.

Properly performed hypothesis-testing 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 either exploratory or hypothesis-generating, but never hypothesis-testing. The results of hypothesis-generating research should only be used to guide future hypothesis-testing research, and it never should be used to draw inferences – especially inferences that have significant public policy implications.

4. Study selection bias

In our RFC, we said EPA staff had displayed a systematic preference for studies that show positive associations even among studies that have important information quality limitations (National Association of Manufacturers 2007, p. 30). For example, where several studies were available to estimate effects on

asthmatics, EPA staff consistently selected studies with positive associations with ozone (e.g., Gent et al. 2003; Mortimer et al. 2002) over studies that do not (e.g., Moolgavkar 2000; Schildcrout et al. 2006). Nowhere in EPA's review plan, or in any other regulatory development document, did EPA discuss – much less establish -- an information quality basis for its selections. This bias is transparent when the EPA staff view of data from personal expiratory flow monitors is compared in the case of ozone ("data are reliable") and the case of nitrogen oxides ("data are unreliable"). The same studies are implicated; the only difference is that positive associations were obtained for ozone but not for nitrogen oxides.[106]

In its Response to Comments, EPA "rejects" our complaint, once again confusing having "discussed" or "considered" negative results and studies as equivalent to having taken them seriously (U.S. Environmental Protection Agency 2008e, pp. 55-56). EPA uses a "weight of evidence" approach that enables it to evade clarity and reproducibility – both hallmarks of good information quality practice. As we have noted, however, information quality principles and practices were missing from the ozone NAAQS review from beginning to end.

EPA also invokes as an all-purpose defense the fact that CASAC reviewed several of the documents subject to our RFC (U.S. Environmental Protection Agency 2008e, pp. 81-82). It is true, as EPA says, that the Agency's Information Quality Guidelines say that "if data are subjected to formal, independent, external peer review the information may generally be presumed to be of acceptable objectivity." EPA's guidelines presume, of course, that an agency's "formal, independent, external peer review" actually subjects the document to an information quality review. Peer review that ignores the information quality principle of objectivity cannot possibly ensure objectivity except by chance. As we documented in our RFC, and we reiterate here, EPA did not include information quality principles, in any shape or form, in its charge to CASAC.

EPA replies saying its 2005 Review Plan (U.S. Environmental Protection Agency 2005d) and first draft Health Risk Assessment (U.S. Environmental Protection Agency 2005b)[107] provided the "criteria for selection of studies and

---

[106] See EPA (2007b, p. 3-16) and the discussion in Section III.A.2(d)(vi).

[107] The in-text citation is to "Abt Associates, 2006." Abt is EPA's contractor. We infer that EPA intended to cite the October 2005 first draft Health Risk Assessment, which Abt produced and EPA published as if it were EPA's own work product. The document has footers on each page ascribing authorship to Abt, and it does not include a disclaimer stating that it was distributed solely for per review. Under the terms of

concentration-response relationships" (U.S. Environmental Protection Agency 2008e, pp. 81-82). However, the Review Plan actually contains no criteria for study selection. It is an outline of the review process and a description of the subjects to be addressed and nowhere mentions how studies would be selected for inclusion or exclusion. The first draft Health Risk Assessment contains a section titled "Selection of epidemiological studies" (4.1.5, p. 4-9), which lists the following criteria for study inclusion:

- It is a published, peer-reviewed study that has been evaluated in the draft O_3 AQCD and judged adequate by EPA staff for purposes of inclusion in this risk assessment based on that evaluation.
- It directly measured, rather than estimated, O_3 on a reasonable proportion of the days in the study.
- It either did not rely on Generalized Additive Models (GAMs) using the S-Plus software to estimate C-R functions or has appropriately re-estimated these functions using revised methods.
- For studies of mortality associated with short-term exposure to O_3, the study reported results for the O_3 season.

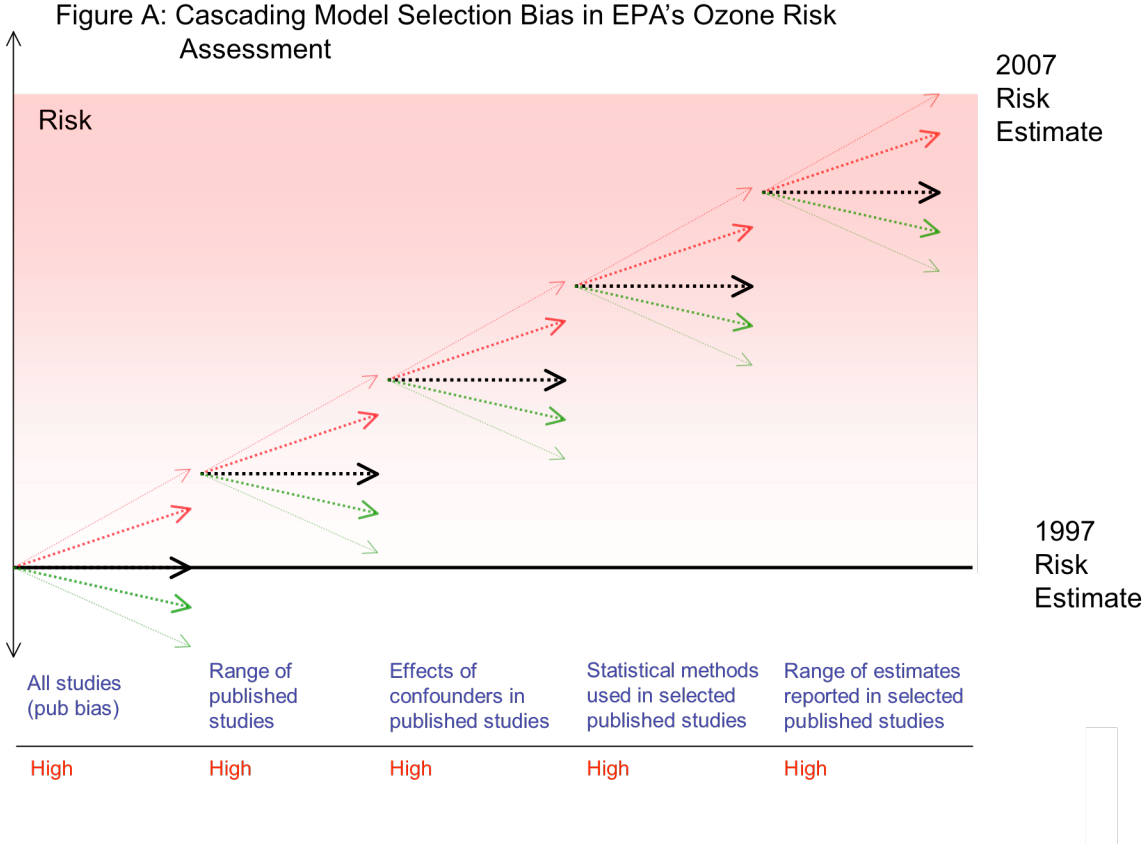
Information quality principles are missing from these criteria, and nonreproducible EPA staff judgment dominates.

5. Model selection bias

In our RFC, we said that EPA staff selected or emphasized models biased in favor of overestimating health risks:

- EPA staff selected models based on criteria other than quality of data or analysis
- EPA staff selected models known to yield upwardly biased risk estimates, such as single-pollutant models that do not control for known confounders
- EPA staff selected models based on statistically convenient but biologically implausible criteria
- EPA staff emphasized results from models known to yield risk estimates that are upwardly biased and more uncertain, such as

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



Generalized Additive Models conducted with insufficient convergence criteria

We used a figure, reproduced above, to illustrate the cascade of bias implied by just some of these practices.

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

We have already shown in the preceding section that EPA staff did not use information quality criteria to select models for risk assessment. In fact, the dominant criterion the staff used was its own non-transparent, non-reproducible, and undefined "judgment." In its Response to Comments, EPA staff defend their "judgment" noting that CASAC "did not express any concerns" about their choice of studies and models (U.S. Environmental Protection Agency 2008e, p. 82). This is hardly surprising, because the CASAC panel was dominated by researchers with strong policy views whose own work EPA was relying upon.[108]

[108] A notable example is Korrick et al. (1998), of which CASAC ozone panel member Frank Speizer is a co-author. Speizer cannot reasonably be expected to have

A peer review cannot be genuinely independent if it is conducted by the same scientists whose work the Agency is summarizing or promoting.

Several commenters recommended that EPA solve the model selection bias problem by adopting by Bayesian model averaging. EPA staff discussed this briefly in the Criteria Document, but they discarded it because it had certain undesirable effects – most notably, the magnitude of estimated effects would be "diluted (i.e., result in smaller coefficients) when variables are highly correlated, as may be the case for air pollution studies" (U.S. Environmental Protection Agency 2006a, p. 7-20).[109] EPA staff discuss the Bayesian model averaging study by Koop and Tole (2004) but discard it because its results "cannot be interpreted meaningfully." This test is related to information quality, but EPA staff do not subject the studies they rely upon to the same rigor. A more plausible explanation is that Koop and Tole found only very small effects. Meanwhile, the principle mortality study EPA staff rely on – Bell, McDermott et al. (2004) – uses a Bayesian data averaging procedure without which the authors could not have reported a statistically significant positive effect. Bayesian methods that push the ozone risk envelope outward are useful and appropriate; Bayesian methods that push the ozone risk envelope inward are not.

This is another practical example of EPA staff use of the Iron Law we presented in Section I.C. Scientific information indicating greater risk pushes the envelope outward; information that is equivocal supports the current location of the envelope; and information indicating lesser risk is discarded. We challenge EPA to refute the Iron Law by providing specific, concrete examples.

(b) Control for known confounders other than air pollution

In our RFC, we alluded to, but perhaps did not make clear, the fundamental information quality defect in the EPA staff's analytic approach. As an example, we listed 12 factors known to cause respiratory effects in asthmatic children: (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 (Sarafino et al. 2001). One way to manage this complexity is to try to estimate the contribution of air pollution (in this case, ozone) while controlling

---

given an objective review of this study or to have refrained from endorsing the EPA staff decision to give it a lot of weight. Moreover, it is difficult to imagine other members of the panel publicly identifying deficiencies in this study or opposing its use.

[109] EPA recycles these objections in its Response to Comments (U.S. Environmental Protection Agency 2008e, p. 44).

for these other factors, because failing to adequately control for known confounders yields upwardly biased estimates of risk. Moreover, such an approach would illuminate the characterization of adversity. It would be problematic, for example, to interpret as adverse physiological effects from ozone no greater than those from benign or positive phenomena. However, we noted that in the ozone epidemiology literature, control for confounders has been spotty, especially in ecologic studies but even in panel studies where individual data are obtained. It is remarkable, for example, that in the panel studies EPA relies on, there is no control for allergens and little control even for medication use.

EPA dismisses our concern about household allergens and exercise confounding the association between asthma and ozone exposure on the ground that they do not vary daily (U.S. Environmental Protection Agency 2008e, p. 42). We would agree if ozone's presumptive contribution to asthma were large relative to allergens, but it is not. Even according to EPA's preferred studies, however, ozone might be responsible for or exacerbate a tiny fraction of asthma cases, Most of the variance in asthma and its symptoms remains unexplained in these models.

Reflecting on EPA's Response to Comments, it now seems obvious that in practice the Envelope Theory requires that as much as possible of any health effect must be ascribed to air pollution (in this case, ozone). EPA staff do not seek to understand a specific health effect and try to discern the most plausible causes and allocate it objectively. The only time that EPA staff face a genuine dilemma is when multiple pollutants are effectively "competing" for a share of the air pollution burden, and in those cases EPA staff is susceptible to the temptation to assign the same health effects to each one.

This analytic defect carries over into the epidemiological research that EPA staff funds and, after publication, relies upon to advance its mission of tightening the NAAQS standards. We cannot find a single EPA-funded research study in EPA's scientific database that is focused on understanding the etiology of a health effect rather than seeking to prove that air pollution is its cause.

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

In our RFC, we criticized EPA for basing its risk estimates on models known to yield upwardly biased estimates of health risk (National Association of Manufacturers 2007, p. 34). In its Response to Comments, EPA "rejects" our assertion that the Agency has done this, then proceeds to defend basing its risk estimates on models known to yield upwardly biased estimates of risk (U.S.

Environmental Protection Agency 2008e, pp. 44-45). EPA defends the use of single-pollutant models on the ground that they are "robust," or some similar formulation such as "fairly robust," "generally robust," or "statistically robust." EPA never defines any of these terms scientifically, but the Agency uses them in an ostensibly scientific context 54 times to describe associations in volume 1 of the Criteria Document, 48 times in the Staff Paper, and 28 times in the NPRM. We're not at all sure what EPA means by "robustness," but we do know that the Agency has not used the term in the same manner as it has been used by the statisticians who pioneered robust methods (e.g., Tukey (1977), Hoaglin et al. (1983).

Elsewhere in its Response to Comments, EPA mischaracterizes our complaint to imply that we think fully disclosing all relevant scientific information and results is a violation of applicable information quality, then proceeds to rebut its mischaracterization (U.S. Environmental Protection Agency 2008e, p. 83). We never made any such claim; we objected to EPA's reliance on models known to be upwardly biased for the purpose of human health risk assessment, not the comprehensive reporting of results. In short, EPA is violating information quality guidelines by purposefully estimating individual risk in a biased manner. With regard to our actual complaint, EPA is silent.

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

In our RFC, we objected to EPA's favorable treatment of several epidemiological studies in which researchers had mined the data to identify the most statistically significant lags and lag structures, then speculated why the results of these mining operations might be biologically meaningful (National Association of Manufacturers 2007, pp. 34-36). While we had no objection to the researchers' use of such exploratory data analysis techniques for EDA purposes, it was disconcerting to note that in some cases researchers drew inferences well beyond what EDA methods permit, and that EPA had treated these inferences as if they were confirmatory rather than exploratory.[110]

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

[110] "Using techniques that adopt specifications on the basis of searches for high R^2 or high t values, is called data-mining, fishing, grubbing or number-crunching. This methodology is described eloquently by [Ronald] Coase: 'if you torture the data long enough, Nature will confess'" (Kennedy 1985, p. 76).

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, pp. 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.

In its Response to Comments, EPA "disagrees" with our characterization, even to the point of ignoring the actual statements of the researchers themselves, which make clear that their statistical analyses were exploratory in nature (U.S. Environmental Protection Agency 2008e, p. 45). EPA infers biological plausibility from statistical significance, rather than using statistical methods to test whether data are consistent with biologically plausible lags.

We emphasized EPA's reliance on Mortimer et al. (2002) as symptomatic of this constellation of information quality defects. The authors used seemingly every conceivable statistical device to discover positive associations: a wide array of lags and lag models; discarding statistically nonsignificant evening post-exposure effects in favor of statistically significant morning pre-exposure effects; then speculating about possible biological mechanisms that might explain their results. This is not controversial as an exercise in exploratory data analysis for the purpose of generating testable hypotheses, but it is completely inappropriate to interpret the results of EDA as confirming biological mechanisms concocted speculatively after the fact. In EPA's exposition, the exploratory nature of the researchers' data mining is downplayed and their results are treated as if they were confirmatory.

6. Assumption of causality

In our RFC, we faulted EPA for basing its conclusions about the causality of statistical associations on policy considerations rather than a plausibly objective scientific procedure. We illustrated a plausibly objective procedure (reproduced again as Figure B below) in which, *ceteris paribus*, effect sizes are treated the same regardless of their signs. We noted that EPA's approach consisted of putting a large policy thumb on the scientific scales:

> 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 (National Association of Manufacturers 2007, pp. 38-39).

We added that EPA's approach is "unambiguously and transparently policy-directed" (p. 38).

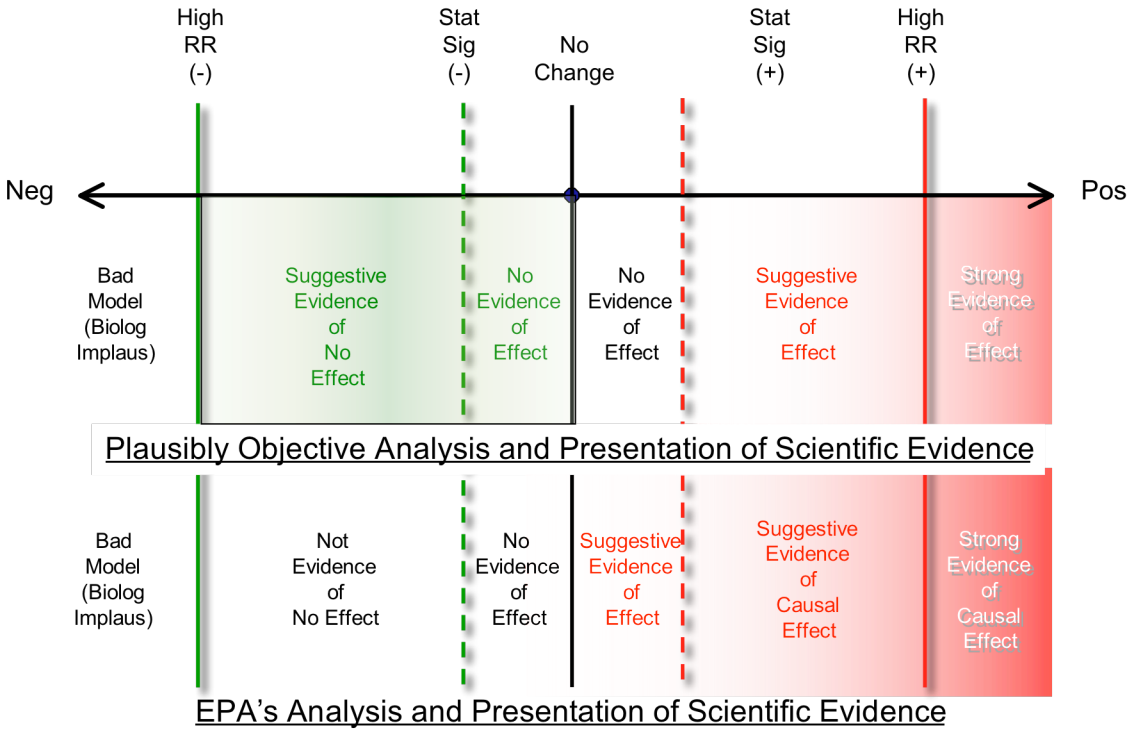
In its Response to Comments, EPA "strongly disagrees" with the latter complaint, but provides only a boilerplate legalistic defense evading the point:

> The critical assessment of epidemiologic evidence presented in [section 7.1.2 of] the Criteria Document is conceptually based upon consideration of salient aspects of the evidence of associations so as to reach fundamental judgments as to the likely causal significance of the observed associations… (U.S. Environmental Protection Agency 2008e, p. 34).

This section of the Criteria Document is a discussion citing the Bradford Hill criteria and other "considerations" EPA staff took into account, including whether associations were "robust." As we noted in section III.C.5(c) on page 102, EPA uses the term "robust" and its variants a hundred times but never defines it.

Just as we said it was in our RFC, the EPA staff's process for determining causality is unambiguously and transparently policy-directed. EPA's Response to Comments appears to misinterpret our complaint to suggest that the Administrator or other policy officials directed the staff to embed policy judgments within their scientific review. We have not found any evidence suggesting such interference. Rather, we see a consistent pattern of EPA staff usurping the decision-making prerogatives of the Administrator and embedding their policy judgments into the science.

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



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

In our RFC, we said EPA's risk assessment lacked objectivity as a matter of policy, rooted in the Agency's narrow mission and implied by the policy views of its career staff (National Association of Manufacturers 2007, pp. 39-40). We noted that this policy is a matter of public record. We cited as our authority a recent EPA Staff Paper on Risk Assessment Principles and Methods that celebrates the staff's practice of producing purposefully biased risk assessments (U.S. Environmental Protection Agency Office of the Science Advisor 2004a). In this Report, EPA staff gave a *pro forma* 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). The Report makes clear, however, that these commitments are subordinate to the greater goal that its risk

assessments be biased in favor of erring on the side of overestimating human health risk, not estimating it objectively:

> 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."

This staff policy explicitly leads to bias that the staff justify on account of the existence of uncertainty and variability:

> Since uncertainty and variability are present in risk assessments, EPA usually incorporates a "high-end" hazard and/or exposure level in order to ensure an adequate margin of safety for most of the potentially exposed, susceptible population, or ecosystem (U.S. Environmental Protection Agency Office of the Science Advisor 2004b, p. 16, emphasis added).

The Clean Air Act delegates to the Administrator, not to his technical staff, the authority to decide what constitutes an "adequate margin of safety." By embedding an "adequate margin of safety" into its risk assessments, EPA staff assures that whatever margin of safety the Administrator chooses, it will be over and above the margin of safety that his staff have already included in its risk assessment and characterization.

In our RFC, we said EPA's ozone risk assessment was faithful to the EPA Staff Paper on Risk Assessment Principles and Methods, and the staff's commitment "op avoid both underestimation of risk and gross overestimation of risk." In its Response to Comments, EPA does not actually deny that its risk assessment adheres to these principles. Rather, EPA simply waves the talisman of the CASAC peer review – and gives a self-serving exposition of the panel's views, at that (U.S. Environmental Protection Agency 2008e, p. 85). EPA notes that CASAC graciously called its risk characterization "well done, balanced and reasonably communicated," but EPA fails to mention any of the important caveats CASAC included in the same paragraph of the same letter:

- "Although a number of issues are raised, their impacts on the estimates have not been thoroughly explored."
- "Additional sensitivity analyses seem warranted."

- "Although the 3 parameter logistic (3PL) model emulates the pattern seen in the five "data points," these points are aggregates of the original data, and may give a misleadingly optimistic picture of the quality of the fit."
- "More importantly, although the problem of model uncertainty is noted it has not been addressed even though methods exist for doing so."
- "Even if only the linear and logistic models were included in the analysis, the error bands around the estimated response probabilities would likely increase to better reflect that uncertainty."
- "In addition, a suggestion to deal with the uncertainties surrounding estimation of PRB, particularly as related to Table 5.5 (for lung function) and Table 5.11 (mortality), would be to change the form of the analyses to assess the impact of the concentration change in the expected number of health effects relative to the current standard. The key advantage of estimating the effect of concentration change is that it does not depend on the choice of the PRB."[111]

As we have noted elsewhere, EPA never asked CASAC to review its scientific work products to ensure that they were objective. The information quality principles that the EPA Staff Paper on EPA Risk Assessment Principles and Practices says the staff is committed to uphold are missing entirely from the panel's Charge, which asks them instead to evaluate its scientific and technical work for "reasonableness." "Objectivity" can be refuted by the application of methods that scientists such as CASAC panel members know well. On the other hand, "reasonableness" is purely a matter of judgment and opinion, and as such, it can never be refuted. Thus, the goal of EPA staff has been to persuade CASAC that their effort has been reasonable, not that the output of that effort is objective.

[111] See Henderson (Henderson 2006c, p. 12). All comments cited here – including the praise cited by EPA – were on the second draft Staff Paper published in August 2006 (U.S. Environmental Protection Agency 2006f). CASAC's review of the final draft Staff Paper apparently was so abbreviated by time constraints that it did not examine the extent to which EPA staff had responded to its concerns. Neither the letter nor the individual comments by CASAC panel members suggests that CASAC actually reviewed Chapter 5, which contains the risk characterization. Because this particular meeting was conducted by teleconference, there is no meeting transcript.

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

In our RFC, we said EPA's risk assessment attributed to ozone health risks that the Agency had previously attributed to other pollutants – most notably, fine PM – through the device of single-pollutant models that exclude control for confounding air pollutants (National Association of Manufacturers 2007, p. 41). In its Response to Comments, EPA points to tables in the risk assessment that provide a range of alternative models (U.S. Environmental Protection Agency 2008e). EPA justifies the use of single-pollutant models for estimating mortality risk on the ground that there is "some evidence" that the effect of PM "may not be very substantial."

EPA's response is fully consistent with the EPA Staff Paper on Risk Assessment Principles and Practices (it is Agency staff policy to not to ever understate risk) and our Envelope Theory of EPA Risk Assessment (all science either points toward high risk or points nowhere at all).

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

In our RFC, we identified several broad information quality errors in EPA's health risk assessment, each of which had the effect of upwardly biasing the Agency's estimate of human health risk.

A. EPA Treats Transient and Reversible Effects as Adverse

We noted that EPA staff defined as "adverse" physiological effects that are transient and reversible (National Association of Manufacturers 2007, p. 42). Such effects have at least a dozen reported triggers (Sarafino et al. 2001), including laughter, which presumably EPA does not intend to count as adverse.

In its Response to Comments, EPA "strongly rejects" our position, claiming that we were contesting "the Administrator's judgments as to when O_3-related effects become regarded as adverse to the health of individuals" (U.S. Environmental Protection Agency 2008e, p. 62, emphasis added). This is false; we contested the EPA staff's characterization of the science of adversity, not any aspect of the Administrator's policy judgment. Any reference to the Administrator's policy judgment is diversionary, for EPA staff has sought to define adversity in technical and scientific terms that are covered by information quality principles; the Administrator's policy judgment is not.

This can be seen in the Criteria Document, for example, where EPA staff devote considerable attention to the task of defining criteria for determining adversity in scientific terms (U.S. Environmental Protection Agency 2006a, pp. 8-

65 to 68-69). The Criteria Document reprints respiratory effect size categories EPA first published in 1997, but neither the CD nor the Staff Paper transparently define adversity. By never defining it, EPA staff implicitly interprets all effects as adverse.[112]

It appears that as a tactical matter, EPA staff relied on CASAC to define adversity in terms of its members' policy views, then recharacterized those policy views as "science" (U.S. Environmental Protection Agency 2007d, Section 2.1). The script can be found in the discussion between EPA staff and CASAC during the CASAC meeting on August 24, 2006 (U.S. Environmental Protection Agency Science Advisory Board Staff Office 2006, pp. 142-150). EPA's Response to Comments implicitly attributes to CASAC the decision to treat transient and reversible effects as adverse, citing this very same discussion (U.S. Environmental Protection Agency 2008e, p. 65).

EPA's Response to Comments also cites a pyramidal spectrum of adverse respiratory health effects listed in guidance developed by a committee of the American Thoracic Society (1985), and it claims that this supports the EPA staff position that transient and reversible effects are adverse (U.S. Environmental Protection Agency 2008e, p. 62, footnote 4). This is false. In a subsequent guideline, also cited by EPA, the American Thoracic Society (2000) said it had "hinged the distinction between adverse and nonadverse effects on medical considerations" whose "boundary is further influenced by societal considerations" (p. 666).[113] Transient and reversible effects were not on this list, and the least adverse effect in the spectrum – "interference with the normal activity of the affected person or persons" – does not admit to objective interpretation.

Like its predecessor, the 2000 ATS guidance is a mixture of medical science and policy considerations, and for that reason it is even more difficult to interpret objectively than was the 1985 list. Nonetheless, with respect to transient and reversible respiratory effects, the ATS did not define them as adverse *per se*:

---

[112] Perhaps the most obvious example is EPA's implicit characterization of the 1.5% to 2.8% group mean FEV_1 decrements reported by Adams (2006a) as adverse – even though in the Criteria Document EPA staff characterize effects ± 3% as equivalent to no effect at all.

[113] The committee apparently considered economics as a factor in determining adversity, but decided against doing so because it recognized that ATS lacked expertise in this area. See American Thoracic Society (2000, pp. 668-689).

> *Physiological impact*. The committee recommends that a small, transient loss of lung function, by itself, should not automatically be designated as adverse. In drawing the distinction between adverse and nonadverse reversible effects, this committee recommended that reversible loss of lung function in combination with the presence of symptoms should be considered adverse. This committee considered that any detectable level of permanent lung function loss attributable to air pollution exposure should be considered adverse (American Thoracic Society 2000, p. 672).

The ATS statement also specifically declined to endorse the EPA staff's definitions of adversity:

> The Environmental Protection Agency has also needed to address the interpretation of such data. The Environmental Protection Agency, in its 1989 review of ozone offered a graded classification of lung function changes in persons with asthma. Reduction of the forced expiratory volume in 1 s (FEV,) was graded as mild, moderate, or severe for reductions of less than 10%, l0-20%, and more than 20%, respectively. <u>This classification has not been validated</u> for acceptability or against other measures (emphasis added).[114]

B. <u>EPA Uses Important Scientific Terms and Language in Policy-directed Ways</u>

In our RFC, we objected on information quality grounds to EPA's use of probabilistic statements without ever defining what they mean in clear, accurate and understandable language (National Association of Manufacturers 2007, pp. 42-44). We focused particularly on EPA's use of the terms "likely" and "unlikely," which appear 144 times in volume 1 of the Criteria Document, 177 times in the Staff Paper, and 134 times in the NPRM – but in no case does EPA ever provide a definition. The terms "robust" and its adverbial variants (e.g., "fairly robust," "generally robust," "statistically robust") 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 – but EPA never defines this term, either.

The model we presented of EPA's approach to causality (Figure B in Section III.C.6) illustrated the implications of EPA's linguistic nontransparency: a large number of studies is assembled, each of which has weak or ambiguous

---

[114] Although the revised statement was published in 2000, the ATS committee did not comment on the graded scheme EPA published in the 1996 ozone Criteria Document (U.S. Environmental Protection Agency 1996a) and republished in the latest edition (U.S. Environmental Protection Agency 2006a).

evidence, but in combination they are transformed into predictions that are "likely," about which EPA staff is "confident" – another term EPA staff do not explain.

In its Response to Comments, EPA agrees in principle "where available information provides a basis for assigning quantitative values to probabilistic statements that it is generally appropriate to do so" (U.S. Environmental Protection Agency 2008e, p. 156). However, EPA does not agree that this principle imposes any duty in practice:

> EPA does not agree that it is appropriate to interpret information in quantitative terms if available information does not provide a basis to do so, which would have the effect of communicating a higher degree of precision than is warranted… (Id.)

There are several rebuttals to this reply.

1. *The definition of "likely" has nothing to do with "precision."*

A simple search of an English dictionary will show that the meaning of "likely" has nothing to do with precision and everything to do with the magnitude of probability.[115] In its response, EPA staff attempts to divert attention away from its persistent and consistent refusal to adhere to the Information Quality Act's requirement that it be transparent about the size of effects and their likelihood. When EPA staff describe a phenomenon or make a prediction that it calls "likely," it must be clearer about what "likely" means. By evading this legal responsibility, the EPA staff invites members of the public to substitute their own definitions of "likely." This abuse of qualitative probabilistic language makes EPA staff determinations neither testable nor reproducible. Moreover, EPA's Response to Comments indicates that this is entirely deliberate and intended for the purpose of obfuscation:

> [T]he word "likely" is intended to convey its common meaning, i.e., having the qualities or characteristics that make something probable. This meaning reflects a judgment, for which EPA provides a reasoned basis in these documents (U.S. Environmental Protection Agency 2008e, pp. 156-157).

---

[115] A comparison of definitions across six online dictionaries reveals none that imply that "likely" conveys any notion of precision unless it is preceded by an adverb. See http://dictionary.reference.com/browse/likely. A similar review of multiple thesauruses reveals none that say "precise" is a synonym. See http://thesaurus.reference.com/browse/likely.

The use of partial synonyms is evasive. The "reasoned basis" EPA provides in the preambles to the NPRM and final rule concern the exercise of the Administrator's policy judgment, not the description or estimation of probabilities. EPA staff have no authority to make policy decisions on behalf of the Administrator, and the Administrator does not have the discretion to base probabilities on policy judgment. The Administrator's policy judgment applies to matters of public policy and the weighting of competing social values. The Clean Air Act does not authorize the Administrator to interpose policy judgments and social values into descriptions of scientific knowledge or facts; they must be objectively determined or estimated and accurately described.

2. *If there is no scientific basis for probabilistic language, EPA should not use it.*

EPA's Response to Comments says the EPA staff cannot be more quantitatively specific about what they mean when they use probabilistic words such as "likely" because "available information does not provide a basis" for it to do so. This implies that EPA staff themselves do not know what they mean when they use probabilistic language despite the frequency with which they invoke it. If this is true, then EPA staff must cease using probabilistic language. EPA cannot expect the public to understand what Agency staff mean by "likely" if Agency staff use the term without a clear factual basis.[116]

The EPA staff's responsibility is to provide the Administrator with objective factual information about such matters as probabilities – for example, the likelihood that a well-defined health effect is occurring at ozone concentrations below the existing primary NAAQS, and if so, to whom. To the extent that this likelihood is variable (i.e., it differs across individuals and subpopulations) or uncertain (i.e., there are limits to the precision with which it can be estimated or described), EPA staff have the duty to inform the Administrator about that as well. It is then the Administrator's responsibility (and indeed, his statutory authority under the Clean Air Act) to decide whether these objectively estimated or described likelihoods are large enough that the existing primary NAAQS is no longer "requisite" to protect public health. By refusing to disclose to the Administrator and the public what they mean when they use probabilistic words such as "likely," EPA staff are violating applicable

[116] EPA follows by saying: "NAM has not identified any specific instance in which the Agency's use of terms such as 'likely' or 'unlikely' is not consistent with the common meaning of these terms." This of course is true, for EPA staff have used these terms such that they cannot be reproduced by any third party, and thus they cannot be refuted.

information quality principles and guidelines and failing to provide the Administrator with information that satisfies the utility standard.

3. *EPA staff have available to them – and have used in other contexts – clearly defined meanings for probabilistic terms.*

In its Response to Comments, EPA implies that science does not provide a superior way to described probabilities when precise figures are not available. However, EPA has a record of being much more clear about the meaning of probabilistic statements in other contexts. EPA has several times cited approvingly a scheme that defines terms including "very likely" (> 90% probability), "likely" (> 66% but > 90% probability), "unlikely: (> 10% but < 33% probability), and "very unlikely" (< 10% probability) (U.S. Environmental Protection Agency 2007a; 2007n, p. 8, footnote 3). It is inconceivable that EPA staff are unaware of these documents.

4. *Ad hoc meanings for probabilistic language are not compatible with information quality.*

In conventional English, proper words mean very specific things and ordinary words are empowered with general or universal meaning. Lewis Carroll was the first to explore the logical implications of reversing this rule. Substituting the word "likely" for "glory," and EPA staff for Humpty Dumpty, the latter's conversation with Alice would have gone like this:

> "I don't know what you mean by 'likely.'" Alice said.
>
> EPA staff smiled contemptuously. "Of course you don't – till I tell you. It means 'there's a nice knock-down argument for you!'"
>
> "But 'likely' doesn't mean 'a nice knock-down argument,'" Alice objected.
>
> "When I use a word," EPA staff said, in rather a scornful tone, "it means just what I choose it to mean -- neither more nor less."
>
> "The question is," said Alice, "whether you can make words mean so many different things."
>
> "The question is," said EPA staff, "which is to be master -- that's all."[117]

"Likely" means whatever EPA staff say it means – nothing more and nothing less. It truly is a "knock-down argument."

---

[117] Carroll (1960, p. 269, "Through the Looking Glass, Chapter VI).

EPA staff "disagree" that they have any obligation under information quality principles to be clear, accurate, and transparent. Doing so would not be an "appropriate use of Agency resources" (U.S. Environmental Protection Agency 2008e, p. 157).

We reiterate here what we said in our RFC:

> 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 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, at a minimum EPA must 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

In our RFC, we noted that EPA had presented scientific and technical information about variability and uncertainty in a confused manner (National Association of Manufacturers 2007, pp. 42-44). Reasons for EPA's confusion were hard to fathom; the distinction between variability and uncertainty is well established and understood in the risk assessment field (Morgan et al. 1990). Nonetheless, 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.

This problem infects more than just the magnitude of risk estimates. Sampling error receives almost all of the EPA staff's attention, but among sources of uncertainty, it may be the smallest. 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. The National Academy offered EPA guidance on this point 13 years ago:

> 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)

In the ozone review, EPA has not followed the Academy's recommendations.

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 human health risks. The EPA staff's characterization of various risks as "likely" is not accompanied by any indication of what probabilities are implied. Point estimates of health risk with confidence intervals capture only statistical variability for the selected models, not scientific uncertainty. 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 sufficient evidence that exposure below the current standard poses a sufficient incremental risk to warrant revising the NAAQS downward. That cannot be done without clear and accurate characterizations of probabilities.

In its Response to Comments, EPA says that it "explicitly discussed" these material analytic weaknesses in its Staff Report (U.S. Environmental Protection Agency 2008e, p. 87). In EPA's view, it is sufficient to acknowledge that "the uncertainty ranges reported in the risk assessment do not reflect all of the uncertainty in the risk estimates" without disclosing the extent to which its risk estimates are reported with unrealistically narrow confidence intervals. Information quality principles and guidelines do not require EPA to perform perfect risk assessments. They require EPA to be honest about the extent to which its risk assessments are imprecise and unintentionally biased, and to avoid utilizing procedures that purposefully impart bias.

EPA "does not agree that the preponderant effect of all of the sources of uncertainty is to create an upward bias in EPA's risk estimates" (U.S. Environmental Protection Agency 2008e, p. 87, emphasis added). We highlight the qualifier "all" because it converts our information quality complaint into a straw man; we never claimed that every aspect of EPA's risk assessment was upwardly biased.

Further, "EPA does not agree that other researchers have presented a credible, balanced, peer-reviewed integrated uncertainty analysis that shows the large majority of probability in the estimates falls far below the primary estimates that EPA reported" (U.S. Environmental Protection Agency 2008e, p. 88). The example of an "integrated uncertainty analysis" that EPA cites as lacking "credibility" and "balance" – terms that EPA nowhere defines, by the way -- is a competing analysis of mortality risks (Smith and Gibbs 2007) that is not, and does claim to be, an "integrated uncertainty analysis." EPA purports to discard Smith and Gibbs (2007) because it deals only with mortality risk and relies on assumptions different from (but at least as plausible as) the assumptions used by EPA staff. EPA's Response to Comments implies that EPA staff intend to reject any competing analyses submitted through the public comment process unless, at a minimum, they improve upon each and every aspect of the Agency's risk

assessment and secure peer review -- all within the public comment period, which in this case was 90 days.

D. EPA Does Not Disclose a Credible Analysis of Uncertainty

In our RFC (National Association of Manufacturers 2007, pp. 46-47), we noted that 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 point estimates, especially when those estimates were described as "plausible upper bounds." 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" (Ibid. p. 185).

Thirteen years later, in a risk assessment supporting one of the Agency's most far-reaching regulatory actions, EPA continues to rely on plausible upper-bound point estimates and declines to conduct or disseminate a formal uncertainty analysis.[118]

In a 2002 report to EPA specifically about the assessment of health risks from air pollution regulations, a committee of the National Academy of Sciences examined previous EPA health risk assessments and reached several conclusions, including:

- 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

[118] EPA's review plan promised very limited efforts to analyze exposure uncertainty (U.S. Environmental Protection Agency 2005e, pp. 10-11), and EPA never wavered from that limited commitment (Langstaff 2006a, 2006b, 2007).

> 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 (National Research Council 2002, p. 146).

In its ozone health risk assessment prepared five years later, nothing changed. Just as EPA staff have cherry-picked data and studies to reverse-engineer scientific support for the new standards they wanted the Administrator to adopt, they have cherry-picked advice from the NRC.[119]

In its Response to Comments, EPA defends its decision to ignore the recommendations of this NRC committee, dismissing the 2002 report as irrelevant for EPA health risk assessment:

> [T]he 2002 NRC report cited by several commenters made recommendations with respect to EPA's regulatory impact analyses which are required under E.O. 120266 [sic] and not EPA's health risk assessments (U.S. Environmental Protection Agency 2008e, p. 88).

EPA misreads the Academy report, and apparently, it has forgotten its own Charge to the committee. The intersection between benefits assessment and health risk assessment is so strong that a retired commissioned officer of the Public Health Service and emeritus professor of public health, John C. Bailar, III, was selected to chair the NRC committee – not an economist familiar with Executive Order 12866 and its Regulatory Impact Analysis requirement. In fact, of the committee's 13 members, 10 were public health scientists and only one was an economist.[120] It is entirely plausible, if not certain, that none of these public health scientists would have agreed to serve if they had known in advance that EPA would dismiss their work as relevant only to Regulatory Impact Analysis.

E. <u>EPA's Particular Use of Default Values Violates Information Quality Principles</u>

In our RFC, we noted that the use of "inference guidelines" (National Research Council 1983) and "default options" (National Research Council 1994)

---

[119] EPA (2008a) states that it adopted NRC (2002) recommendations for the selection of human health endpoints (Table 6.1), the choice of concentration-response functions associated with these endpoints (Table 6.2), reductions in school absences resulting from lowering the primary NAAQS (p. 6-18).

[120] See NRC (2002, pp. 166-170). Given the dearth of economics expertise on the committee, it is remarkable that the report contains as much economics content as it does.

has a long and checkered history (National Association of Manufacturers 2007, pp. 47-49). 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). 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, pp. 601-627, Appendix N-1) and those who said they ought not (National Research Council 1994, pp. 629-640, Appendix N-2). 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" (National Research Council 1994). Fourteen years later, EPA has not established that procedure.[121]

More importantly, the federal Information Quality Act and its implementing guidance have superseded these debates. Information of a scientific nature now disseminated by federal agencies must be objective, in both 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:

> 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

---

"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).

> discretion when she does so (*Whitman v. American Trucking 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.

We said in our RFC that the documents subject to our information quality challenge systematically incorporate default options that fail the substantive objectivity test (National Association of Manufacturers 2007, pp. 48-49). Moreover, the degree to which policy judgments that belong solely to the Administrator's discretion have been 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, a clear, complete and unbiased picture of human health risks posed by ozone exposure below the 1997 NAAQS. The documents we challenge 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 EPA staff have given him an inaccurate scientific record.

In its Response to Comments, "EPA rejects NAM's contention that it used default values and assumptions in its assessments" (U.S. Environmental Protection Agency 2008e, p. 157). EPA said we had failed to provide "specific examples of where EPA had used default values," by which we infer that the Agency takes an exceedingly narrow view of the concept of defaults and inference guidelines. At the cost of even greater redundancy, we list just a handful of defaults, each of upwardly biases EPA's estimates of human health risk or portrays these estimates as more precise than they actually are:

- Data, model selection, coefficient selection, and publication biases are negligible.
- Pulmonary tests are capable of distinguishing very small differences.
- Inter-maneuver variance in pulmonary function tests is zero.
- Ambient ozone concentrations can be assumed to be highly correlated with personal exposure even if they are not.
- Results from controlled human studies of personal exposure can be applied to ambient concentrations without adjustment for differences between personal and ambient exposure.

- Samples analyzed in critical epidemiological study are representative.
- There is no nonresponse bias in critical epidemiological studies.
- Self-reported data recorded in diaries are accurate.
- Weak epidemiological effects are causal if they are statistically significant and/or positive. Weak epidemiological effects have no information value if they are not positive.
- Asthmatic children are exposed to ozone the same as nonasthmatic children.
- All asthmatic children are equally susceptible.

EPA may reply that in each of these cases the staff was compelled by data gaps to exercise "judgment." We submit that EPA staff's exercise of judgment consistently imparted upward bias and excess precision to the Agency's risk estimates, consistent with the 2004 Staff Paper on Risk Assessment Principles and Practices (U.S. Environmental Protection Agency Office of the Science Advisor 2004b) and the Iron Law of EPA Staff Ozone Risk Assessment and Characterization. The use of "judgment" to impart purposeful bias and excess precision is incompatible with the information quality principles of substantive and presentational objectivity.

F. EPA Assumes Confidence Intervals Adequately Describe Variability and Uncertainty

In our RFC, we said that the EPA staff's approach to the various studies in its scientific database overstated confidence by assuming that variability and uncertainty were adequately described by reported confidence intervals (National Association of Manufacturers 2007, p. 49). EPA staff did this irrespective of whether the population studied was representative, irrespective of the sample response rate, irrespective of publication bias, and irrespective of model uncertainty. We cannot find any example in which EPA staff did more than "discuss" or "consider" these weaknesses before acting as if they did not exist.

In its Response to Comments, EPA addresses an unrelated question – the relative importance of statistical significance compared to "the pattern of results across various studies" (U.S. Environmental Protection Agency 2008e, p. 33). EPA further

> disagrees that most of the epidemiologic literature evaluated in the O_3 Criteria Document is based on non-random research designs. Not all epidemiologic studies evaluated in the O_3 Criteria Document use study populations that are generalizable to the entire population, but this does

> not mean that the study population was non-random (emphasis in original).

EPA's qualifying reference to the Criteria Document, which includes many studies EPA did not rely upon, is revealing. Focusing on the panel studies EPA cites in the NPRM as supporting the conclusion that the 1997 primary NAAQS is not requisite to protect public health, we see that they have a constellation of research design limitations:

- Sampling methods assumed but not demonstrated to be representative (Gent et al. 2003; Mortimer et al. 2002)
- Explicit convenience sampling (Korrick et al. 1998; Romieu et al. 1997; Romieu et al. 1996; Sarnat et al. 2005; Sarnat et al. 2000; Sarnat et al. 2001)
- Significant to severe non-response bias (Gent et al. 2003; Korrick et al. 1998; Mortimer et al. 2002; Sarnat et al. 2000)
- Reliance on unvalidated data recorded in diaries (Gent et al. 2003; Mortimer et al. 2002; Romieu et al. 1997; Romieu et al. 1996; Sarnat et al. 2000; Sarnat et al. 2001)
- The discard of inter-maneuver variability and uncertainty in FVC, FEV_1, or PEF testing (Korrick et al. 1998; Mortimer et al. 2002; Romieu et al. 1997; Romieu et al. 1996)

The amount of inflation in statistical significance is unknown, but it becomes increasingly important as effect sizes involved approach zero. Nevertheless, the EPA staff assume that the confidence intervals in the epidemiological studies accurately capture variability and uncertainty.[122]

Presentational objectivity demands at least a transparent acknowledgement of this problem and its importance, with the added advice that the results of such interpret such results with extreme caution. The NPRM shows that, in fact, EPA staff never acknowledged the problem of understated confidence intervals and interpreted their results with very little caution.

---

[122] EPA's limited uncertainty analysis consists of a Monte Carlo simulation of concentration-response functions assuming that the confidence intervals reported in the epidemiological studies accurately and completely capture uncertainty. See Langstaff (2007).

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

In our RFC, we objected to EPA's reliance on ambient ozone levels as proxies for personal exposure despite overwhelming evidence that ambient and personal exposures are uncorrelated (National Association of Manufacturers 2007, p. 50). In its Response to Comments, EPA replies that this is okay because the epidemiological studies upon which it constructed its risk assessment also rely on ambient ozone levels rather that personal exposure (U.S. Environmental Protection Agency 2008e, p. 83). Further, EPA "does not agree that there is any requirement [under applicable information quality guidelines] to provide unbiased estimates of exposure for each subpopulation group of concern before it can use concentration-response relationships in its risk assessments" (emphasis added).

This response misrepresents our complaint, for we never claimed that EPA had any such duty. Rather, we said the use of ambient ozone levels "violates the objectivity requirement of information quality because it imparts purposeful and avoidable bias to the risk estimate." Moreover, EPA lacks an unbiased estimate of exposure for any subpopulation of concern, or for the population as a whole. While EPA (sort of) denies that the use of ambient ozone data results in biased risk estimates,[123] the Agency's Response to Comments never replies to any of the public commenters who provided evidence otherwise – or, most ironically, CASAC:

> **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

[123] "The fact that ambient concentrations may overstate actual personal exposure does not imply that the risk estimates are biased." See EPA (2008e, p. 83).

> 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.
>
> 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 (Henderson 2006b, pp. 3-4).

In the second draft Staff Paper, EPA staff responded to CASAC by digging in their bureaucratic heels[124] and erecting a huge impediment to objective

---

[124] "O_3 concentrations measured at central ambient 10 monitoring sites may explain, at least partially, the variance in individual exposures; however, this relationship is influenced by other factors such as air exchange rates in housing and time spent outdoors which may vary from city to city. Other studies conducted in various cities observed that the daily averaged personal O_3 exposures from the population were well correlated with ambient O_3 concentrations, although substantial variability existed among the personal measurements. Thus, there is 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 the most relevance for time-series studies. This is especially true for respiratory hospital admission studies, for which much of the response is attributable to O_3 effects on people with asthma. Ambient monitors are more likely to correlate reasonably well with the personal exposures of children, who spend more time outdoors in the warm season and who are also more likely to have asthma than adults. Conversely, there is some concern about the extent to which ambient concentrations are representative of personal O_3 exposures of another particularly susceptible group of individuals, the debilitated elderly, and what impact that may have on mortality and hospitalization time-series studies. The correlation between ambient concentrations and personal exposure measurements has not been examined in this population. 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 with ambient O_3 concentrations (U.S. Environmental Protection Agency 2006f, p. 3-39).

exposure assessment – a default assumption that, absent the routine collection of personal ozone exposure data, they were committed to using ambient ozone for reasons of expedience:

> [P]opulation health risk estimates derived using ambient O_3 levels from currently available observational studies, with appropriate caveats about personal exposure considerations, remain useful (U.S. Environmental Protection Agency 2006f, p. 3-40).[125]

The practical consequence of EPA staff using ambient concentrations in lieu of personal exposures is to significantly bias the scientific record provided to the Administrator. Under the NAAQS program, EPA sets standards for ambient concentrations, not personal exposure. EPA acknowledges that ambient concentrations exceed personal exposures by 2- to 4-fold,[126] interprets this as implying that ozone is more potent,[127] then discards this algebraic relationship. Adams (2002, 2006a) estimated group mean decrements in FEV_1 of approximately 1.5% compared to 0.04 ppm (2.8% compared to filtered air) when subjects were exposed to personal exposures of 0.06 ppm. EPA staff thus should be multiplying by 2- to 4-fold to obtain the ambient concentration equivalent. Instead, they treat personal exposures in controlled experiments as if they were the same as ambient concentrations in epidemiological studies. The results obtained by Adams at 0.06 ppm in personal exposure are roughly equivalent to 0.12 to 0.24 ppm in ambient concentration equivalents, using EPA's own conversion metric.

---

[125] EPA never defines the meaning of "useful," nor does it explain the significance of these "appropriate caveats."

[126] "Using ambient concentrations to determine exposure generally overestimates true personal O_3 exposures (by approximately 2- to 4-fold in the various studies described in the Criteria Document, section 3.9)…" EPA (2008b, p. 16458).

[127] "[A]ssuming the relationship is causal, [this] would result in biased descriptions of underlying concentration-response relationships (i.e., in attenuated effect estimates). From this perspective, the implication is that the effects being estimated in relationship to ambient levels occur at fairly low personal exposures and the potency of O_3 is greater than these effect estimates indicate" EPA (2008b, p. 16458).

H. EPA Assumes that Associations Observed in Short-Term Time Series Studies Are Significant and Meaningful, but the Absence of Associations in Long-Term Cohort Studies Is neither Significant nor Meaningful nor Logically Inconsistent

In our RFC, we asked EPA to reconcile the Agency staff's view that short-term time-series studies which show positive associations with mortality are supportive evidence of risk, but long-term cohort studies which do not show such associations are not evidence of the absence of risk (National Association of Manufacturers 2007, p. 50). We inferred that EPA was concluding "ozone causes premature mortality in the short-term that cannot be observed over the long-term."

In its Response to Comments, "EPA rejects NAM's contention that it has reached inappropriate conclusions about associations between O_3 exposure and premature mortality" (U.S. Environmental Protection Agency 2008e, p. 53, emphasis added). EPA resolves inconsistency by implying that long-term epidemiological studies also would have supported the staff's inference that ozone causes mortality, if only they too had been statistically significant and/or positive. Precisely because these studies were "not consistent," they were effectively discarded.

EPA repackages our complaint about scientific inconsistency into a sterile debate about "appropriateness," a complaint we never raised because "appropriateness" has no scientific meaning. Wherever EPA's scientific statements are illogical, inconsistent, non-reproducible, or otherwise controlled by undisclosed and illegitimately exercised staff views about air pollution policy, Agency staff abandon any pretense to be evaluating science and instead assert the right to exercise unfettered judgment under the cloak of science.

I. EPA Assumes Causality

In our RFC, we objected to EPA's method of handling causality, which may be the most important scientific issue in the entire ozone review (National Association of Manufacturers 2007, pp. 50-51). Nowhere in any of EPA's supporting documents does the staff make its method of determining causality transparent to the Administrator or the public, nor is its method reproducible by third parties. It is therefore impossible to test or refute it utilizing scientific methods and procedures. The EPA staff have discarded causality as a scientific concept and replaced it with opinion.

In its Response to Comments, EPA denies that it has any obligation under information quality guidelines to describe causality in a probabilistic (i.e., scientific) manner (U.S. Environmental Protection Agency 2008e, pp. 84-85).

Similar to other crucial scientific concepts that EPA staff do not want to be transparent about, "causality" is whatever the EPA staff say it is; nothing more, and nothing less.[128]

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

We have pointed out several times elsewhere that the EPA staff approach is best explained as an Iron Law of EPA Staff Ozone Risk Assessment and Characterization. In our RFC, we noted that for a presentationally objective characterization of human health risks actually or purportedly due to ozone exposure below the NAAQS it was necessary to place ozone-associated health risks in context (National Association of Manufacturers 2007, p. 51). In its Response to Comments, EPA says it

> believes it has provided sufficient context in its discussion of respiratory effects in the Criteria Document and Staff Paper and that there is no specific requirement to make the type of comparison suggested by [NAM] (U.S. Environmental Protection Agency 2008e, p. 99).

It is hard to understand the basis for EPA's conclusion that it faithfully adhered to the information quality standard of presentational objectivity given that none of EPA's supporting documents ever mention the subject. EPA needs thousands of pages to explain what it knows about the health risks from ozone, but zero pages to explain why these thousands of pages are presentationally objective.

K. Double-counting

In our RFC, we said we sympathized 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 NO_x. (National Association of Manufacturers 2007, pp. 51-52) Still, as we said in our RFC, double counting is simply not acceptable under information quality principles. EPA's risk assessment relies on many studies that estimate effects of ozone along with other air pollutants. We said EPA had an obligation to allocate health risk across these competing sources to ensure that it was not double-counting.

In its Response to Comments, EPA admits that double-counting is possible but says that if it occurred its magnitude was small (U.S. Environmental

[128] See the discussion in Section IV.B beginning on page 111 about EPA's serial abuse of probabilistic language.

Protection Agency 2008e, pp. 87-88).[129] The basis for EPA's confidence is a single meta-analysis (Levy et al. 2005). The authors considered only single-pollutant models in their main analysis (p. 459), and performed only a sensitivity analysis with respect to the confounding effect of $PM_{2.5}$ (p. 463).[130] It is the one-paragraph description of this sensitivity analysis that EPA staff rely on as the basis for dismissing our concern about double-counting of risks.

L. EPA's Alternative Risk Estimates

In our RFC, we noted that substantive objectivity requires that information be presented in an "accurate, reliable, and unbiased" manner, and we observed that EPA did not adhere to this requirement in the reporting of alternative risk estimates (National Association of Manufacturers 2007, pp. 52-53). EPA characterizes some of its risk estimates as "primary" and others as "secondary." This language implies that one set of estimates have a stronger scientific foundation and are more likely to be correct than the other set of estimates. However, nowhere does the Agency use science or statistical method to show that this distinction is grounded in either science or probability.

We said that EPA's "primary" risk estimates were those that most tended to support a policy preference for a more stringent NAAQS, and EPA's "secondary" risk estimates are those that provided less support. This distinction is purely arbitrary. 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

---

[129] EPA also admits that it has a preference for single-pollutant models because, in multi-pollutant models, the coefficients for ozone lose stability. Coefficient instability across model specifications is a common indicator of model specification error (Kennedy 1985). EPA ignores this and characterizes its results as "robust." See the discussion about "robustness" in Section IV.B beginning on page 111.

[130] EPA's inferences are much stronger than these made by the authors: "The less robust influence of NO_2, along with the weak effect of $PM_{2.5}$, is hard[] to interpret. Given the evidence demonstrating a relationship between ambient $PM_{2.5}$ and mortality, a stronger association for he $PM_{2.5}$-ozone association may have been anticipated… Our findings could be related to difficulties in identifying causal factors in a multivariate context, limitations in our ambient pollution data, or might indicate that the use of air pollution regression coefficients in hierarchical linear models is not the optimal approach for evaluating confounding" (Levy et al. 2005, p. 465).

prohibit it from mischaracterizing these policy preferences as scientific, or informed by science, when they are not.

In its Response to Comments, EPA says "NAM's contention that EPA's risk estimates are characterized as 'primary' or 'secondary' in the Staff Paper or proposal notice is incorrect" (U.S. Environmental Protection Agency 2008e, p. 100, emphasis added). In our RFC, however, we said nothing about the Staff Paper or the NPRM. The section in which this complaint appeared concerned EPA's risk assessment. EPA's risk assessment makes a very clear distinction between the staff's "primary" risk estimates…

> [T]he exposure-response functions used in the primary analyses are based on the assumption that the relationship between exposure and response has a logistic form with 90 percent probability and a linear (hockeystick) form with 10 percent probability.

… and its "secondary" risk estimates:

> In this sensitivity analysis, we considered the impact of two alternative exposure-response functions, based on an 80 percent logistic/20 percent linear split and a 50 percent logistic/50 percent linear split, in five locations – Atlanta, Chicago, Houston, Los Angeles, and New York.[131]

Bias in EPA's risk assessment is rather obvious. Its most controversial aspect is the assumption that the extraordinarily weak associations observed in selected epidemiological studies are causal.

The public must look to EPA's Regulatory Impact Analysis – a document that was not completed until after the Administrator made his decision -- to uncover the implications of assuming causality, especially for mortality risk. In the RIA, EPA acknowledges that the value of mortality risk reductions from NAAQS standards has historically comprised 85% to 95% of total estimated health benefits (U.S. Environmental Protection Agency 2008a, p. 6-6). It would be negligent for the Administrator to have ignored that ratio, and of course, the EPA staff risk assessment reasonably led him to believe that these benefits were real.[132]

---

[131] EPA (2007c, pp. 3-76 to 73-77).

[132] A common myth surrounding NAAQS rulemakings is that the Administrator cannot use the RIA to inform decision-making. The Clean Air Act prohibits the Administrator from taking account of the cost of achieving the NAAQS, but it does not compel him to also ignore benefits. Indeed, the whole point of regulating air pollution is to generate benefits. According to the RIA, the value in 2020 of assumed mortality

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

In this section of our RFC, we discussed a wide range of information quality errors in EPA's management of peer review by the Clean Air Scientific Advisory Committee (CASAC) (National Association of Manufacturers 2007, pp. 53-58). We noted that CASAC review is complicated by the inherently conflicted mission Congress established for it – to perform both a scientific review (which requires scrupulous attention to facts and data) and policy advice (which is fettered by no such constraints).[133] We said that this conflicted mission requires EPA to be extraordinarily careful in how it listens to CASAC to ensure that it clearly distinguishes between CASAC's scientific insight and its policy prescriptions. We noted that, as an independent body outside of the Agency's control, CASAC is exempt from federal information quality guidelines, but that EPA is not exempt when it disseminates or uses information provided by

---

reductions from lowering the primary NAAQS to 0.075 ppm is 23% to 44% of total benefits. Between 50% and 99% of these benefits come from serendipitous reductions in $PM_{2.5}$. See EPA (2008a, p. ES-3).

[133] 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.
>
> (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.

CASAC. We were pleased to read in EPA's Response to Comments that the Agency agrees wholeheartedly with this synopsis and demarcation of responsibilities (U.S. Environmental Protection Agency 2008e, p. 150).

Where we disagreed with EPA – and continue to disagree – concerns EPA's implementation of this common understanding. We said EPA cannot simply cite CASAC as a scientific authority without regard for whether the contents of statements are scientific and whether scientific statements adhere 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 information quality standards. It cannot simply attribute the information to CASAC and assume that it is scientifically objective, or assume that it speaks to science and not policy.[134]

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.[135] If it fails to make this distinction, EPA voids the "opinion exemption" in the definition and subjects policy advice to the same level of scrutiny to which scientific information must adhere. 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

---

[134] 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).

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

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), emphasis added). 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. Without violating its statutory assignment, EPA cannot disseminate or use for decision-making a risk assessment that is inaccurate, incomplete, or fails to represent the latest scientific knowledge.[136] The problem facing CASAC is clear: how does the panel perform this scientific responsibility without allowing the infiltration of its members' policy views?

EPA could have made CASAC's job much easier if it had structured its charge around the information quality principles the Agency promulgated in 2002 (U.S. Environmental Protection Agency 2002) and elaborated upon in 2003 (U.S. Environmental Protection Agency 2003). Unfortunately, EPA instead decided to exclude from the charge to CASAC all information quality content. Nowhere in the charge did EPA discuss the crucial information quality concepts of *utility* and *objectivity*. Nowhere did it reference the Agency's own foundational information quality documents. CASAC can be forgiven for knowing nothing about information quality, because EPA apparently worked hard to keep its members in the dark.

In its Response to Comments, EPA acknowledges our complaint about the absence of information quality content from the CASAC charge, then proceeds to obfuscate the matter with statements that are irrelevant or literally fantastic (U.S. Environmental Protection Agency 2008e, p. 150). Irrelevancies include CASAC's separate status,[137] which has nothing to do with EPA's charge to CASAC; and the fact that CASAC's policy recommendations are exempt from information

[136] § 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").

[137] "CASAC is a separate entity from EPA and, as such, assesses scientific and other documents produced by EPA independently of Agency oversight."

quality,[138] something no one disputes. EPA claims to have incorporated information quality throughout its Action Development Process (ADP),[139] a claim that cannot be tested because the primary guidance document is hidden on the Agency's Intranet where it cannot be publicly examined.[140] Looking elsewhere for evidence, we note that in 2006 EPA publicly disseminated an ADP guidance document for children's health, and this document is silent about information quality (U.S. Environmental Protection Agency 2006d).

By far the most fantastic element in EPA's reply is its claim that the Agency has no responsibility to actually perform pre-dissemination review just because it had promised to do so:

> EPA's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated [sic] does not require the Agency to discuss, separately, whether the pre-dissemination review actually occurred (U.S. Environmental Protection Agency 2008e, p. 150).[141]

B. CASAC's Policy Advice Charge

CASAC's review of the EPA Staff Paper is necessarily different, for the Staff Paper contains 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

---

[138] "EPA cannot subject CASAC recommendations to information quality standards."

[139] "The [Information Quality] Guidelines, rather, provide a process for developing quality actions, of which the pre-dissemination review procedures are a part. This process is also a part of EPA's Action Development Process (ADP). EPA's ADP is a mechanism that assists the Agency in achieving the objectivity and transparency of information used in developing regulations."

[140] Environmental Protection Agency, "EPA's Action Development Process: Guidance for EPA Staff on Developing Quality Actions," June 2004; http://intranet.epa.gov.adplibrary/index.htm.

[141] An unknown office within EPA issued pre-dissemination review guidelines in September 2006 (U.S. Environmental Protection Agency Office of Environmental Information 2006) and made them publicly available (http://www.epa.gov/region2/science/qmp/pdfs/pdr-guidelines.pdf). The text of these guidelines makes abundantly clear, however, that they were issued because program offices such as the Office of Air and Radiation and staff offices like the Office of Research and Development had failed to implement the pre-dissemination review requirements in the Agency's Information Quality Guidelines.

policy advocacy roles.[142] CASAC does not seem to have adhered to that principle; it is difficult to discern where it is commenting on science and opining about policy. To take just one obvious example mentioned in our RFC (p. 56), the list of bullets in its letter review of the Staff Paper contains both scientific comments and policy advice, often within the same bullet (Clean Air Scientific Advisory Committee 2006a, pp. 2-3).[143]

We noted in our RFC that CASAC's members are of course expected to provide the Administrator with their policy advice concerning how he ought to exercise his statutory discretion in revising or retaining the NAAQS. Because their principal charge is scientific, however, 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, not a scientific one. When a person "representing" State air pollution control agencies gives advice, it is presumed that this advice will favor intensifying the stringency of federal air pollution standards if that is what the

---

[142] 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 2007g).

[143] In its Response to Comments, EPA says:

> In this rulemaking, EPA is confident that it has been able to clearly differentiate CASAC's science advice from the policy advice on the appropriateness of new or revised NAAQS. NAM has not identified examples where it believes EPA has failed to so differentiate, nor examples where CASAC has improperly mixed science and policy in providing its advice." (U.S. Environmental Protection Agency 2008e, p. 149).

It is indisputable that CASAC mixed science and policy, so EPA must be saying that it was not "improper" for CASAC to do so. If that is so, then EPA also convicts itself of failing to differentiate science from policy in its use of input from CASAC.

governing authorities in that State prefer. It would be newsworthy only if this person recommended against more stringent federal standards.[144]

CASAC members also are asked to "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 have scientific training and have distinguished themselves in one or more scientific fields, there is nothing scientific about giving policy advice.

The provision of policy advice by scientists is further confounded by two other phenomena, one that applies to scientists in general and one that applies specifically to this panel. The general fact is that all scientists are susceptible to the temptation to believe that their status as scientists endows them with special insights about public policy. Some scientists don't care about policy, but they are the least likely to be recruited to serve on panels such as CASAC or be interested in doing so. CASAC members work long hours for token financial compensation;[145] the ability to influence public policy is their primary reward.

The phenomenon that is specific to this panel is that many of them are authors of research papers in the scientific database on ozone. It is entirely natural for them to think that their own research is most relevant to the questions at hand.[146] This raises a serious question: Are CASAC members being asked to indirectly review their own work? This practice is permitted under the National Academy of Sciences' conflict of interest rules, but with an important limitation that, if it had been rigorously applied to CASAC, probably would have required many of them to be recused:

---

[144] EPA selected as a State representative an official from Vermont. Among other things, Vermont has been a party to litigation against EPA advocating more stringent air pollution standards. The Administrator would have received completely different policy advice if he had appointed an official from a State whose elected leadership opposed more stringent air pollution standards. The act of selecting the statutorily-required State representative determines the content of "State" stakeholder input.

[145] See footnote 86 for an interesting exception in which a CASAC ozone panel member reveals having devoted about 12 hours per year to the review task.

[146] Some CASAC members are especially fond of their own work. CASAC's letter review of EPA's final draft Staff Paper cites for special emphasis six peer reviewed papers authored or co-authored by CASAC members Drs. Morton Lippman and/or Frank Speizer, all published between 1988 and 1993 (i.e., prior to the 1997 NAAQS review).

> [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 (The National Academies 2003, p. 5 , document not paginated).

We noted previously that at least one crucial study for EPA's health risk assessment was co-authored by a CASAC ozone panel member.[147]

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

The NPRM contains numerous subsections in which the input it received from CASAC is summarized. In our RFC, we noted that in some places this input is clearly described as scientific information or policy advice. In most instances, however, the line between science and policy is difficult to discern. We appreciate EPA's challenge because in many cases – particularly in its review of the Staff Paper -- CASAC itself did not make these distinctions clear. Nevertheless, adherence to information quality guidelines is EPA's responsibility and not that of CASAC. EPA's decision to shield CASAC from information quality principles and standards in its charge does not alleviate the Agency's responsibility.

D. EPA's Lack of Pre-Dissemination Review

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 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, p. 8459, emphasis added).

OMB's use of the imperative "shall" signifies that these requirements are not optional or merely suggestive, but rather they are mandatory. This is entirely consistent with Information Quality Act, which gave OMB similarly imperative language to implement in its government-wide guidelines, to which EPA and its guidelines are subordinate (Information Quality Act 2000).

---

[147] See footnote 108

EPA's own guidelines commit the Agency to obey the directives of statute and OMB's government-wide guidelines for the establishment and implementation of 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 (U.S. Environmental Protection Agency 2002, p. 29, emphasis added).

EPA's now-acknowledged failure to actually perform pre-dissemination review, combined with its steadfast claim it doesn't actually have to do so despite this commitment, implies that the Agency thinks both pre-dissemination review and compliance with the Information Quality Act is not "appropriate." If EPA really believes that it should be exempt from the principles of information quality because those principles are not "appropriate" for the scientific information supporting the ozone NAAQS, the Agency should say so transparently and explain its reasoning.[148]

VI. Information Quality Errors in the Rollback Assumption

In our RFC, we objected on information quality grounds to EPA's rollback assumption (National Association of Manufacturers 2007, pp. 58-60). EPA's approach violated information quality standards by failing to approximate how the States actually would respond to a lower NAAQS. This is relevant for estimating the incidence of various health effects avoided. In particular, EPA's model assumes that compliance with a lower NAAQS will result in reductions not just at the peaks, where a determination of attainment is made, but also throughout the entire distribution. We suggested that EPA validate its model by testing it against actual data from State implementation of the 1997 NAAQS. We also expressed concern that reductions at the low end of the distribution were

---

[148] The dearth of pre-dissemination review is particularly notable for the one instance in which information quality principle of *objectivity* appears in the NPRM: 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" (U.S. Environmental Protection Agency 2007h, p. 37881). In the Staff Paper and RIA, EPA's argument for failing to account for UV-B shielding is that "this beneficial effect of [UV-B] radiation has not previously been studied in sufficient detail" (U.S. Environmental Protection Agency 2008a, p. 6-21). This issue was first raised before the 1997 ozone NAAQS was issued (Lutter and Wolz 1997) and it became a central element of litigation. Since then, EPA has steadfastly refused to account for UV-B because it is incompatible with the Envelope Theory.

particularly problematic given both the uncertainty about true background and EPA's controversially low values for Policy Relevant Background (PRB). EPA may be crediting its new ozone NAAQS with reducing background ozone concentrations.

In its Response to Comments, EPA "concluded" that its model "generally best represented the pattern of reductions across the O_3 air quality distribution observed over an 8-year period in areas implementing control programs designed to attain the O_3 NAAQS" (U.S. Environmental Protection Agency 2008e, p. 90). Furthermore, EPA says "only reducing peak 8-hour daily maximum values that are at or near the standard level is unrealistic in that most O_3-related air pollution control measures are continuous in nature and have an impact on the entire distribution of 8-hour O_3 concentrations"

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

In the Staff Paper, EPA defines Policy Relevant Background (PRB) in a way that makes it ambiguous as to whether it is a scientific estimation or a policy-driven default assumption:

> For purposes of this document, background or policy relevant background (PRB) O_3 is defined as the distribution of O_3 concentrations that would be observed in the U.S. in the absence of anthropogenic (man-made) emissions of precursor emissions (e.g., VOC, NO_x, and CO) in the U.S., Canada, and Mexico (U.S. Environmental Protection Agency 2007g, p. 2-47).

Despite the word "policy," in the title of the concept, *PRB* is a strictly scientific concept. That is, PRB should be defined as the level of ozone that would be present if all controllable anthropogenic U.S. sources did not exist.[149] EPA's PRB is unambiguously biased both by definition and in implementation.

A. EPA's *Definition* of Policy Relevant Background is Biased

As we noted in our RFC, EPA's PRB is biased because it assumes that ozone precursors from anthropogenic sources in Canada and Mexico are subject to control by U.S. air pollution policy and regulation (National Association of Manufacturers 2007, p. 60). This assumption is false. By treating these emissions as if they were controllable by State Implementation Plans, EPA understates the level of ozone that would exist if all U.S. anthropogenic sources were "turned

[149] The prefatory clause should be discarded, for this definition applies not just in the Staff Paper but throughout the package of documents.

off." This yields upwardly biased estimates of baseline risk and risk reduction from lowering the NAAQS.

In its Response to Comments and the preamble to the Final Rule, EPA asserts, in virtually identical language, that the Agency has the capacity to "influence" emissions from Canada and Mexico; that this capacity to "influence" arises from its ability to negotiate international agreements with Canada and Mexico; and that Canadian and Mexican emissions must be assumed to be controllable by EPA because EPA has defined PRB this way "over more than two decades" (U.S. Environmental Protection Agency 2008b, p. 16468; 2008e, p. 93). The first two of these arguments demonstrates that what should have been a scientifically defined quantity is purposefully biased by EPA staff, in violation of information quality principles. The EPA staff definition is not scientific but policy-driven; it deflates the estimated level of background ozone, inflates the amount of ozone reduction that in principle could be achieved by lowering the ozone NAAQS, and therefore inflates estimated reductions in risk.

The third argument is an appeal to tradition: EPA has erred for more than 20 years, and errors committed over that long a period ought to be exempt from information quality principles. Of course, nothing in the Information Quality Act or any of the relevant implementation guidance documents exempts information that is inaccurate or biased just because it has been used before, or for a long time. The only test for applicability is met if EPA is currently disseminating the information. That test is clearly satisfied. Moreover, though our RFC we have invoked the statutorily prescribed process for correcting information quality error. It is illegal for EPA to decline to correct error because it has a history of committing similar errors and correcting the error now is inconvenient.[150]

A closer look at the history of the 1997 ozone NAAQS review shows that EPA also was not transparent about the exclusion of Canadian and Mexican emissions from the definition of PRB. A search of the 1996 Criteria Document, the 1997 Staff Paper, the 1996 NPRM and the 1997 final rule preambles reveals no discussion whatsoever on this point. In that review, EPA stated that background was assumed to be 0.04 ppm (U.S. Environmental Protection Agency 1996b, p. 65726), and there does not seem to have been much controversy over the point. If

[150] The NPRM did not disclose to the public this important aspect of EPA's definition of Policy Relevant Background. That alone was a violation of the presentational objectivity standard. We have noticed that EPA has rectified this error in the final rule by explaining that precursor emissions from Canada and Mexico are not included in PRB because EPA assumes that its regulatory actions can and will target them (U.S. Environmental Protection Agency 2008b, p. 16433, footnote 13).

in fact EPA has for more than 20 years counted Canadian and Mexican emissions as controllable by Agency action, then these prior actions also were biased and violated information quality guidelines.

CASAC appears to have accepted this policy-driven assumption at the outset because EPA staff built it into CASAC's charge, thereby removing it from the scope of the panel's scientific – and policy -- review:

> 1. Policy Relevant Background (PRB) Ozone. PRB ozone concentrations will ultimately be taken into account by OAQPS in analyses to be included in the Ozone Staff Paper that attempt to estimate risks to human health and environmental effects associated with exposures to ozone concentrations attributable to anthropogenic sources of precursors emitted in the United States, Canada and Mexico (i.e., to ozone levels above PRB concentrations). The estimation of PRB ozone concentrations precludes the use of observational data alone because of substantial production and transport from anthropogenic sources in the United States and bordering countries. Contributions to PRB ozone arise from intrusions of stratospheric ozone, biogenic and other natural sources of ozone precursors, and anthropogenic sources outside of the U.S., Canada and Mexico. The modeling approach that has been adopted for estimation of PRB concentrations is based on peer reviewed journal articles describing the GEOS-CHEM model, its evaluation and application to the calculation of PRB ozone values. See Henderson (2005a, pp. B-1 to B-2, emphasis added).[151]

Still, CASAC ultimately distanced itself from the EPA staff's policy-driven approach:

> [W]ith respect to policy-relevant background (PRB), the Ozone Panel wishes to point out that the Final Ozone Staff Paper does not provide a sufficient base of evidence from the peer-reviewed literature to suggest that the current approach to determining a PRB is the best method to make this estimation. One reason is that part of the PRB is not controllable by EPA. It would require international cooperation beyond the bounds of North America. A better scientific understanding of the PRB and its

[151] Note also that the charge also precludes CASAC review of the merits of observational data. EPA staff faced some resistance on this point; see, e.g., the comments by CASAC panel member Barbara Zielinska (Henderson 2005a, p. C-133). For CASAC as a group to have objected, however, they would have had to decide to overrule their charge – an unlikely and highly controversial act.

> relationship to intercontinental transport of air pollutants could serve as the basis for a more concerted effort to control its growth and preserve the gains in air quality achieved by control efforts within the U.S.[152]

The NPRM acknowledges that CASAC was disturbed by other technical aspects of EPA's model for estimating PRB and, in a footnote, committed to reopen the matter:

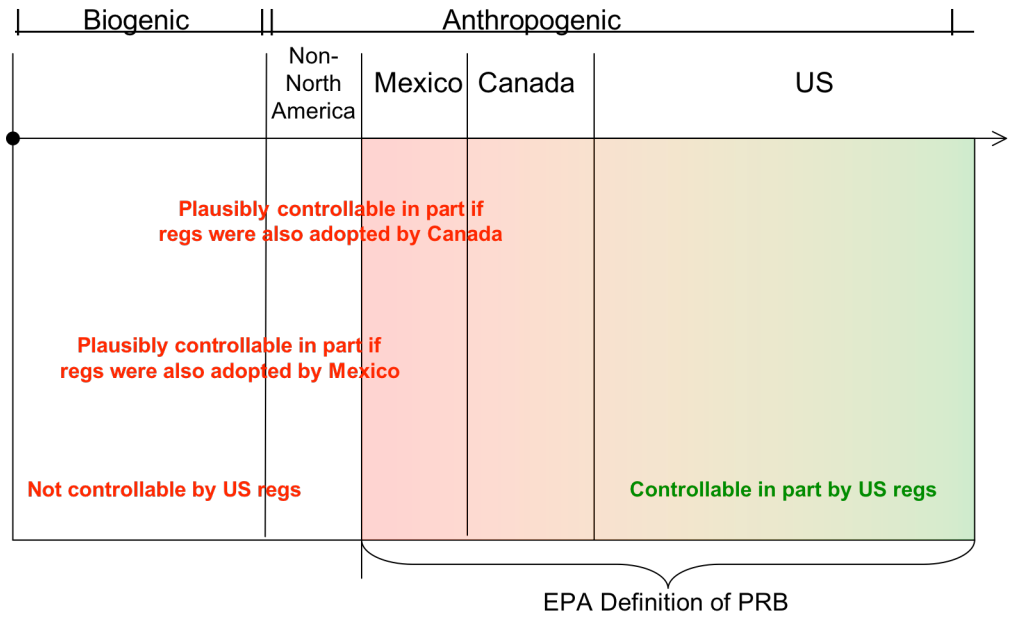
> Recognizing the importance of this issue, EPA intends to conduct additional sensitivity analyses related to policy-relevant background and its implications for the risk assessment (U.S. Environmental Protection Agency 2007h, p. 37857, footnote 40).

---

[152] Henderson (2007a, pp. 2-3, emphasis added, internal citations omitted). We have omitted the remainder of the paragraph (reprinted verbatim below) because it is not germane to the issue of whether Canadian and Mexican emissions of ozone precursors belong in background:

> In any case, there is no apparent need to define PRP [sic] in the context of establishing a health-based (primary) ozone NAAQS. The effects of inhaled ozone on decreases in respiratory function have been seen in healthy children exposed to ozone within ambient air mixtures in summer camps. Furthermore, the concentration-response functions above 40 ppb are either linear, or indistinguishable from linear. Thus, PRB is irrelevant to the discussion of where along the concentration-response function a NAAQS with an 8-hour averaging time that provides enhanced public health protection should be.

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



During the interregnum between proposal and final, however, EPA seems to have abandoned its intent to conduct additional sensitivity analysis. To allay legitimate concerns that what EPA "intends" to do is bury this issue until the next ozone NAAQS review, EPA must, at a minimum, publicly disclose the contribution of Canadian and Mexican precursor emissions, show what effects including them in PRB has with respect to risks and benefits, and inform the public concerning what impacts this will have on affected States' efforts to achieve attainment.

In our RFC, we illustrated the combined effect of these two critical assumptions (see Figure C reprinted below). Ozone emissions were 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 were 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.

B. EPA's *Estimates* of the Magnitude of Policy Relevant Background Are Biased

In our RFC, we objected to EPA's estimates of PRB because they are biased by design (see Section A above), and because they are based on modeling that appears not to have been validated (National Association of Manufacturers 2007, p. 61). Other public commenters have raised concerns about the lack of spatial and dynamic resolution in EPA staff's modeling approach (Brauer et al. 2007; Smith and Gibbs 2007). Limited resolution is not *per se* an information quality defect. Nevertheless, it appears to be undisputed that the public health significance of any choice of primary ozone standard depends crucially on how PRB is modeled or estimated, and that makes the estimation or modeling of PRB an information quality issue of paramount concern and relevance.

Smith and Gibbs performed a sensitivity analysis to determine how health risk estimates differ depending on the choice of PRB. They report that EPA's health risk estimates "would typically be 90% to 100% lower" if 0.04 ppm had been used as PRB instead (Smith and Gibbs 2007, p. 16). They also attempted to validate EPA's modeling results by comparing them to data from Trinidad Head CA, and found that if these observational data had been used as background, health risk estimates would be 65% lower in Sacramento and 72% lower in Los Angeles. They did not find any city in which EPA's new approach to PRB resulted in a lower risk estimate.

In its Response to Comments, EPA dismisses these information quality concerns on the ground that they "were considered by EPA's scientific staff and the CASAC Panel during the course of reviewing the Criteria Document" (U.S. Environmental Protection Agency 2008e, p, 94, emphasis added) – a boilerplate reply. Having "considered" an information quality error and done nothing about it is not compatible with EPA's obligations under the Agency's Information Quality Guidelines, nor can EPA hide behind a peer review in which information quality principles, policies and procedures played no role.

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."

VIII. Conclusion

We identified a large number of information quality errors in our RFC. In its Response to Comments, EPA dismisses virtually all of them, often without bothering to provide either as logical or evidentiary basis. In many case, EPA's Response to Comments mischaracterizes our complaint and responds only to its own mischaracterization. Sometimes, EPA describes the information quality complaint correctly but "answers" it by discussing irrelevant or unrelated matters. Finally, the general tone of EPA's Response to Comments is one of opinion – that is, EPA "disagrees" with or "rejects" our information quality complaints as if they are matters of opinion rather than knowledge or fact. In this broad sense, EPA's Response to Comments fails to fulfill the Agency's duty under information quality guidelines to fairly and objectively address challenges to its representations of knowledge or fact. EPA apparently seeks to evade the discipline of information quality principles by erroneously characterizing all disputes as matters of opinion.

It has been said that the absence of evidence is not the same as evidence of absence. That adage does not hold sway in this case, however. The absence of evidence of information quality principles in every EPA staff work product; the absence of any pre-dissemination review; the absence of information quality from the EPA staff's charge to CASAC, and its corresponding absence from CASAC's review; and the absence of information quality principles and analysis in the preambles to both the NPRM and final rule, make clear beyond any reasonable doubt that EPA staff did not comply with the Agency's information quality principles and guidelines at any time since the ozone review began in 2005.

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 legally 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 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, and failure to adhere to information quality principles prevents the EPA staff from producing the accurate scientific record that the Clean Air Act requires.

IX. References

Adams WC. 2000. Ozone Dose-Response Effects of Varied Equivalent Minute Ventilation Rates. *Journal of exposure analysis and environmental epidemiology* 10(3): 217-226.

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. 2003. 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. 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 Profiles. *Inhalation Toxicology* 18: 413-422.

Adams WC. 2007. *Public Comment to CASAC Ozone Review Panel Teleconference.*

American Petroleum Institute. 2006. *Comments on National Ambient Air Quality Standards (or Ozone, OAQPS Staff Paper - Second Draft (Docket No. EPA-HQ-OAR-2005-0172-0057).*

American Thoracic Society. 1985. Guidelines as to What Constitutes an Adverse Respiratory Health Effect, with Special Reference to Epidemiologic Studies of Air Pollution. *American Review of Respiratory Disease* 131: 666-669.

American Thoracic Society. 2000. What Constitutes an Adverse Health Effect of Air Pollution? *American Journal of Critical Care Medicine* 161: 665-673.

Bell ML, Dominici F, Samet JM. 2005. A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. *Epidemiology* 16(4).

Bell ML, McDemott A, Zeger SL, Samet JM, Dominici F. 2004. Ozone and short-term mortality in 95 US urban communities, 1987-2000. *Journal of the American Medical Association (JAMA)* 292(19): 2372-2378.

Bell ML, McDermott A, Zeger SL, Samet JM, Dominici F. 2004. Ozone and Short-term Mortality in 95 US Urban Communities, 1987-2000. *JAMA* 292(19): 2372-2378.

Bell ML, Peng RD, Dominici F. 2006. The Exposure-Response Curve for Ozone and Risk of Mortality and the Adequacy of Current Ozone Regulations. *Environmental Health Perspectives* 114(4): 532-536.

Brauer M, Frampton MW, Koutrakis P, McClellan RO, McDonnell WF, Moolgavkar S, *et al.* 2007. *Comments on National Ambient Air Quality Standards for Ozone: Proposed Rule, Federal Register, Vol. 72, No. 132/Wednesday, July 11, 2007, pp 37818-37919: Critical Considerations in Evaluating Scientific Evidence of Health Effects of Ambient Ozone; Docket No. EPA-HQ-OAR-2005-0172-4727.1*

Brown JS. 2006. *Biological Plausibility of Human Effects Due to Air Pollutants: Evidence from Human Clinical Studies.* "US EPA Workshop on Interpretation of Epidemiologic Studies of Multipollutant Exposures and Health Effects; December 13-14, 2006". Chapel Hill, N.C.: U.S. Environmental Protection Agency National Center for Environmental Assessment.

Brown JS. 2007a. *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.: U.S. Environmental Protection Agency National Center for Environmental Assessment.

Brown JS. 2007b. *The effects of ozone on lung function at 0.06 ppm in healthy adults (January 2007 draft)*. U.S. Environmental Protection Agency National Center for Environmental Assessment.

Carroll L. 1960. The Annotated Alice: Alice's Adventures in Wonderland & Through the Looking Glass. New York, N.Y.: Forum Books.

Chen L, Jennison BL, Yang W, Omaye ST. 2000. Elementary School Absenteeism and Air Pollution. *Inhalation Toxicology* 12: 997-1016.

Chen Y, Blaser MJ. 2007. Inverse associations of *Helicobacter pylori* with asthma and allergy. *Archives of Internal Medicine* 167(April 23): 821-827.

Chen Y, Blaser MJ. 2008. *Heliobacter pylori* colonization is inversely associated with childhood asthma. *Journal of Infectious Disease* 198: 553-560.

Chowienczyk PJ, Parkin DH, Lawson CP, Cochrane GM. 1994. Do Asthmatic Patients Correctly Record Home Spirometry Measurements? *BMJ* 309: 1618.

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.

Cox WM, Camalier L. 2006. *Memorandum: The Effect of Measurement Error on 8-hour Ozone Design Concentrations*. "Cox & Camalier Memo". Research Traingle Park, N.C.: U.S. Environmental Protection Agency Ofice of Air Quality Planning and Standards.

Eaton T, Withy S, Garrett JE, Mercer J, Whitlock RML, Rea HH. 1999. Spirometry in Primary Care Practice: The Importance of Quality Assurance and the Impact of Spirometry Workshops. *Chest* 116(2): 416-423.

Folinsbee LJ, McDonnell WF, Horstman DH. 1988. Pulmonary function and symptom responses after 6.6 hour exposure to 0.12 ppm ozone with moderate exercise. *Journal of the Air Pollution Control Association* 38: 28-35.

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.

Gilliland FD, Berhane K, Rappaport EB, Thomas DC, Avol E, Gauderman WJ, *et al.* 2001. The Effects of Ambient Air Pollution on School Absenteeism Due to Respiratory Illnesses. *Epidemiology* 12: 43-54.

Gong H, Jr,, Wong RJ, 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(2): 1080-1087.

Gradient Corporation. 2007. *Comments on the National Ambient Air Quality Standard for Ozone: US EPA Federal Register Proposed Rule; Docket No. EOPA-HQ-OAR-2005-0172-4183.1, Attachment 2*. Cambridge, Mass.

Hazucha MJ, Folinsbee LJ, Seal Jr. E. 1992. Effects of steady-state and variable ozone concentration profiles on pulmonary function. *American Review of Respiratory Disease* 146(6): 1487-1493.

Henderson R. 2005a. *June 22, 2005 Letter to Stephen L. Johnson (Administrator), 'Clean Air Scientific Advisory Committee (CASAC) Ozone Review Panel's Peer Review of the Agency's Air Quality Criteria for Ozone and Related Photochemical Oxidants (First External Review Draft), Volumes I, II, and III, (EPA/600/R–05/004aA, bA, and cA, January 2005)'*. "CASAC June 22, 2006 Letter on O3 Review". Washington, D.C.

Henderson R. 2005b. *October 25, 2005 Letter to Stephen L. Johnson from CASAC on Clean Air Scientific Advisory Committee (CASAC) Ozone Review Panel, 'Consultation on EPA's Draft Ozone Environmental Assessment Plan: Scope and Methods for Exposure, Risk and Benefits Assessment (August 2005)'*. "CASAC October 25, 2005 Letter on O3 Review". Washington, D.C.

Henderson R. 2006a. *February 10, 2006 Letter to Stephen L. Johnson from CASAC on Clean Air Scientific Advisory Committee's Peer Review of the Agency's Air Quality Criteria for Ozone and Related Photochemical Oxidants (Second External Review Draft), Volumes I, II, and III, (EPA/600/R–05/004aB, bB, and cB, August 2005).* "CASAC February 10, 2006 Letter on O3 Review". Washington, D.C.: U.S. Environmental Protection Agency Office of the Science Advisory Board.

Henderson R. 2006b. *June 5, 2006 Letter to Stephen L. Johnson from CASAC on 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).* "CASAC October 24, 2006 Letter on O3 Review". Washington, D.C.

Henderson R. 2006c. *October 24, 2006 Letter to Stephen L. Johnson from CASAC on Clean Air Scientific Advisory Committee's (CASAC) Review of the Agency's Second Draft Ozone Staff Paper.* "CASAC October 24, 2006 Letter on O3 Review" (Committee CAASA, ed). Washington, D.C.

Henderson R. 2007a. *March 26, 2007 Letter to Stephen L. Johnson on Clean Air Scientific Advisory Committee's Review of the Agency's Final Ozone Staff Paper.* "CASAC March 26, 2007 Letter on O3 Review". Washington, D.C.

Henderson R, 2007b, *March 26, 2007 Letter: Clean Air Scientific Advisory Committee's (CASAC) Review of the Agency's Final Ozone Staff Paper.* Committee CAASA, Secondary March 26, 2007 Letter: Clean Air Scientific Advisory Committee's (CASAC) Review of the Agency's Final Ozone Staff Paper.

Henderson R. 2008a. *April 7, 2008 Letter to Stephen L. Johnson from CASAC on 'Clean Air Scientific Advisory Committee Recommendations Concerning the Final Rule for the National Ambient Air Quality Standards for Ozone'*. "CASAC April 7, 2008 Letter on O_3 NAAQS". Washington, D.C.: U.S. Environmental Protection Agency Office of the Science Advisory Board.

Henderson R. 2008b. *January 23, 2008 Letter to Stephen L. Johnson from CASAC on Clean Air Scientific Advisory Committee Recommendations Concerning the Implementation of EPA's Revised NAAQS Review Process.* "CASAC January

23, 2008 Letter on NAAQS Review Process". Washington, D.C.: U.S. Environmental Protection Agency Office of the Science Advisory Board.

Henderson R. 2008c. *June 7, 2008 Letter to Stephen L. Johnson on Clean Air Scientific Advisory Committee Recommendations Concerning the Final Rule for the National Ambient Air Quality Standards for Ozone*. "CASAC June 7, 2008 Letter on O3 Review".

Hoaglin DC, Mosteller F, Tukey JW, eds. 1983. *Understanding Robust and Exploratory Data Analysis*. New York, N.Y.: Wiley Interscience.

Horstman DH, Ball BA, Brown J, Gerrity T, Folinsbee LJ. 1995. Comparison of Pulmonary Responses of Asthmatic and Nonasthmatic Subjects Performing Light Exercise While Exposed to a Low Level of Ozone *Toxicology and Industrial Health* 14(4): 369-385.

Horstman DH, Folinsbee LJ, Ives PJ, Abdul-Salaam S, McDonnell WF. 1990. Ozone Concentration and Pulmonary Response Relationships for 6.6-hour Exposures with Five Hours of Moderate Exercise to 0.08, 0.10, and 0.12 ppm. *American Review of Respiratory Disease* 142(5): 1158-1163.

Information Quality Act, 44 USC 3516 note.

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

Kennedy P. 1985. A Guide to Econometrics. 2d ed. Cambridge, Mass.: MIT Press.

Koop G, Tole L. 2004. Measuring the Health Effects of Air Pollution: To What Extent Can we Really Say that People are Dying from Bad Air? *Journal of Environmental Economics and Management* 47(1): 25.

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. *Environmental Health Perspectives* 106: 93-99.

Langstaff J. 2006a. *Analysis of Uncertainty in Ozone Population Exposure Modeling (Draft)*. "Draft Langstaff Memo". Research Triangle Park, N.C.

Langstaff J. 2006b. *Memorandum: Analysis of Uncertainty in Ozone Population Exposure Modeling (Draft)*. Research Triangle Park, N.C., 1-87.

Langstaff J. 2007. *Analysis of Uncertainty in Ozone Population Exposure Modeling (Final)*. "Final Langstaff Memo". Research Triangle Park, N.C.

Levy JI, Chemerynski SM, Sarnat JA. 2005. Ozone Exposure and Mortality: An Empiric Bayes Metaregression Analysis. *Epidemiology* 16(4): 458-468.

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

Lutter R, Wolz C. 1997. UV-B Screening by Tropospheric Ozone: Implications for the National Ambient Air Quality Standard. *Environmental science & technology* 31(3): 142A-146A.

MacIntyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CPM, Brusasco V, *et al.* 2005. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *European Respiratory Journal* 26: 16.

McCluney L. 2007. *Memorandum: Ozone 1-hour to 8-Hour Ratios for the 2002-2004 Design Value Period*. "McCluney Memo". Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 1-11.

McCluney L, Rizzo M, Ross M. 2006. *Development of Descriptive Statistics for 8-hr Daily Maximum Ozone Data from Epidemiological Studies*. "McCluney/Rizzo/Ross Memo". Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

McDonnell WF. 1996. Individual Variability in Human Lung Function Responses to Ozone Exposure. *Environmental Toxicology and Pharmacology* 2: 171-175.

McDonnell WF, Kehrl HR, Abdul-Salaam S, Ives PJ, Folinsbee LJ, Devlin RB, *et al.* 1991. Respiratory Response of Humans Exposed to Low Levels of Ozone for 6.6 Hours. *Archives of Environmental Health* 46(3): 145-150.

Meyers RJ. 2008. *Letter from Robert J. Meyers, Principal Deputy Assistant Administrator, Office of Air and Radiation, U.S. Environmental Protection Agency, to John Engler, President and CEO, National Association of Manufacturers*. Washington, D.C.: U..S. Environmental Protection Agency, Office of Air and Radiation.

Miller MR, Crapo R, Hankinson J, Brusasco V, Burgos F, Casaburi R, *et al.* 2005. General considerations for lung function testing. *European Respiratory Journal* 26(1): 153-161.

Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, *et al.* 2005. Standardisation of spirometry. *European Respiratory Journal* 26(2): 319-338.

Moolgavkar S. 2007. *A Critical Review of the Staff Paper on Ozone; Docket No. EPA-HQ-OAR-2005-0172-0493.1*.

Moolgavkar SH. 2000. Air Pollution and Mortality in Three U.S. Counties. *Environmental 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 Journal* 19(4): 699-705.

National Association of Manufacturers. 2007. *Letter and Comments Sent to Molly A. O'Neill (Assistant Administrator, Office of Environmental Information and Chief Information Officer) re: Proposed Rule – National Ambient Air Quality Standards for Ozone; Docket No. OAR-2005-0172-4275*.

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. 1999. *Circular A-110: Uniform Administrative Requirements for Grants and Agreements With Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations*. "OMB Circular A-110". Washington, D.C.

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.

Office of Management and Budget. 2006. *Standards and Guidelines for Statistical Surveys*. Washington, D.C.

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

Peacock M, 2005, *Memorandum for Assistant Administrator (Research and Development) George Gray and Acting Assistant Administrator (Air and Radiation) Bill Wehrum, 'National Ambient Air Quality Standards'*. U.S. Environmental Protection Agency Office of the Deputy Administrator, Secondary Memorandum for Assistant Administrator (Research and Development) George Gray and Acting Assistant Administrator (Air and Radiation) Bill Wehrum, 'National Ambient Air Quality Standards'.

Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, *et al.* 2005. Interpretative strategies for lung function tests. *European Respiratory Journal* 26: 948-968.

Pinkerton KE, Balmes JR, Fanucchi MV, Rom WN. 2007. Ozone, a Malady for All Ages. *American Journal of Respiratory and Critical Care Medicine* 176(2): 107-108.

Reddel HK, Toelle BG, Marks GB, Ware SI, Jenkins CR, Woolcock AJ. 2002. Analysis of Adherence to Peak Flow Monitoring when Recording of Data Is Electronic. *BMJ* 324(7330): 146-147.

Rizzo M. 2005. *Memorandum: A Comparison between Different Rollback Methodologies Applied to Ambient Ozone Concentrations (Draft)*. Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

Rizzo M. 2006. *Memorandum: A Comparison between Different Rollback Methodologies Applied to Ambient Ozone Concentrations (Corrected)*. Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

Romieu I, Meneses F, Ruiz S, Huerta J, Sienra JJ, White M, *et al.* 1997. Effects of intermittent ozone exposure on peak expiratory flow and respiratory symptoms among asthmatic children in Mexico City. *Archives of Environmental Health* 52: 368–376.

Romieu I, Meneses F, Ruiz S, Sienra JJ, Huerta J, White MC, *et al.* 1996. Effects of Air Pollution on the Respiratory Health of Asthmatic Children Living in Mexico City. *American Journal of Respiratory and Critical Care Medicine* 154(2): 300-307.

Ross MA, Persky VW, Scheff PA, Chung J, Curtis L, Ramakrishnan V, *et al.* 2002. Effect of Ozone and Aeroallergens on the Respiratory Health of Asthmatics. *Archives of Environmental Health* 57: 568-578.

Sarafino EP, Gates M, DePaulo D. 2001. The role of age at asthma 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, Koutrakis P, Suh HH. 2000. Assessing the Relationship between Personal Particulate Matter and Gaseous Exposures of Senior Citizens

Living in Baltimore, MD. *Journal of the Air & Waste Management Association* 50: 1184-1198.

Sarnat JA, Schwartz J, Catalano PJ, Suh HH. 2001. Gaseous pollutants in particulate matter epidemiology: confounders or surrogates? *Environmental 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.

Smith AE, Gibbs B. 2007. *Comments on EPA's Proposed Rule to Revise the National Ambient Air Quality Standards for Ozone (72FR 37818); Docket No. EPA-HQ-OAR-2005-0172-4183.1, Enclosure 3*. Washington, D.C.: CRA International.

Smith RL. 2007a. *Public Comment on EPA Draft Ozone Standard; Docket No. EPA-HQ-OAR-2005-0172-4143.1*.

Smith RL. 2007b. *Public Comment to CASAC Ozone Review Panel; Docket No. EPA-HQ-OAR-2005-0172-0080.1*.

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

Thurston GD, Bates DV. 2004. Air Pollution as an Underappreciated Cause of Asthma Symptoms. *JAMA* 290(14): 3.

Tukey JW. 1977. Exploratory Data Analysis. Reading, Mass.: Addison-Wesley.

U.S. Environmental Protection Agency. 1996a. *Air Quality Criteria for Ozone and Related Photochemical Oxidants*. "O3 AQCD". Washington, D.C.

U.S. Environmental Protection Agency. 1996b. National Ambient Air Quality Standards for Ozone: Proposed Decision. *Federal Register* 61(241): 65716-65750.

U.S. Environmental Protection Agency. 1997. *Review of National Ambient Air Quality Standards for Ozone; Assessment of Scientific and Technical Information (OAQPS Staff Paper)*. "1997 Ozone NAAQS Staff Paper". Research Triangle Park, N.C.: U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards.

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)*. "EPA IQG".

U.S. Environmental Protection Agency. 2003. *A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information*.

"General Assessment Factors Guidance" (U.S. Environmental Protection Agency, ed). Washington, D.C.

U.S. Environmental Protection Agency. 2005a. *First Draft Ozone Health Risk Assessment for Selected Urban Areas: October 2005 (Appendices)*. "2005 1st Draft O_3 Health Risk Assessment Appendices" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 135.

U.S. Environmental Protection Agency. 2005b. *First Draft Ozone Health Risk Assessment for Selected Urban Areas: October 2005 Draft Report*. "1st Draft O3 Heath Risk Assessment" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 135.

U.S. Environmental Protection Agency. 2005c. *First Draft Ozone Population Exposure Analysis for Selected Urban Areas; November 2005*. "2005 1st Draft O_3 Exposure Analysis" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 135.

U.S. Environmental Protection Agency. 2005d. *Plan for Review of the National Ambient Air Quality Standards for Ozone*. "O3 Review Plan" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 1-22.

U.S. Environmental Protection Agency. 2005e. *Plan for Review of the National Ambient Air Quality Standards for Ozone*. "O_3 Review Plan" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2006a. *Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume I of III (EPA 600/R-05/004aF)*. "2006 O_3 CD Vol I". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Research and Development National Center for Environmental Assessment-RTP Office.

U.S. Environmental Protection Agency. 2006b. *Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume II of III (EPA 600/R-05/004bF)*. "2006 O_3 CD Vol II". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Research and Development National Center for Environmental Assessment-RTP Office.

U.S. Environmental Protection Agency. 2006c. *Air Quality Criteria for Ozone and Related Photochemical Oxidants: Volume III of III (EPA 600/R-05/004cF)*. "2006 O_3 CD Vol III".

U.S. Environmental Protection Agency, 2006d, *EPA's Action Development Process: Guide to Considering Children's Health When Developing EPA Actions; Implementing Executive Order 13045 and EPA's Policy on Evaluating Health Risks to Children*. U.S. Environmental Protection Agency, Secondary EPA's Action Development Process: Guide to Considering Children's Health When Developing EPA Actions; Implementing Executive Order 13045 and EPA's Policy on Evaluating Health Risks to Children.

U.S. Environmental Protection Agency, 2006e, *Peer Review Handbook*. Secondary Peer Review Handbook.

U.S. Environmental Protection Agency. 2006f. *Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information: Second Draft OAQPS Staff Paper*. "2006 2nd Draft O_3 Staff Paper". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2006g. *Second Draft Ozone Health Risk Assessment for Selected Urban Areas (Appendices); June 2006*. "2006 2nd Draft O_3 Health Risk Assessment Appendices" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency, 151.

U.S. Environmental Protection Agency. 2006h. *Second Draft Ozone Health Risk Assessment for Selected Urban Areas; June 2006*. "2006 2nd Draft O_3 Health Risk Assessment" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2006i. *Second Draft Ozone Population Exposure Analysis for Selected Urban Areas; July 2006*. "2006 2nd Draft O_3 Exposure Analysis" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2007a. *Climate Change -- Science; State of Knowledge*. Available: http://www.epa.gov/climatechange/science/stateofknowledge.html [accessed September 7, 2008].

U.S. Environmental Protection Agency. 2007b. *Draft Integrated Science Assessment for Oxides of Nitrogen – Health Criteria; August 2007*. "1st Draft NOx ISA ". Washington, D.C.

U.S. Environmental Protection Agency. 2007c. *Final Ozone Health Risk Assessment for Selected Urban Areas; Updated July 2007 (EPA 452/R-07-009)*. "2007 Final O_3 Health Risk Assessment (Updated July)" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2007d. *Final Ozone Health Risk Assessment for Selected Urban Areas; Updated July 2007 (EPA 452/R-07-009)*. "Final O3 Health Risk Assessment (Updated)" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2007e. *Final Ozone Population Exposure Analysis for Selected Urban Areas; Updated July 2007 (EPA-452-/R-07-010)*. "Final O3 Exposure Analysis (Updated)" (U.S. Environmental Protection Agency Office of Air Quality Planning and Standards, ed). Research Triangle Park, N.C.: U.S. Environmental Protection Agency.

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

U.S. Environmental Protection Agency. 2007g. *Final Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information: OAQPS Staff Paper Appendices (EPA-452/R-07-007)*. "2007 Final O_3 Staff Paper Appendices (Updated July)". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2007h. National Ambient Air Quality Standards for Ozone: Proposed Rule. *Federal Register* 72(132): 37818-37919

U.S. Environmental Protection Agency. 2007i. *Regulatory Impact Analyses: 2007 Proposed Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone*. "O3 Regulatory Impact Analysis".

U.S. Environmental Protection Agency. 2007j. *Review of the National Ambient Air Quality Standards for Ozone: Policy Assessment of Scientific and Technical Information: Final OAQPS Staff Paper (EPA-452/R-07-007)*. "2007 Final O_3 Staff Paper (Updated July)". Research Triangle Park, N.C.: U.S.

Environmental Protection Agency Office of Air Quality Planning and Standards.

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

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

U.S. Environmental Protection Agency. 2007m. *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)*. "Staff Paper (Final Updated July 2007)".

U.S. Environmental Protection Agency. 2007n. *SF_6 Emission Reduction Partnership for Electric Power Systems: 2006 Annual Report*. Washington, D.C.: U.S. Environmental Protection Agency.

U.S. Environmental Protection Agency. 2008a. *Final Regulatory Impact Analyses: Revisions to the National Ambient Air Quality Standards for Ground-Level Ozone*. "2007 Final O_3 NAAQS RIA". Research Triangle Park, N.C.: U.S. Environmental Protection Agency Office of Air Quality Planning and Standards.

U.S. Environmental Protection Agency. 2008b. National Ambient Air Quality Standards for Ozone; Final Rule. *Federal Register* 73(60): 16436-16514.

U.S. Environmental Protection Agency. 2008c. Responses to NAM Comments on the 2007 Proposed Rule on the National Ambient Air Quality Standards for Ozone. In: *Email from Lydia Wegman, USEPA to Office of Management and Budget (OMB); Staff Subject: National Association of Manufacturers (NAM) and exposure/risk comment/responses (March 11, 2008); Docket No EPA-HQ-OAR-2005-0172-71521* (U.S. Environmental Protection Agency, ed). Washington, D.C.: U.S. Environmental Protection Agency, Docket No. EPA-HQ-OAR-2005-0172.

U.S. Environmental Protection Agency, 2008d, *Responses to Significant Comments on the 2007 Proposed Rule on the National Ambient Air Quality Standards for Ozone (July 11, 2007; 72 FR 37818); Docket ID EPA-HQ-OAR-2005-0172-7152.1*. U.S. Environmental Protection Agency, Secondary Responses to Significant Comments on the 2007 Proposed Rule on the National

Ambient Air Quality Standards for Ozone (July 11, 2007; 72 FR 37818); Docket ID EPA-HQ-OAR-2005-0172-7152.1.

U.S. Environmental Protection Agency. 2008e. *Responses to Significant Comments on the 2007 Proposed Rule on the National Ambient Air Quality Standards for Ozone (July 11, 2007; 72 FR 37818); Docket ID EPA-HQ-OAR-2005-0172-7152.1.* "2008 O_3 Response to Comments" (U.S. Environmental Protection Agency, ed). Washington, D.C.: U.S. Environmental Protection Agency

U.S. Environmental Protection Agency Office of Environmental Information. 2006. *Pre-Dissemination Review Guidelines: Review Guidelines to Ensure That Disseminated Information Is Consistent with EPA Information Quality Guidelines.* "EPA Pre-Dissemination Review Guidelines".

U.S. Environmental Protection Agency Office of the Science Advisor. 2004a. *An Examination of EPA Risk Assessment Principles and Practices.* Available: http://www.epa.gov/osainter/pdfs/ratf-final.pdf [accessed August 30 2005].

U.S. Environmental Protection Agency Office of the Science Advisor. 2004b. *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].

U.S. Environmental Protection Agency Science Advisory Board Staff Office. 2006. *Public Advisory Meeting: Meeting to Conduct a Peer Review of EPA's 2nd Draft Ozone Staff Paper and Related Draft Technical Support Documents (TSDs).* "CASAC O_3 Panel August 24, 2006 Meeting Transcript". Durham, N.C.: U. S. Environmental Protection Agency, EPA Science Advisory Board Staff Office.

Vaughan TR, Weber RW, Tipton WR, Nelson HS. 1989. Comparison of PEFR and FEV1 in Patients with Varying Degrees of Airway Obstruction; Effect of Modest Altitude. *Chest* 95: 558-562.

Vingarzan R. 2004. A Review of Surface Ozone Background Levels and Trends. *Atmospheric Environment* 38: 3431-3442.

Vu V, 2005a, *Additional CASAC Member Individual Comments on EPA's NAAQS Process*. U.S. Environmental Protection Agency Office of the Deputy Administrator, Secondary Additional CASAC Member Individual Comments on EPA's NAAQS Process.

Vu V. 2005b. *Memorandum for Assistant Administrator (Research and Development) George Gray and Acting Assistant Administrator (Air and Radiation) Bill Wehrum, 'Clean Air Scientific Advisory Committee Members' Comments on the*

Agency's Process for Establishing National Ambient Air Quality Standards'. Washington, D.C.: U.S. Environmental Protection Agency Science Advisory Board.

Wanger J, Coates A, Pedersen OF, Brusasco V, Burgos F, Casaburi R, *et al.* 2005. Standardisation of the measurement of lung volumes. *European Respiratory Journal* 26: 511-522.

World Health Organization. 2000. *Air Quality Guidelines for Europe*. "WHO Regional Publications, European Series, No 91".

World Health Organization. 2006. *WHO Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxides and Sulfur Dioxide: Global Update 2005; Summary of Risk Assessment*.