UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON D.C. 20460

OFFICE OF THE ADMINISTRATOR
SCIENCE ADVISORY BOARD

September 27, 2007

EPA-CASAC-07-007

Honorable Stephen L. Johnson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, NW
Washington, DC 20460

Subject: Clean Air Scientific Advisory Committee's (CASAC) Review of the 2nd Draft Lead Human Exposure and Health Risk Assessments Document

Dear Administrator Johnson:

The Clean Air Scientific Advisory Committee (CASAC or Committee), augmented by subject-matter-expert Panelists — collectively referred to as the CASAC Lead Review Panel (Lead Panel) — met on August 28–29, 2007, in Durham, NC, at the request of EPA's Office of Air Quality Planning and Standards (OAQPS) to conduct a peer review of the Agency's *Lead Human Exposure and Health Risk Assessments for Selected Case Studies, Draft Report* (2nd Draft Lead Human Exposure and Health Risk Assessments, July 2007). This letter provides the Lead Panel's advice and recommendations to you concerning the Agency's exposure/risk assessment that supports the setting of a primary National Ambient Air Quality Standard (NAAQS) for Lead, the secondary Lead NAAQS, and implementation issues associated with EPA's revised NAAQS review process. The CASAC roster is found in Appendix A of this report, and the Lead Panel roster is attached as Appendix B. The charge questions provided to the Lead Panel by EPA staff are contained in Appendix C to this report, and Panelists' individual review comments are provided in Appendix D.

Review of EPA's 2nd Draft Human Lead Exposure and Health Risk Assessments

1. Overall Evaluation

Overall, the CASAC Lead Review Panel judges that *the Agency's 2nd Draft Lead Human Exposure and Health Risk Assessments is not yet a complete, well-documented exposure and risk assessment that presents the full range of pertinent data and analyses.* In particular, in order to support the establishment of National Ambient Air Quality Standards, it is especially important that the EPA develop exposure estimates that will have *national* implications for, and relevance to, urban areas. Agency staff needs to undertake additional case studies of several urban locales

with varying lead exposure levels. More generally, in the view of the Committee, the second draft exposure/risk assessments document was missing certain critical information and is deemed incomplete, which is described in greater detail below. *Therefore, this EPA document, while representing a worthy effort by Agency staff, is not yet deemed to be adequate for regulatory decision-making.* The CASAC looks forward to reviewing OAQPS' Final Lead Exposure and Risk Assessments document — and, especially, the Final Staff Paper for Lead — when these are released by EPA in early November 2007.

2. Reiteration of CASAC Support for Continuing to List Lead as a Criteria Air Pollutant and the Lead Panel's Preliminary Analyses Concerning the Level of a Primary (Health-Based) NAAQS for Lead

Before discussing the Lead Panel's review of the EPA's 2nd Draft Lead Exposure and Risk Assessments document, *the Committee wishes to strongly reiterate its opposition to any considered de-listing of Lead as a criteria air pollutant and its concomitant and unanimous support for maintaining fully-protective NAAQS.* The details of the CASAC's rationale for this recommendation are contained in the Committee's March 27, 2007 letter to you concerning the Lead Panel's review of the Agency's 1st Draft Lead Staff Paper and the Draft Lead Exposure and Risk Assessments documents (EPA-CASAC-07-003).

Furthermore, as described in detail in Appendix D of the Committee's March 2007 letter/report, the Lead Panel previously considered three separate, though related, population-based analytical approaches aimed at deriving an estimated range of alternative levels for the primary Lead NAAQS. On the basis of the CASAC's preliminary scientific analyses and risk management assumptions, *EPA needs to substantially lower the level of the primary NAAQS for Lead, to 0.2 µg/m^3 or less. In the unanimous opinion of the Lead Panel, EPA has not presented any rigorous analyses or other information in its 2nd Draft Lead Exposure and Risk Assessments document that leads the CASAC to reconsider its previous recommendation to you that the upper limit the Agency should consider in revising the Pb NAAQS should be 0.2 µg/m^3 on a monthly average.*

3. Need Population-Based Risk Assessments of Urban Areas of National Significance

In the CASAC's previous letter to you on this topic (March 2007), the Lead Review Panel recommended using a "population-based" risk assessment to supplement the case-study approach used in the "pilot-phase" risk assessment. In addition, the Panel noted that a risk assessment of this type would typically include two key components:

(1) A quantitative description of the relationship between concentrations of lead in ambient air in various parts of the U.S. and resulting distributions of blood lead concentrations; and

(2) A quantitative description of the relationship between blood lead concentrations levels and impacts on IQ.

The Lead Panel was pleased to see that EPA introduced an *urban* model in the second draft of the exposure/risk assessments document, and further recommends that Agency staff focus on the *hybrid* urban model, using available information for several urban areas where there

are multiple monitors. The Committee recognizes that the nature of the national airborne lead database is limited by both its size (*i.e.*, a total of 189 PM-TSP [total suspended particulates] sites measuring ambient lead), as well as the selective location of monitors to mostly source-oriented sites. Some additional PM_{10} measurements are available (see the section below on the use of PM_{10} samplers to monitor lead).

In spite of these limitations, the CASAC strongly believes that it is important that EPA staff make estimates of exposure that will have *national* implications for, and relevance to, *urban* areas; and that, significantly, *the case studies of both primary lead (Pb) smelter sites as well as secondary smelter sites, while relevant to a few atypical locations, do not meet the needs of supporting a Lead NAA*QS. *The Agency should also undertake case studies of several urban areas with varying lead exposure concentrations*, based on the prototypic urban risk assessment that OAQPS produced in the 2nd Draft Lead Human Exposure and Health Risk Assessments. In order to estimate the magnitude of risk, the Agency should estimate exposures and convert these exposures to estimates of blood levels and IQ loss for children living in specific urban areas. In addition, the Agency should make IQ change estimates across the range of exposures to provide estimates for the change in median, as well as the 5th- to the 95th percentile of the population for different standards using the hybrid urban model with a geometric standard deviation (GSD) of 2.0 or 2.1.

The Lead Panel recognizes that there are few urban areas with multiple TSP monitors to estimate distributions of lead exposure. The Panel urges that PM_{10} monitors, with appropriate adjustments, be used to supplement the data. If necessary, other data with lead concentrations from special monitoring studies (*e.g.*, speciation studies of particulate matter [PM]) may provide estimates of the GSD of air lead over urban areas, which could also be used to supplement the limited TSP data. Discussion of the risk estimates obtained should carry appropriate caveats that document where estimates fall outside the range of data used to generate the estimates. EPA should also provide a qualitative discussion of how typical the chosen cities are of the range of what is seen in a broad spectrum of U.S. urban areas.

4. Completeness of 2nd Draft Lead Exposure and Risk Assessments

In a more general sense, the CASAC considers that EPA's 2nd Draft Lead Exposure and Risk Assessments document was missing certain key information and analytical components and was therefore incomplete. Specifically, the Committee believes that a properly comprehensive exposure/risk assessment in support of reviewing either primary or secondary NAAQS should be accompanied by documentation that includes a discussion of the four, policy-relevant elements of selecting a specific NAAQS — that is, indicator variable, averaging time, statistical form, and ranges of alternative levels of the standard — along with analyses that model the impact of current ("as is") standards and any proposed alternatives (*i.e.*, a truly *quantitative* risk assessment). In addition, the Agency needs to provide details on both the scope of the exposure assessment and the results from the modeling, to include components that would describe: selection of urban areas to be modeled; the time periods and the populations that were modeled; and the results from modeling the current lead standard and proposed alternative NAAQS.

Furthermore, with respect to the primary (public-health based) Lead NAAQS, a complete public-health risk assessment should include documentation that provides appropriate details on

both the scope of the health risk model and the associated modeling results vis-à-vis: selection of health endpoint categories; selection of study areas; and air quality considerations for both the current NAAQS and any proposed alternative levels of the standard. Finally, the exposure/risk assessment should include a thorough discussion of the key uncertainties.

Accordingly, the CASAC requests that EPA tailor its future exposure/risk assessments to provide the above documentation and analyses for the CASAC's and the public's review during the comparable phase of the NAAQS review process — that is, *prior* to the Agency's issuance of the Policy Assessment (PA) in the form of an Advance Notice of Proposed Rulemaking (ANPR) — for any given criteria air pollutant.

5. Choice of Biokinetic Model and Steady-State Dose Metric for Both Dose-Response Functions and Use in the Full Risk Assessment

Agency staff selected the Integrated Exposure Uptake Biokinetic (IEUBK) model for lead in children as the preferred method for estimating blood lead as the dose metric in the risk assessment, based on its overall assessment of blood lead estimating methods. The Lead Panel concurs in both OAQPS' selection of the IEUBK model as the biokinetic modeling method for blood lead estimates and its detailed rationale for doing so. The rationale provided by EPA staff for its choice is, overall, scientifically-sound.

In addition, OAQPS selected concurrent blood lead estimates as the best expression for the dose parameter used in both the dose-response functions and the risk assessment. The Panel also concurs in both this selection by Agency staff and the scientific rationales for doing so, particularly its reliance on the findings of the international pooled analysis of many longitudinal studies of cognitive deficits at lower exposures (doses) reported in Lanphear *et al.* (2005).

6. Predicting IQ Changes Based on Concurrent Blood Lead Concentrations

The Panel recommends using the two-piece linear function for relating IQ alterations to current blood lead levels with a slope change or "hinge" point closer to 7.5 µg/dL than 10.82 µg/dL as used by EPA staff in the second draft exposure/risk assessments document. The higher value used by staff underestimates risk at lower blood Pb levels, where most of the population will be located. Epidemiologic data indicate that the slope of the line below 7.5 µg/dL is approximately minus three (-3) IQ decrements per 1 µg/dL blood lead and the vast majority of children in the U.S. have maximal baseline Pb blood levels below 7.5 µg/dL (Lanphear *et al.*, EHP 2005; MMWR 2005). On a population level, the mean increase in blood lead concentration from airborne lead would generally be up to, but not exceeding, a blood lead concentration of 7.5 µg/dL. This approach should also account for sensitive subpopulations of children.

7. Level and Averaging Time for Primary Lead Standard with a Margin of Safety

The most recent epidemiologic studies demonstrate a statistically significant relationship between blood Pb and IQ loss well below 5 µg/dL. The CASAC recognizes that lead is a multimedia pollutant and that most of the country is in compliance with the current Lead NAAQS of 1.5 $\mu g/m^3$. However, the risk analysis scenarios presented by EPA for current conditions using the Agency's hybrid dust model — which the Lead Panel judges to be the most scientifically-defensible and robust dust model currently available — show that the "recent" air exposure path-

way contributes anywhere from 28 to 57% of the total amount of ingested lead. Additionally, recent air exposures still contribute 27% under an alternative primary Pb NAAQS of 0.2 $\mu g/m^3$ maximum monthly average, and only fall to 13% under an alternative primary Lead standard of 0.05 $\mu g/m^3$ maximum monthly average. *Since there is no known threshold in the relationship between blood Pb and IQ loss, the level of the current primary Lead standard clearly provides no margin of safety from ambient air lead exposures.* However, nor would any lower primary Lead NAAQS level provide a margin of safety, and hence, the question becomes:

> What percentage of the population of children in various parts of the U.S. will suffer what amount of IQ loss and other harmful effects due to the contribution of air exposures to the overall toxicity of Pb?

EPA staff should identify what levels of the primary Lead NAAQS would be deemed as being adequately-protective of human health. As a preliminary target, staff should identify the level of the standard that would ensure that 95% or more of the children in the U.S. do not experience decreased IQ from exposure to ambient concentrations of recent airborne lead. As noted in the Committee's previous letter to you on this subject, target levels of IQ decrements that would be of great concern would be one to two (1–2) IQ points or more. After identifying such a level of the standard, Agency staff should investigate alternative levels around this level, including much lower levels, to provide guidance as to how alternative standards would lead to changes in health. For example, if the analyses conducted by EPA staff suggest that a 0.1 $\mu g/m^3$ standard would lead to a decrease in IQ of one point or less for 95% of the children in the U.S., staff should assess other levels of the standard near 0.1 $\mu g/m^3$, both above and below, as well as much lower levels, *e.g.*, on the order of 0.05 and 0.01 $\mu g/m^3$. Further, the Agency should provide additional analyses to adequately inform both the Administrator and CASAC as to how uncertainties impact the level of protectiveness of the proposed alternative standards.

8. Use of PM_{10} Samplers to Monitor for Airborne Lead

Another recommendation that the CASAC provided in its March 2007 letter was to consider use of PM_{10} samplers to monitor lead. A substantial reduction in the level of the Pb NAAQS, combined with a shortening of the averaging time from quarterly to monthly, will require increases in both the number of lead monitoring sites, as well as the frequency of sample collection. Improved sampling precision will also be needed as more locations fall closer to standards and to support future health assessments as ambient lead concentrations are further reduced. For these and other reasons outlined in our previous advisory letter, *the Lead Panel strongly encourages the Agency to consider revising the Pb reference method to allow sample collection by PM_{10}, rather than TSP samplers,* accompanied by analysis with low-cost multi-elemental techniques like X-Ray Fluorescence (XRF) or Inductively Coupled Plasma–Mass Spectroscopy (ICP-MS). (See EPA-CASAC-07-003 for additional details.)

The Lead Panel also recognizes the importance of coarse dust contributions to total Pb ingestion and acknowledge that TSP sampling is likely to capture additional very coarse particles which are excluded by PM_{10} samplers. However, the precision of TSP samplers is poor, the upper particle cut size varies widely as a function of wind speed and direction, and the spatial non-homogeneity of very coarse particles cannot be efficiently captured by a national monitoring network. Generally, it can be expected that PM_{10} Pb will represent a large fraction of, and be

highly correlated with TSP Pb. Ambient Lead data from the (few) collocated TSP and PM_{10} sites presented in the 1st Draft Pb Staff Paper exhibited a high correlation (r = 0.96), with slopes (PM_{10} Pb/TSP Pb) ranging from 0.85 to 1. A single quantitative adjustment factor could be developed from a short period of collocated sampling at multiple sites; or a PM_{10} Pb/TSP Pb "equivalency ratio" could be determined on a regional or site-specific basis.

9. Other, Non-IQ-Related Effects of Lead in Ambient Air

While the CASAC agrees with the Agency's choice of IQ alterations in young children as the priority health effect and population for the risk assessment, the Lead Panel cautions against focusing only on IQ loss (or gain). There are ramifications of lead exposure on other endpoints that have societal and individual implications of great importance. Neurological developmental and functional effects in children exposed to Pb can lead to negative and disruptive behaviors well into teenage years. Moreover, while the adult nervous system has long been recognized as a target of Pb toxicity, epidemiologic and experimental toxicology data are emerging that support the relationship between Pb exposure and increased adverse cardiovascular outcomes, including increased blood pressure, increased incidence of hypertension, and cardiovascular morbidity and mortality at lower and lower levels of exposure.

Secondary Lead NAAQS

The "pilot phase" lead exposure and risk assessment document addressed both human health and environmental effects, but the July 2007 "full-scale" exposure and risk assessment document is focused entirely on human exposures and health risks. Agency staff made it quite clear in the 1st Draft Lead Staff Paper that OAQPS did not anticipate having either sufficient funding or time available to perform additional quantitative ecological risk assessment work during this current review cycle for the NAAQS for Lead. Thus, EPA staff did not conduct a full-scale *ecological* risk assessment for this second draft exposure/risk assessments document. Nevertheless, *the CASAC requests that EPA revise the ecological portion of the "pilot-phase" lead risk assessment on the basis of Lead Panel members' individual review comments found in Appendix E of the Committee's March 2007 letter; and that this be reflected in the welfare-effects sections of both the Final Lead Exposure and Risk Assessments document and the Final Lead Staff Paper.*

With respect to secondary Pb standards, the Lead Panel notes that the secondary Lead NAAQS was initially set equal to the primary Lead standard in 1978 "due to a lack of relevant data at that time." Nearly 30 years later, it now appears that *the Agency still lacks the relevant data to provide a clear, quantitative basis for setting a secondary Pb NAAQS that differs from the primary in indicator, averaging time, level or form.* To collect such data for the next Lead NAAQS review cycle, *the EPA needs to initiate new measurement activities in rural areas* — including those that are remote, close to urban and other sources, and located at high elevations — which quantify and track changes in lead concentrations in the ambient air, soils, deposition, surface waters, sediments and biota, along with other information as may be needed to calculate and apply a critical loads approach for assessing environmental lead exposures and risks in the next review cycle. Depending on the results of these Pb monitoring activities, the Agency may need to set the level of the secondary Lead NAAQS as a to-be-determined fraction of the level of

the primary standard, given the likelihood that many of the millions of animal and plant species are more sensitive to environmental lead pollution than are humans.

Importantly, EPA needs to move away from the traditional practice of simply setting a secondary Lead NAAQS that is *equal* to the primary standard — a practice that may technically meet the Clean Air Act requirements but has no scientific or technical basis. Nevertheless, in the *absence* of essential monitoring or other research data, the Lead Panel continues to recommend that, "at a minimum, the level of *the secondary Lead NAAQS should be at least as low as the lowest-recommended primary lead standard*," as the CASAC wrote in its previous letter to you on this subject. Furthermore, the Panel also continues to recommend that the Agency "*identify the necessary funds to support needed continuing research on the ecological effects of airborne lead pollution and to consider developing alternative secondary standards such as critical loads for lead, which may be different from primary standards in indicator, averaging time, level or form*." (For additional details, see the individual written comments of Drs. Ellis Cowling and Samuel Luoma found in Appendix D; and also see EPA-CASAC-07-003 and Lead Panelists' individual written comments found in Appendix E.)

The large environmental burden of historically-deposited lead is currently decreasing. Accordingly, *the goal should be to set the secondary Lead NAAQS such that there is no reversal of the current downward trend in lead concentrations in the environment.* The limited funds available for monitoring environmental lead should be focused on this critical task.

Comments on Implementation of the Agency's Revised NAAQS Review Process

The review of the NAAQS for ambient Lead has been a *hybrid* process, which began under the EPA's long-standing NAAQS review process and has since evolved into the Agency's new, revised process. Specifically, the Lead Panel in 2006 conducted a peer-review of the 1st and 2nd Drafts of EPA's Lead Air Quality Criteria Document (AQCD) and, in February 2007, reviewed OAQPS' 1st Draft Lead Staff Paper and the Draft Lead Exposure and Risk Assessments documents. The understanding of the CASAC at the time of the February review was that, since the Agency was transitioning to the revised review process midway through the current Lead NAAQS review, this was to be the last version of either of these documents that OAQPS would develop and that the Lead Panel would have an opportunity to review. Indeed, in the CASAC's last letter to you on this subject (March 27, 2007), the Panel expressed its concern that, in the absence of being given an opportunity to review even a 2nd Draft Lead Staff Paper, or to review a second draft or a final lead exposure/risk assessment document, prior to EPA's issuance of the Lead Policy Assessment in the form of an ANPR, the Committee would not have the information — that is, both the data and the analyses — needed to properly advise you concerning the setting of NAAQS for Lead that would be adequately-protective of both human health and welfare. In response to CASAC's concerns, the Agency agreed to produce a second draft of the exposure/ risk assessments document, the peer-review of which is the topic of this letter/report.

However, immediately prior to our August 28–29 public advisory meeting on the 2nd Draft Lead Human Exposure and Health Risk Assessments, OAQPS staff informed me as the CASAC Chair that a recent (August 24, 2007) Federal court order was requiring the Agency to produce a Final Lead Staff Paper by the previously-agreed-upon date of November 1, 2007 —

and, thus, "the pendulum has swung back" to the former NAAQS review process, at least until the issuance of the Lead PA by means of an ANPR no later than November 30, 2007. Although this seesaw process has admittedly been both confusing and vexing, the Committee is looking forward to reviewing the Final Lead Exposure and Risk Assessments document, the Final Staff Paper for Lead, and the Lead PA in a meeting to be held in mid-December 2007.

Both these process-related perturbations and, as detailed above, the absence of certain critical information in the Agency's 2nd Draft Lead Exposure and Risk Assessments document — which as previously noted was intended to be the CASAC's final source of information from EPA with respect to the review of the NAAQS for Lead prior to the issuance of a Policy Assessment in an ANPR — underscore the Committee's concerns about the Agency's revised NAAQS review process as it is presently being implemented and may be implemented in the future. The CASAC has a statutory mandate to provide the EPA with expert advice and recommendations on scientifically-appropriate standards for criteria air pollutants. In order to be able to fulfill this, it is axiomatic that the CASAC must receive, in a timely manner, and be afforded an opportunity to review and comment on the complete suite of relevant risk- and exposure-related data and analyses that will presumably underpin the Agency's regulatory decisions — not only for the Lead standards but also for forthcoming risk/exposure assessments associated with the NAAQS reviews for other criteria pollutants.

In closing, the CASAC is pleased to advise you and OAQPS staff on the 2nd Draft Lead Human Exposure and Health Risk Assessments document. As both EPA and the Committee continue to work through the details associated with implementation of the revised NAAQS review process, we would ask that the Agency ensure that the CASAC receive the full breadth of information and supporting analyses necessary to provide timely, expert advice and recommendations to the EPA. As always, we wish Agency staff well in this important task.

Sincerely,

/Signed/

Dr. Rogene Henderson, Chair
Clean Air Scientific Advisory Committee

Appendices (A–D)

NOTICE

This report has been written as part of the activities of the U.S. Environmental Protection Agency's (EPA) Clean Air Scientific Advisory Committee (CASAC), a Federal advisory committee administratively-located under the EPA Science Advisory Board (SAB) Staff Office that is chartered to provide extramural scientific information and advice to the Administrator and other officials of the EPA. The CASAC is structured to provide balanced, expert assessment of scientific matters related to issue and problems facing the Agency. This report has not been reviewed for approval by the Agency and, hence, the contents of this report do not necessarily represent the views and policies of the EPA, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use. CASAC reports are posted on the SAB Web site at: http://www.epa.gov/sab.

Appendix A – Roster of the Clean Air Scientific Advisory Committee

U.S. Environmental Protection Agency
Science Advisory Board (SAB) Staff Office
Clean Air Scientific Advisory Committee (CASAC)

CHAIR

Dr. Rogene Henderson, Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

MEMBERS

Dr. Ellis Cowling, University Distinguished Professor At-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Dr. James D. Crapo, Professor, Department of Medicine, National Jewish Medical and Research Center, Denver, CO

Dr. Douglas Crawford-Brown, Director, Carolina Environmental Program; Professor, Environmental Sciences and Engineering; and Professor, Public Policy, Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC

Mr. Richard L. Poirot, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

Dr. Armistead (Ted) Russell, Georgia Power Distinguished Professor of Environmental Engineering, Environmental Engineering Group, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

Dr. Frank Speizer, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA

SCIENCE ADVISORY BOARD STAFF

Mr. Fred Butterfield, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov) (Physical/Courier/FedEx Address: Fred A. Butterfield, III, EPA Science Advisory Board Staff Office (Mail Code 1400F), Woodies Building, 1025 F Street, N.W., Room 3604, Washington, DC 20004, Telephone: 202-343-9994)

Appendix B – Roster of the CASAC Lead Review Panel

U.S. Environmental Protection Agency
Science Advisory Board (SAB) Staff Office
Clean Air Scientific Advisory Committee (CASAC)
CASAC Lead Review Panel

CHAIR

Dr. Rogene Henderson*, Scientist Emeritus, Lovelace Respiratory Research Institute, Albuquerque, NM

MEMBERS

Dr. Joshua Cohen, Research Associate Professor of Medicine, Tufts University School of Medicine, Institute for Clinical Research and Health Policy Studies, Center for the Evaluation of Value and Risk, Tufts New England Medical Center, Boston, MA

Dr. Deborah Cory-Slechta, Professor of Environmental Medicine, Department of Environmental Medicine, University of Rochester School of Medicine and Dentistry, Rochester, NY

Dr. Ellis Cowling*, University Distinguished Professor At-Large, North Carolina State University, Colleges of Natural Resources and Agriculture and Life Sciences, North Carolina State University, Raleigh, NC

Dr. James D. Crapo [M.D.]*, Professor, Department of Medicine, National Jewish Medical and Research Center, Denver, CO

Dr. Douglas Crawford-Brown*, Director, Carolina Environmental Program; Professor, Environmental Sciences and Engineering; and Professor, Public Policy, Department of Environmental Sciences and Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC

Dr. Bruce Fowler, Assistant Director for Science, Division of Toxicology and Environmental Medicine, Office of the Director, Agency for Toxic Substances and Disease Registry, U.S. Centers for Disease Control and Prevention (ATSDR/CDC), Chamblee, GA

Dr. Andrew Friedland, Professor and Chair, Environmental Studies Program, Dartmouth College, Hanover, NH

Dr. Robert Goyer [M.D.], Emeritus Professor of Pathology, Faculty of Medicine, University of Western Ontario (Canada), Chapel Hill, NC

Mr. Sean Hays, President, Summit Toxicology, Allenspark, CO

Dr. Bruce Lanphear [M.D.], Sloan Professor of Children's Environmental Health, and the Director of the Cincinnati Children's Environmental Health Center at Cincinnati Children's Hospital Medical Center and the University of Cincinnati, Cincinnati, OH

Dr. Samuel Luoma, Senior Research Hydrologist, U.S. Geological Survey (USGS), Menlo Park, CA

Dr. Frederick J. Miller, Consultant, Cary, NC

Dr. Paul Mushak, Principal, PB Associates, and Visiting Professor, Albert Einstein College of Medicine (New York, NY), Durham, NC

Dr. Michael Newman, Professor of Marine Science, School of Marine Sciences, Virginia Institute of Marine Science, College of William & Mary, Gloucester Point, VA

Mr. Richard L. Poirot*, Environmental Analyst, Air Pollution Control Division, Department of Environmental Conservation, Vermont Agency of Natural Resources, Waterbury, VT

Dr. Michael Rabinowitz, Geochemist, Marine Biological Laboratory, Woods Hole, MA

Dr. Armistead (Ted) Russell*, Georgia Power Distinguished Professor of Environmental Engineering, Environmental Engineering Group, School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA

Dr. Joel Schwartz, Professor, Environmental Health, Harvard University School of Public Health, Boston, MA

Dr. Frank Speizer [M.D.]*, Edward Kass Professor of Medicine, Channing Laboratory, Harvard Medical School, Boston, MA

Dr. Ian von Lindern, Senior Scientist, TerraGraphics Environmental Engineering, Inc., Moscow, ID

Dr. Barbara Zielinska, Research Professor, Division of Atmospheric Science, Desert Research Institute, Reno, NV

SCIENCE ADVISORY BOARD STAFF

Mr. Fred Butterfield, CASAC Designated Federal Officer, 1200 Pennsylvania Avenue, N.W., Washington, DC, 20460, Phone: 202-343-9994, Fax: 202-233-0643 (butterfield.fred@epa.gov)

* Members of the statutory Clean Air Scientific Advisory Committee (CASAC) appointed by the EPA Administrator

Appendix C – Agency Charge to the CASAC Lead Review Panel

Charge to the CASAC Pb Panel

Within each of the main sections of the draft risk assessment report, questions that we ask the Panel to focus on in their review include the following:

Design of the Exposure and Risk Assessments (Chapter 2):

1. To what extent is the presentation of evidence from the health studies assessed in the Pb AQCD and the integration of information from across the various health-related research areas drawn from the Pb AQCD technically sound, appropriately balanced, and clearly communicated?

2. What are the views of the Panel on the appropriateness of staff's discussion and conclusions on key issues related to quantitative interpretation of epidemiologic study results, including, particularly, the form of a blood Pb-response function for neurocognitive effects, and the form of the associated blood Pb metric?

3. What are the views of the Panel on the appropriateness, technical soundness and clarity of staff's presentation of the approaches used for aspects of the assessment such as air and surface soil/dust concentrations for alternate scenarios, the temporal aspects of each scenario, and the differentiation of blood Pb and risk estimates with regard to policy-relevant exposure pathways.

Exposure Assessment (Chapter 3):

1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis, including characterization of Pb concentrations in media, the modeling of multi-pathway Pb exposure and application of biokinetic blood Pb models, as presented in Chapter 3 technically sound, appropriately balanced, and clearly communicated?

2. Are the methods used to conduct the exposure analysis, including the modeling of population-level distributions of total blood Pb levels and the pathway-apportionment of those blood Pb levels (*e.g.*, air-inhalation, versus soil-ingestion versus dust-ingestion, versus background) technically sound? Specifically, regarding the indoor dust ingestion pathway, is the apportionment of exposure between recent air-related and other sources (*e.g.*, paint, outdoor soil/dust and additional sources), appropriate?

3. Regarding the hybrid indoor dust model developed for the general urban case study, what are the Panel's views on (a) the mechanistic component used to link outdoor ambient air Pb with indoor dust Pb concentrations and (b) the empirically-based component used to represent other sources of indoor dust Pb (*e.g.*, paint, outdoor soil/dust and additional sources)? Does

the performance evaluation of this model (along with the results of that evaluation) support application of this model in the Pb NAAQS risk assessment?

4. Specifically regarding blood Pb modeling, what are the Panel's views regarding the decision, based on performance evaluation results, to use the IEUBK model as the primary method for predicting blood Pb levels in the analysis and use the Leggett model in the sensitivity analysis?

5. What are the Panel's views on the staff interpretation of the various performance evaluations completed for the analysis (described in Section 4.3.3) with regard to the representativeness of individual modeling steps completed for the analysis (*e.g.*, characterization of ambient air Pb levels, outdoor soil and indoor dust Pb levels and the estimation of blood Pb levels for specific case studies)?

Risk Assessment (Chapter 4):

1. In general, are the concentration-response functions and blood Pb metrics used in the analysis appropriate for this review? Given the uncertainty associated with the shape of the concentration-response relationship at low blood levels (*i.e.*, < 5 µg/dL), does the suite of concentration-response functions included in the analysis adequately span the range of likely relationships?

2. Are the methods used to conduct the health risk assessment, including the method used to combine the concentration-response functions with predicted blood Pb levels in order to generate IQ loss estimates for specific population percentiles, technically sound?

3. In recognition of uncertainty in various key modeling elements, the analysis employs multiple modeling approaches, representing combinations of different models and input parameters, for each case study.

    a. Does the resultant range of risk estimates for each case study provide a reasonable indication of the uncertainty in the modeling elements?

    b. Is this set of modeling approaches biased, either towards low- or high-risk, or does it provide a reasonably balanced treatment of uncertainty?

    c. That is, does the range of risk results for a given population percentile provide an appropriate perspective on the range of potential risk, given key sources of uncertainty in the analysis?

    d. For each of the case studies, would the Panel place significantly greater confidence in results associated with any one (or more) of the modeling approaches, relative to the others?

4. To what extent does the uncertainty characterization, including the sensitivity analysis, adequately identify key sources of uncertainty and the nature of their impact on risk results?

5. Given the hypothetical nature of the general urban case study, and associated simplifications and generalities, as well as the use of multiple modeling approaches, what is the panel's view with regard to how various aspects of the results might be used to inform our understanding of risk in U.S. in U.S. urban areas?

    a. What are important considerations in identifying different modeling approach results that may be more appropriate to some urban areas, and associated populations, than others? For example, would risk results generated using the lower GSD be more informative for smaller U.S. urban areas (*e.g.*, neighborhoods), and risk results generated using the larger GSD more informative to larger urban areas?

    b. What limitations and uncertainties does the use of spatially uniform media concentrations carry into interpretations of the case study results with regard to U.S. urban areas? For example, would less uncertainty be associated with consideration of small, as compared to larger areas, where a greater spatial variability in media concentrations might be expected?

<u>Ambient Pb information and analyses (Appendix A)</u>:

To what extent are the emissions and air quality characterizations and analyses clearly communicated, appropriately characterized, and relevant to the context for the risk assessment?

Appendix D – Review Comments from Individual CASAC Lead Review Panel Members

This appendix contains the preliminary and/or final written review comments of the individual members of the Clean Air Scientific Advisory Committee (CASAC) Lead Review Panel who submitted such comments electronically. The comments are included here to provide both a full perspective and a range of individual views expressed by Panel members during the review process. These comments do not represent the views of the CASAC Lead Review Panel, the CASAC, the EPA Science Advisory Board, or the EPA itself. The views of the CASAC Lead Review Panel and the CASAC as a whole are contained in the text of the report to which this appendix is attached. Panelists providing review comments are listed on the next page, and their individual comments follow.

Dr. Deborah Cory-Slechta

D. Cory-Slechta Comments

Since I am not a modeler, and most of the issues regarding the use of the models are being addressed by other panel members, I restrict my comments to one component of Chapter 4, specifically the performance analysis. With respect to model performance, this is where the rubber really meets the road. However, the current text related to the tests of performance analysis does not provide a sufficiently rigorous or quantitative description of the outcomes of these analyses. It uses descriptors such as “generally”, and adjectives such as ‘low’ which really cannot be evaluated since there is no calibration factor. This makes it difficult to ascertain the overall utility of the models in predicting human blood leads and consequent changes in IQ. Some more in depth information is included in the Appendix materials, but since the document itself is likely to be the sole source of information in many cases, a more precise description of the outcomes of these performance analyses needs to be included in the chapter.

Dr. Ellis Cowling

Dr. Ellis Cowling
North Carolina State University
September 4, 2007

Individual Comments Following the August 28-29, 2007 CASAC Lead Panel Discussions on the Lead Human Exposure and Health Risk Assessment Document

One of the most vivid general impressions from these late August 2007 discussions regarding the current review of the primary National Ambient Air Quality Standards for Lead is the continuing evidence that public-welfare-based (mainly ecological) concerns are indeed very far removed from the central concerns of the USEPA.

In the July 31, 2007 transmittal note from OAQPS for the August 28 and 29, 2007 CASAC Lead Panel meeting called to conduct a peer review of the Agency's the July 2007 2nd Draft Report titled "*Lead Human Exposure and Health Risk Assessments for Selected Case Studies*," the lead panel was reminded that: "This draft risk assessment report describes the full-scale human exposure and risk assessment" but that "a full-scale ecological risk assessment is not being performed for this review."

In other words, the public-health-based concerns of the USEPA with regard to lead pollution in our country appear to be much more important than the public-welfare-based concerns of the USEPA during the current (2007) review of the National Ambient Air Quality Standards for Lead.

As all of us are aware, the current Secondary (public-welfare based) Standard for lead, many years ago was set equal to the Primary (public-health based) Standard. During the time allocated for discussion of the Secondary standard for lead, this transmittal-note from OAQPS led to a review of the following CASAC Lead Panel consensus statements in the March 27, 2007 letter to Administrator Johnson:

"Secondary (Welfare-Based) NAAQS for Lead

> Chapter 6 of the 1st Draft Lead Staff Paper and Chapter 7 of the "Pilot Phase" Draft Lead Exposure and Risk Assessments technical support document present compelling scientific evidence that current atmospheric lead concentrations and deposition — combined with a large reservoir of historically-deposited lead in soils, sediments and surface waters — continue to cause adverse environmental effects in aquatic and/or terrestrial ecosystems, especially in the vicinity of large emission sources. These effects persist in some cases at locations where current airborne lead concentrations are below the levels of the current primary and secondary lead standards. *Thus, from an environmental perspective, there are convincing reasons to both retain lead as a regulated criteria air pollutant and to lower the level of the current secondary standard.*

> Since concentrations of historically deposited lead in soils throughout the U.S. (averaging 0.5 to 4 g/m^2 of land area) are changing only slowly — with a half-life exceeding a century — these concentrated deposits of lead are expected to remain accessible for exchange with the atmosphere and the rest of the biosphere into the foreseeable future. Fires, changes in land use, or climatic events such as regional dust storms could mobilize significant quantities of lead that would be harmful both to human health and ecosystems downwind. This potential for harm is not adequately recognized in the 1st Draft Lead Staff Paper and the Draft Lead Exposure and Risk Assessments technical support document, but is a concern that warrants careful continued monitoring in the future."

Thus, after our August 28-29 meeting, I reexamined the magnitude of the reservoir of "historically deposited lead" and found that — 0.5 to 4 grams per square meter of land area translates into 4.4-35.7 pounds of lead that is available for potential redistribution from every acre of land area in this country! This simple recalculation adds significance to the CASAC assertion last March that "Fires, changes in land use, or climatic events such as regional dust storms could mobilize significant quantities of lead that would be harmful both to human health and ecosystems downwind." Among these several potential causes of atmospheric redistribution of lead, wild fires in forest and grassland areas — which have been occurring in many parts of the U.S. with increasing frequency and intensity in recent decades — are potentially very important for both human and ecosystem health and stability.

In summary, many of us in CASAC continue to believe that:

> "The principal goal of the NAAQS review process is to answer the following policy question: 'What scientific evidence and/or scientific insights have been developed since the last review to indicate if the current Primary (public-health based) and Secondary (public-welfare based) National Ambient Air Quality Standards are satisfactory — or if new and different standards are needed to protect both public health and public welfare and the environment.'"

It seems clear that it is not just the present ambient concentrations of lead that are emitted into the air by contemporary lead emissions sources that are hazardous to the present and future health and productivity of terrestrial and aquatic ecosystems, but rather, in very large part, also, the fraction of the historically deposited lead that is redistributed. Thus, maintaining a Secondary (public-welfare based) NAAQS that is equivalent to the Primary (public-health based) NAAQS — and thus aims only to manage current air concentrations of lead by decreasing contemporary emissions of lead instead of processes and procedures that also decrease the redistribution of historically deposited — will not provide satisfactory protection of terrestrial and aquatic ecosystems from risks of poisoning by atmospherically deposited lead.

Dr. Douglas Crawford-Brown

Comments on the Lead Human Exposure and Health Risk Assessments for Selected Case Studies (Draft Report), Volume I

Doug Crawford-Brown, August, 2007

This review focuses largely on Chapter 3 of the document (Exposure Assessment), but this chapter is reviewed in the context of the overall document and its appendices. I begin with a few general observations and then move to more specific comments on Chapter 3.

General Remarks

This is a sophisticated and complex assessment that makes use of a wide range of data and models. While I will raise some questions about specific computational steps, the overall framework of assessment is sound. It certainly comports with current state-of-the-art in regulatory risk assessment, and with past EPA practice. The authors have identified the relevant exposure pathways and age groups; have developed at least approximate measures of exposure by each pathway; have converted this exposure to measures of blood lead through the use of a model that is both well accepted and appropriate for the current scenarios of exposure; and have employed the most recent data in developing the relationship between blood lead and IQ loss needed to characterize risk. They have considered both variability and uncertainty to the degree it is feasible to incorporate these into the assessment, including both spatial (geographic) and inter-subject variability where these are appropriate. My overall impression is of a well designed assessment containing all of the key components of a risk assessment, and executed using computational approaches that are defensible.

As I mentioned above, there are places in the assessment where I might have preferred a slightly different model, or a slightly different selection of data, or a different regression function, etc. These concerns, however, need to be placed in a much larger picture of the overall uncertainty in risk assessments of this kind, uncertainties that can be large. I would be hard pressed to claim that an assessment of this kind, by its nature, would be accurate to within better than a factor of 2 for any specific risk measure, even using state-of-the-art computational approaches and data. As a result, I am comfortable that the authors have at least produced a study that is as good as one that would result from any of the methodological changes I might introduce, and that our two assessments would be contained within the same error bounds. There comes a point in scientific assessments at which one must simply accept the current condition of the science, produce an assessment that is “good enough” to make decisions, and caveat the claims of the assessment with some idea of the likely uncertainty. I believe the authors are at this stage, and so I support their computational framework.

The major problem I have is in understanding the relationship between the three case studies and any regulatory decisions to be made. I realize that such decisions are part of a policy process that involves more than the risk numbers appearing in this document and the appendices, but whether the results of an assessment are relevant to and properly inform a specific policy decision

depends on the way in which that decision is framed by consideration of risk (and other factors). There certainly is no claim in the document that these case studies are to be taken as an approximation of national risks, except somewhat in the case of the General Urban scenario. So I am left assuming that the decision logic will be something like: *We will establish a new NAAQS for lead (if needed) if there are existing scenarios (i.e., geographic locations and facilities) at the high end of exposure characterized by unacceptably large risks to a significant fraction of the local population, and where those risks could be reduced by controls on lead emissions from regulated sources*. If that is the decision logic, then the current assessment certainly produces a wealth of information needed to make such a decision. What the current assessment cannot do is to provide the scientific basis for a full-scale cost benefit assessment of a proposed new NAAQS, or even of meeting the current NAAQS.

As to a more specific comment, I find the prediction of risks at the 99.5^{th} percentile completely unwarranted. To produce such an estimate, one needs an inter-subject variability distribution with stable and accurate estimates out at the extremes of the tails. I am completely unconvinced that the current assessment (indeed, any assessment that might be done using the current science) is capable of making reliable predictions out in these tails. This is despite the fact that the authors have chosen to run the sampling 50,000 times, which I agree will produce stability out in these tails. But stability is not the same thing as accuracy.

The problem is two-fold. First, almost all inter-subject variability distributions can be given parametric forms (such as the lognormal distribution used in this study) that are accurate near the center. But the deviations usually become quite large outside the 95 percent confidence bounds, producing large errors in the first few or last few percentiles of the distribution. And secondly, these errors usually result from some biochemical or physiological mechanisms that cause the distributions to be truncated on each end. *I would be hard pressed to support a policy application of this assessment that focused on results in this extreme tail (again, the 99.5^{th} percentile; in fact, I am skeptical of any results beyond the 95^{th} percentile).*

Finally, the major remaining weakness is the inability to fully specify the exposures from "policy background" sources, including lead paint in the home, and the very rough approximations the authors have had to make in estimating the contributions of re-suspension from soil in a yard. They have done a good job of describing these limitations, and so the reader is not led to believe that there is greater accuracy here than is warranted, but the tables of risk at the end of the report suggest that the way these pathways are treated may have a significant influence on the issue of the percentage of risk attributed to policy relevant sources. I don't think there is anything the authors can do at present to improve this situation other than to characterize the uncertainty and recommend further research (which I am not proposing here — or I should say that I am not proposing that such research precede any policy decisions).

Specific Comments

1. Figure 2-1 is a good figure to include, but there must be a way to better highlight the boxes considered in this assessment. In the reproduced copy I used, I could barely distinguish between the two categories of boxes. A better approach would be to put dashed rather than solid boxes around the parts not being considered.

2. At the bottom of page 2-13, the authors suggest that low exposures may be affecting "very different biological mechanisms" than at higher exposures. I think the intent of this sentence is in part to justify the claim that the exposure-response slope is different at low and high exposures. I am convinced that the slope is indeed different, but this doesn't mean that there are mechanisms affected at low levels that are not affected at higher levels. It may simply mean that there are unrecognized sensitive and less sensitive subpopulations (with the more sensitive individuals dominating the slope at low exposures until they are all affected and the less sensitive individuals dominating the slope at higher exposures), or it may mean that there is a sensitive mechanism that is saturated at relatively low exposures (but still present at all levels of exposure).

3. On page 2-18, the authors raise the issue of a substantial amount of data being available at a site, data of many different kinds. This is the kind of situation in which methods such as Bayesian Maximum Entropy can be used to produce spatial and temporal maps of exposure with greater accuracy than those currently used in the assessment. *I am not suggesting the authors go back and do this, but that the EPA begin to explore the use of such modern methods in future assessments of this kind*.

4. I fully support the approach taken to develop ratios of maximum monthly average to annual average based on the monitoring results. While these ratios can vary significantly between locations, I think the approach taken by the authors is the only one feasible at present. I can't, however, find any discussion of the impact of the variation in this ratio across monitors on the overall uncertainty in the assessment. This needs to be examined and at least some statement made.

5. The authors have chosen to set the outdoor soil/dust concentrations equal to the current conditions, even for simulations of a reduced NAAQS. The reasoning appears to be that it takes decades for the soil concentration to re-equilibrate with the air, and even at equilibrium the air is only a partial contributor to this soil concentration. At first, I was not convinced by this argument, but I became more comfortable with it over time as I considered that the EPA goal is also to protect the population in the short-term, when the soil will not yet have re-equilibrated. Perhaps a simple graph of the expected temporal change in soil concentration at a representative home following a shutting-down of a source of lead to ambient air would be useful to make that point more clearly. As a reader, I was provided no insight into the temporal scale of the new equilibrium, other than a sentence stating that it would take a long time.

6. I support the use of the IEUBK model as the primary model rather than that of Leggett *et al.* The Leggett *et al.* model does seem to me an appropriate upper-bound estimate. However, I am less convinced of the validity of the model verification process in which the two models, and the empirical correlation, are compared against the NHANES results. To make such a comparison, one would need to ensure that the actual exposure conditions simulated in the model and underlying the NHANES data were the same. This has not been accomplished in the present study. As a result, I am not comforted by the fact that the IEUBK model produces results near the NHANES data, nor dissuaded from the Leggett et al model by its tendency to predict results higher than those in the NHANES study.

7. The use of an elasticity analysis as a basis for the sensitivity analysis is fine, but the authors suggest too strongly that this is also a form of uncertainty analysis. They have also performed an uncertainty analysis, so I recommend dropping the term "uncertainty" whenever they are discussing sensitivity. And they need at least a statement in the document that the terms for which sensitivity is assessed are statistically independent. Otherwise, their local sensitivity analysis will produce an inaccurate representation of the global sensitivity, which is really the sensitivity one wants to assess.

8. In Table 3-1, and throughout the assessment, there is the assumption that the Multiple Pathways of Exposure model can produce reliable estimates of the *relative* concentrations in soil at different locations, if not the absolute values. To convert to absolute values, the authors then scale using site-specific data. While I am supportive of this approach, it does require that one assume that a model that cannot generate absolute values can generate the correct relative values across a spatial region. The authors should simply state this and perhaps give a justification for that belief. It is quite a common assumption in environmental modeling IF one assumes that the chief source of uncertainty lies in the emissions term (so that the exposure field scales with emissions).

9. The use of a block and block group approach in establishing the spatial template is both appropriate here and comports with past EPA practice and with current state-of-the-art in estimating risks in spatially inhomogeneous fields. And the authors' use of Monte Carlo sampling to sample individuals from the blocks/block groups is both conceptually sound and properly executed.

10. I have one quibble with the use of the NATA results, discussed beginning on page 3-5. The NATA study came with some significant caveats concerning the spatial scale on which the results should be used, with a distinct suggestion in the documentation that the results might be applied at the level of counties but not at individual blocks or block groups. The reasons are stated in the NATA documentation, but generally revolve around the inability to accurately locate emission sources spatially.

11. I support the use of multiple modeling approaches and data sources as a way to understand the uncertainty in the overall assessment. The authors have done this at each of the steps of the calculations, and have carried this through to the results in Chapter 4 and in the appendices. It is a bit hard for the reader, however, to sort through all of the combinations of approaches in Chapter 4. The final tables in that chapter are quite useful in this regard, and so perhaps the authors could give the reader a sense that these tables are coming later, or move the tables forward a bit.

12. I cannot understand the discussion at the top of page 3-13. The authors point to data in which there is a negative correlation between yard soil concentration and indoor dust lead. They suggest this may be an error caused by some sort of confounding due to spotty remediation efforts. I might agree that this confounding would cause a bias towards the null, so that any positive slope would be reduced, but I am unconvinced it would explain a *negative* correlation. This is an area of the document in which the authors need to provide better support for their claims, or at least resolve this apparent discrepancy in the existing studies.

13. I do not completely agree with the reasoning on page 3-14 for using the Lanphear *et al.* model as a performance evaluation tool rather than a primary source of calculations of blood lead levels. The argument given is that the model applies only to children of 18 months. This is true, but the relevant question is whether a mechanistic model such as IEUBK is more accurate. This will depend on whether the greatest source of uncertainty is due to the development of a mechanistic model or to the use of a surrogate population (18 month old children) as representative of the entire modeled population. The document gives no reason to believe the latter is a greater source of uncertainty than the former.

14. The greatest problem with the assessment occurs when the authors are trying to estimate the percentage contribution (to exposure and IQ loss) from the different pathways, and at different percentiles of the distribution. The problem lies largely in the use of a generic blood lead GSD that is independent of pathway of exposure. I don't believe this is valid when there are both inhalation and ingestion exposures, since the ratio of blood lead to exposure differs significantly in these two pathways. There is no way to get around the approach the authors have used, but there is a need to mention this issue prominently within any uncertainty analysis, particularly the uncertainty associated with statements about the percent contribution from policy relevant sources.

15. The authors spend several paragraphs (starting on page 3-24) developing different GSDs for the three scenarios. I think this is too much resolution. The ability to distinguish between a GSD of 1.7 and 2 is simply not present in the existing data. I am not asking that the authors change their approach, using a single GSD instead, but that they at least be aware that this may be a case of false accuracy.

16. On page 3-25, the authors confront the issue of significant figures. They state that this should be no more than "two or three". Depending upon what they mean by the term "significant figure" here (they don't say), I would think that 2 or 3 significant figures is too much of a claim to accuracy in the case of estimates that are lucky to be within a factor of 2 of the actual values.

17. The chapters on model performance assessment are the weakest of this assessment. This is understandable, since we are speaking here of a suite of models and not any one model in isolation, and so there simply aren't data that can be used to reliably test the suite of models in the conditions under which the models are being applied. We will discuss this further in the CASAC meeting.

18. A very minor point (especially given my statement about significant figures), but Table 3-23 shows the median for the secondary smelter to be 63 ppm and the text near the bottom of page 3-46 says 67 ppm.

19. I agree with the results of Table 3-25, but it was very hard for me to understand how this table was generated based solely on the text on the previous page. The authors should improve this description.

Some general thoughts on the Lead Risk Assessment

Doug Crawford-Brown

8-29-07

I summarize here what I believe to be the four big issues that face the CASAC in responding to the EPA. They are quite broad issues rather than the more specific topics we have included in our individual comments. For what they are worth, here they are:

1. The first issue is the conceptual approach to be taken in going about establishing the human health risk benefits associated with reducing the NAAQS for lead (and I am not dealing with ecosystems here). Two very broad approaches have been proposed. The first is one that begins with existing air monitoring results and blood lead results (or even historical patterns in these); links these in some way; and uses this link and the IQ exposure-response functions to examine the impact of different NAAQS values on health. It is the approach championed by Sean Hays (with echoes of [Panel members] Bruce [Lanphear] and Joel [Schwartz]) and is in essence an epidemiologically-based assessment. The second is the computational approach that begins with emissions inventories, models exposure pathways, etc. *On this first issue, while I would love to see more use of epidemiologically based approaches within the EPA, and believe there is merit to comparing such an approach with the risk assessment already performed by the EPA, I agree with the EPA staff that the computational approach is the one for which there has already been significant investment and so will need to be the one used here*. It is an approach that has strong precedent, and that counts for something. Also, I am not convinced the alternative would produce a better, more scientifically sound, answer — only a different answer. Still, it is an idea whose time has come and the EPA should be urged to consider how methodologies related to it can be used in the future.

2. The second issue, given that the first is resolved, is the broad computational framework for producing this kind of assessment rooted in emissions inventories, exposure pathways, etc. By this, I don't mean the specific algorithms used, or the data used to parameterize those algorithms, but rather the logical flow of the calculations. *On this second issue, I believe the EPA staff has generally got the flow of calculations correct, including the idea of folding in inter-subject variability at some point, performing sensitivity analyses, performing model performance assessments, etc*. That doesn't mean they have the details of the algorithms or data selection correct (an issue raised next), but I am comfortable with the overall computational structure. The one point I would raise here is that the treatment of the general urban area probably significantly understates inter-subject variability because it includes no spatial inhomogeneity.

3. The third issue is whether the details of the individual computational steps within this structure are being selected and executed well. *On this third issue, it seems to me we are providing a number of recommendations for improvement*. The most evident ones relate to the specific percentiles of the inter-subject variability distribution to be examined; the way in which inter-

subject variability is incorporated (it currently is a post-processing step after the exposure assessment central tendency values are obtained); the ways in which the multi-pathway assessment is performed; the degree to which the assessment can properly apportion exposures between the various pathways; the treatment of the exposure-response relationship; etc. I personally feel that some of these issues lie at the heart of why Figure 3-2 on Page 3-52 strikes some of us as being odd, with the general urban area results clearly well above those of the secondary smelter. This surely has something to do with how past contamination of the soil is being treated in these two cases.

4. Finally, the fourth issue is how the EPA will make the leap from the current case studies to whatever kind of analysis is needed for a NAAQS. As I said in the meeting, I can't imagine the EPA would set a national rule based on a specific facility (this seems too broad of a policy tool for a one-off problem), and so I am left with the impression that the general urban area results will be more of the driver. If that is the case, the challenge facing the EPA is to develop a bit more robust understanding of the general urban areas by simulating a larger number of such areas based in some way on an understanding of the national distribution of exposures in such areas, and then to convert this in some way to what others have been calling the "body count." *While they haven't yet done this, the EPA team has experience doing such assessments and so I am confident they can perform this step once issue 3 is dealt with.*

Well, those are my four thoughts that might be used to provide structure to a response from the CASAC.

Dr. Bruce Fowler

Bruce A. Fowler

Comments on Chapter 3 of the Draft Lead Document from August 28-29, 2007 CASAC Lead Panel Meeting, Research Triangle Park, NC

General:

A review of the Draft EPA Report on lead shows it to be a solid and thorough document which considers a number of aspects of the health issues related to human exposure to lead in air. The documents does a good job of discussing the IEUBK and Leggett models for predicting blood lead and related IQ decrements associated with lead exposure in children. There is increasing consensus in the scientific community that for some individuals there may be no threshold for the neurotoxic effects of lead. Hence, loss of cognitive ability may occur in the general population at blood lead values well below the current blood level of concern of 10µg/dl. For example a recent paper by Miranda et al.(Environmental Health Perspectives 115: 1242-1247, 2007) reported that decreased end of grade (EOG) test scores in North Carolina elementary school children were correlated with blood lead values in the range found in the general U.S. population.

This effect may have broad and costly lifetime consequences for such individuals and society as a whole in terms of poor school performance, ability of compete for college admission, possible subsequent addictions, loss of lifetime income and even increased risk of criminality. The measured decrements in IQ at an early age may hence have long-term and costly consequences later in life. Additionally, it must also be noted that early short-term exposure to elevated amounts of lead in children may lead to problems later in life since subsequent blood lead values are a reflection of both recent exposures but also lead mobilization from skeletal stores which have a half –life on the order of decades for compact bone.

This lead exposure scenario from internal stores is also an important consideration in terms of lifetime consequences and the ability so exposed individuals to function in society.

In prior discussions with EPA staff involved in modeling efforts, it was recommended, on general principles, that the EPA utilize both the IEUBK and the Leggett models in order to obtain a more complete understanding of predicted blood lead values. The recent appreciation that the IEUBK model is better for predicting more chronic lead exposures while the Leggett model is better suited for modeling short term higher dose exposures provides further scientific support for utilizing both models to address the 2 lead exposure scenarios noted above.

Mr. Sean Hays

Comments on EPA's 2nd Draft Risk and Exposure Assessment in Support of the Lead NAAQS

By: Sean Hays
Date: September 18, 2007

The case study approach used by EPA helps to inform how one might go about setting an ambient air standard for a small community near a smelter. The case study approach seems to fall short of helping to inform a decision for a National Ambient Air Quality Standard (NAAQS) for lead. Ultimately, what is needed is an understanding of how the NAAQS for lead could help to lower the level of lead in U.S. children's blood, and what portions of the population will be impacted (and by how much). Of particular importance is developing an understanding of the sources of exposure responsible for the higher blood lead levels among children in the U.S. Given the multi-media exposure source nature of lead, it is more difficult to determine this as compared to one of the more classic criteria air pollutants where the only source of exposure is via air.

It is important for EPA to conduct a risk assessment which addresses the likely (or potential) changes in the nations' children's blood lead levels resulting from various options for the NAAQS. In particular, it will be important for EPA to predict the likely change in the distribution of blood leads among children with the various NAAQS being considered. Potential changes in the profile of the U.S. children's population blood lead levels for each NAAQS option include:

- The levels of lead among the most highly exposed will be substantially lowered (with the rest of the U.S. population remaining largely unchanged). This might result by lowering the NAAQS so that only the regions of the U.S. with the highest airborne lead concentrations are lowered and the remaining significant fraction of the U.S. remains the same as a result of their airborne lead levels already being under the proposed NAAQS option.
- The levels of lead among most children in the U.S. will decline.
- A combination of the two.

It is important for EPA to be transparent about how children are being exposed today by source and to estimate the likely changes in these exposures resulting from the various NAAQS options. Unfortunately, the current case study approach does not answer this question.

Geometric Standard Deviation (GSD)
The arguments about which GSD to use in the case studies highlight the uncertainty/variability associated with the blood lead estimates. The use of a GSD is arbitrary, and only becomes important as the USEPA chooses to try and quantify improvements in blood lead levels at the upper percentiles. Instead, the USEPA should rather consider performing an "incremental" rather than "absolute" assessment of changes in blood lead levels.

Incremental versus Absolute Changes in Blood Lead

The available models that estimate blood lead levels as a function of age and multi-media exposures are less reliable at estimating absolute blood lead levels. The models are more reliable at predicting incremental changes in blood lead levels associated with a delta in exposure. The USEPA should perform an incremental risk assessment by calculating the incremental changes in blood lead levels associated with changes in the NAAQS. This type of risk assessment will be far more scientifically valid.

Dr. Samuel Luoma

September 14, 2007
Comments from Samuel N. Luoma 2nd draft lead risk assessment

The agency's decision to delay addressing a secondary lead standard is a disappointment. While there are important uncertainties in the science concerning ecological implications of atmospheric lead emissions, failure to complete the risk assessment is not a constructive response. The science helps us see several points:

1. Historic lead concentrations in sediments, at their highest (e.g., from multiple sediment cores from the Great Lakes, as shown in the draft risk assessment), were high enough to exceed levels of concern defined by empirical sediment quality criteria (so-called ERMs). The probability of ecological effects from Pb in those sediments was strong. Common trends in multiple cores indicate that the source of the Pb was the atmosphere. These concentrations were reached under the old atmospheric standard; even before that standard was broadly exceeded. Thus there are ecological reasons to justify the existence of a secondary standard, and valid reasons to suggest that standard should be lower than the present one.

2. Monitoring studies show lead concentrations in sediments have now declined below the level of concern defined by the empirical sediment quality criteria. The important goal for the future should be to assure no reversal of those downward trends. There are several ways to accomplish this:

    a. By abdicating the use of science (not completing a risk assessment or a justification for a secondary standard) the agency makes itself vulnerable to application of the precautionary principle. One of our committee members has argued that given the millions of species in the world, there surely are species more sensitive than humans. Thus precaution in protection of the environment would dictate a standard that is fraction of the human health standard. In the end such a standard is indefensible scientifically on either theoretical or empirical grounds. Theoretically the human health standard is designed to protect individuals; an ecological standard is designed to assure survival of populations. The latter requires a lesser level of protection than the former, perhaps balancing the millions of species theory. Empirically, sediment and ocean data suggest Pb concentrations have declined below any known levels of concern. There does not seem to be an ecological lead crisis resulting from atmospheric inputs at this point. Thus the justification for imposition of extreme precaution seems weak. But the important point is that the abdication of science makes room for an uncertain future in which precaution of the sort described above could become a consideration (as it is in some other jurisdictions).

    b. Setting the secondary standard equal to the primary standard is an arbitrary choice, as noted in the committee's letter. But there are not enough good data, or there has not been enough analysis of the existing data, to be sure what the secondary standard should be, beyond lower than the existing standard. Setting an interim secondary

standard at the same level as the human health standard could be as justifiable as any other choice (and probably the best choice from a cost-benefit point of view) if it were accompanied by dedication of a small research fund to further study ecological impacts at different soil/sediment lead concentrations. Such a program should include development of critical load modeling schemes along with ecological studies. The caveat would be that the standard could change based upon the results. This is nearly adaptive management, and clearly a better choice than a.

c. If we have a simple, easily interpretable goal (no reversal of the downward trajectory of concentrations in the environment) and a means to judge where we are with that goal (Pb concentrations in the environment), then the most important aspect of a solution seems obvious. It is essential that the agency immediately establish a long-term environmental monitoring program (beyond the atmospheric monitoring done for human health) to watch trends in lead concentrations in environmental receptors of atmospheric lead (indicators). Forest soils, lake sediments and ocean determinations of lead could be collected in a fashion comparable to the excellent data available from the past. If trends reverse themselves, then clearly whatever standard was chosen is inadequate. If concentrations continue to decline then the standard is adequate to the goal. This assumes concentrations at present are low enough that there is little ecological risk from direct atmospheric inputs (justifying a reactive strategy rather than a proactive strategy).

The existence of a standard is important. The standard should be sufficiently low to prevent a return to conditions of the past. The exact level of the atmospheric standard is the where the greatest uncertainty lies. That exact choice is less important than the choice to follow up the risk assessment with keeping track of the implications of the choice and doing the studies to prepare a justifiable standard in the future.

Dr. Frederick J. Miller

Fred J. Miller, Ph.D.
August 27, 2007

Chapter 4

This chapter is entitled "Risk Assessment. While various analyses are presented that describe the amount of IQ loss for specific scenarios and potential NAAQS strategies, the chapter does not really represent a risk assessment on a national scale. Yet such a national assessment is required for NAAQS pollutants so that policy decisions can be used for the final selection of the indicator variable, level, statistical form, and averaging time. Including a "General Urban Case Study" does not meet this criterion, and EPA staff have basically ignored CASAC's recommendation to do so (see our March 27, 2007 letter to the EPA Administrator).

The chapter reaches "no bottom line" as to the most scientifically defensible Air Quality Scenario (among those that were examined) to base either retaining the current NAAQS or revising it. One is left with the feeling that this may be intentional in this modified NAAQS review process so that policy makers can have more latitude in selecting standards. Given the failure of EPA staff to state the most scientifically defensible air quality scenario, complete with the indicator variable, range of levels, statistical form, and averaging time, arising from their analyses, I would argue that the CASAC Pb Panel should be proactive and provide its advice on this matter in our letter to the EPA Administrator rather than waiting until the proposed rule making is announced and then having to be reactive.

For the material presented in this chapter, the text overall is well written and clearly organized. Some clarifications are needed such as why in Figure 4-1 does the two-piece linear model not have two pieces — I assume that the other segment occurs above a blood Pb level of 10 µg/dL. Also, as I note below in my comments on Appendix L, I seriously doubt the validity and usefulness of reporting values for the 99.5^{th} percentile as the extreme tails of the distribution are highly variable for the population and the 50,000 simulations are addressing precision rather than accuracy of this estimate.

The chapter just ends without a summary or statement of major conclusions. This needs to be rectified.

Appendix K

Exhibit K-1 is a table that gives all combinations of scenario type, dust model, GSD, Pb blood metric and IQ decrement models that were examined for the risk assessment. The table could be simplified and would be more informative to the reader by using a format similar to the way Figure 2-3 appears on page 2-35 of the main report. Also, this appendix would benefit from a summary section that described the major trends in results from the three case study scenarios.

I am puzzled by the following. Why have the authors departed from the way exhibits like Exhibit K-2 are laid out where whole columns having the same data value are clearly shown and then in

exhibits like Exhibit K-10 all of the numbers are repeated for every row (realizing that a couple of the rows are the only ones that are different in, for example, K-10)?

Appendix L

This appendix is well written and laid out in a way that is easy for the reader to follow and to extract the salient points about which variables have the greatest impact on the results (and thus the variables for which the analyses are the most sensitive to the choice of the value for the variable).

The only variable for which the range of values used in the sensitivity analyses appears to be too small is the absorption fraction for absolute diet, drinking water where values of 40% and 60% in comparison to the base case of 50% are used.

On line 6 of page L-7, the authors state that data and resource limitations prevented more detailed analyses such as Monte Carlo or probabilistic methods. What is meant by resource limitations? I suspect it is a time line limitation due to the court order, and if so, that should be made explicit.

Clarify what is meant by "professional judgment" on line 11 of page L-7.

While specified in the title to Exhibit L-3, the table itself does not make clear that the data entries are amount of IQ loss. I seriously question the utility of presenting 99.5^{th} and 99.9^{th} percentile values as there is undoubtedly great uncertainty surrounding these extremes of the distribution. I would advise not presenting anything greater than the 99^{th} percentile.

Appendix M

This appendix is succinct, well written, and easy to follow. I have only a few minor comments that are listed below.

To state that a full quantitative uncertainty analysis is not possible (line 4, page M-1) because the amount of uncertainty is not well-quantified begs the question. It would be reasonable to make assumptions about the potential range of uncertainty and then proceed to quantify its impact on the overall risk outcome.

There are other sources of uncertainty that are not listed on page M-1 such as activity patterns of children, polymorphisms in the population for the way Pb is handled in various organs, etc. The authors should either expand this list or else add a statement that the ones listed are in the opinion of the authors the most important ones or whatever.

Why are some of the entries under type of case study in Exhibit M-1 bolded and others are not? Either clarify or remove the bolding.

Dr. Paul Mushak

REVIEW COMMENTS: DRAFT REPORT OF LEAD HUMAN EXPOSURE AND HEALTH RISK ASSESSMENTS FOR SELECTED CASE STUDIES (Vols. I & II)

Reviewer: Paul Mushak, Ph.D.

This second OAQPS draft report, like the first Draft, represents a great deal of work and work done under serious time constraints. The authors are to be commended for their efforts on both drafts along with the additional analyses submitted 8/16/07. I have a number of comments about the draft's main elements and comments in response to Charge Qs. One important concern deals with the extent to which the full assessment in this draft incorporates the earlier recommendations of the Panel in its 3/27/07 letter.

A. Preface Footnote, p. 1

Juxtaposition of the timelines for CASAC's Pb NAAQS Panel review with EPA's revising of the Agency's NAAQS review mechanisms — away from the traditional Staff Paper to something quite different — continues to make things problematic, in my view, for a thorough peer review.

Specifically, the Preface footnote in this second draft makes it clear, *unlike earlier language from OAQPS,* that there is a reasonably good likelihood EPA will still be required to go the route of a final Staff Paper (i.e., to include OAQPS conclusions, options and recommendations) as part of its various responses to the Federal court order providing judicial oversight for this Pb NAAQS review *[Missouri Coalition for the Environment v. EPA: Case No.: 4:04CV00660 ERW; 9/14/05]*. The legal context for a "staff paper" in this footnote was not spelled out earlier (Intro, pp. 5 & 6, 12/06 First Staff Paper Draft).

Failure of the Agency to prevail in its current motion before the Court would mean, as I read the footnote, that this CASAC Pb NAAQS Panel would probably have to do future review of additional material such as conclusions, options and recommendations within a Staff Paper.

Could OAQPS or other Agency staff clarify how this concurrent legal action could or would affect how the Panel goes about the balance of its review role were the Agency not to prevail in its current motion?

B. Major Revisions or Absence Thereof in the Second Draft

1. Case Study Approach

The Panel's 3/27/07 Letter presenting the Panel's OAQPS First (Pilot) Draft review sought to draw focus away from case-specific, i.e., locale-specific, scenarios for lead ambient air lead exposures and associated child health risks indexed as IQ point losses and towards a "national" exposure population assessment strategy.

I do not believe the Panel expected that the three-case-studies approach would largely remain intact, whatever the specific language of the 3/27/07 Letter regarding the Panel's take on alternatives to case studies. Two of the three case studies were preserved largely intact in going from the pilot to full assessments in the second draft and comprise the point sources, i.e., smelter emissions. The roadway-reentrained-dust lead scenario appearing in the first draft as a Houston, TX "near roadway" dust lead case study has been expunged in favor of one broadly defining a large urban population case study (p. 2-15, Sec. 2.2.1, et seq.).

Introduction in this draft of a "general urban" population case study for urban areas of the nation with populations > 1 million serves to extend the reach of the draft to a sizeable fraction of U.S. children exposed to air lead, a fraction broadly distributed across major urban areas where most children reside, and a fraction much greater numerically than those children impacted in the two point-source (smelter) case studies.

However, the interested reader does not get the sense from this second draft that a more "national" child population approach has taken center stage. All three case studies are presented as being basically of *equivalent national relevance.* I believe the interested reader must be provided a sense of the relative magnitude of this major urban case study in terms of child numbers versus the child numbers provided for the two smelter locales.

The authors need to quantify as much as possible this sizeable fraction of exposed children and their numbers within the "general urban" case study, using risk assessment methodology specific to this illustration versus those for the other two case studies. This effort at estimating numbers of lead-exposed children and the magnitude of their health risk across major urban areas would go a long way towards adding a "national" dimension to the various full-scale risk assessments in the second draft.

The *qualitative* distinctions drawn between the nature of the general urban case study and the two point-source illustrations were reasonably expressed. However, the authors can do much more to fill *quantitative* gaps in estimations of numbers of potentially exposed children.

Lastly, there is the sense that the Agency does not assign as much reality or validity to the "general urban" case study as it does to the two smelter site illustrations. Specifically, the "Fact Sheet" accompanying the release of this full assessment 7/07 draft notes in its fifth bullet:

> "EPA evaluated three case studies in this draft risk assessment: 1) a hypothetical general urban case ..."

In what sense, exactly, is the general urban case study "hypothetical" to the Agency? First, the critical issue in all of this is really how well the case studies or any other conceptual approach to NAAQS assessment reflect an assessment of the national picture for ambient air lead exposures in past and current time? This criticality was captured in the recommendations made to the Administrator in the Panel's 3/27/07 Letter. The critical issue is not whether the two smelter/point-source cases can be precisely characterized within themselves or are more thoroughly characterizable than the general urban illustration, but whether they are as relevant to

the national picture for ambient air lead exposures of the nation's children as the general urban case study.

The authors should re-read the Panel's 3/27/07 letter to the Administrator. Nowhere in the principal text of the Letter does the Panel in its recommendations (p. 6, top) take them to be or characterize them as "hypothetical." Specifically,

> "...there are some additional considerations and analyses that the Lead Panel strongly feels could help inform a scientifically-defensible NAAQS for lead. In particular, the Panel believes that the risk assessment would be better informed with a "*population-based*" risk assessment to supplement the current case-study approach. A population-based risk assessment would typically include two key components:
>
> 1. A quantitative description of the relationship between concentrations of lead in national ambient air and distributions of resulting blood lead concentrations: and
>
> 2. A quantitative description of the relationship between blood lead concentrations and impacts on IQ.
>
> There are multiple ways in which EPA could conduct a population-based analysis, and the Panel illustrates some possibilities in Appendix D..."

The authors' methodology for estimations of impacted children for the two smelter case studies included use of census tract data spatially pegged to the emitting source. One can arguably use broader descriptors of population distributions within broader areas. If one is looking at the types of urban areas embodied in the "general urban" case study with population centers > 1 million in this second draft, one can consider distributions of individuals by, e.g., the early childhood age band, contained within urban statistical areas.

There is nothing hypothetical about reported ambient urban air monitoring data for 2003-2005 or the existence of major metropolitan statistical units with many exposed children and areas encompassing air lead monitoring results.

I suggest authors use the approach of matching locale-specific air Pb monitoring data sets with corresponding population distribution data from readily available 2000 Census Bureau enumeration or inter-census estimation data for numbers of children residing in the various U.S. Metropolitan Statistical Areas (MSAs; formerly Standard Metropolitan Statistical Areas, SMSAs) within which air lead monitoring data were gathered. These population statistical units are broken out by such strata as population size, presence of an identifiable Central city, residence within or outside a city center, etc. and are to be found in Census Bureau archived data sets. There are currently 362 OMB-designated MSAs within the U.S. There are currently over 40 U.S. MSAs where populations are > 1 million. One can either break out MSA-specific air lead-MSA pairs or average the air lead over the summed numbers of young children in all the MSAs obtained from the Census Bureau data.

We employed the MSA-based approach as part of the CDC/ATSDR 1988 Report to Congress on the nature and extent of lead poisoning in children in the United States. We specifically calculated the numbers of children in each of the U.S. MSAs (termed SMSAs at the time) as a function of population size, presence inside or outside of a central city, etc. These data were used in several ways, such as determining national prevalences of elevated Pb-Bs at various Pb-B cutpoints by combining with NHANES II data or determining lead paint as a likely source of lead exposure using housing age as a surrogate for lead paint exposure.

Enumerations were from those U.S. Census Bureau raw data from the 1980 decennial census that covered the population within each of the specified MSAs. Children selected were those in the age band useful for combining with NHANES II data: 0.5 to 5 years of age. The data for 1980 were adjusted for the year 1984 and further steps were employed. For estimates of lead paint-impacted children, the data were further arrayed for housing ages (pre-1950, 1950-1969, 1970-1980) and family income (under $6,000, $6,000-14,999, 15,000 or more).

As an illustration, we calculated that, for the large urban SMSA designated New York, NY-New Jersey, 571,000 children 0.5 to 5 years of age lived in the Central City, 132,500 children lived outside of the Central City for a total of 703,500 children in that MSA. Of these children, 422,800 resided in housing built before 1950.

In brief illustration, if the authors have air lead monitoring data for the New York City area, they can presumably calculate the numbers of potentially exposed children within whatever desired age band using 2000 Census Bureau population data for the relevant MSA, the New York, NY-New Jersey MSA. The numbers obtained in this way may be less precise than those obtained for the much smaller area census tract approach used for the smelter cases, but they would be much more appropriate FOR A NATIONAL POPULATION RISK ASSESSMENT than point-source tweaking to endless detail.

Refs:

U.S. CDC/ATSDR. 1988. The Nature and Extent of Lead Poisoning in Children in the United States: A Report to Congress. Atlanta GA: U.S. Department of Health and Human Services. (P. Mushak, A.F. Crocetti, principal coauthors; J. Schwartz, principal contributing coauthor; M Bolger, J Briskin, J Cohen, JM Davis, H Falk, LD Grant, S Lee, R Levin, DC Moore, RS Murphy, Staff of the Division of Vital Statistics, National Center for Health Statistics, U.S. CDC, contributing authors).

AF Crocetti, P Mushak, J Schwartz. 1990. Determination of numbers of lead-exposed U.S. children by areas of the United States: An integrated summary of a report to the U.S. Congress on childhood lead poisoning. Environ. Health Perspect. 89: 109-120.

2. Comparative Magnitude of Case-Study Risk Estimates, Second vs. First Drafts

Overall, the magnitude of the health risks indexed as IQ point losses are greater under a number of scenarios in this draft than was the case in the first draft. They are greater for a number of valid reasons. This comparison can be made for the two point-source cases, the

primary and secondary smelters. The “general urban” case study, newly added in response to CASAC previous review, does not have an earlier version in these drafts.

In addition, the newest analyses, provided by OAQPS as an August 16, 2007 document, provide a closer-in, sub-area analysis of the primary smelter main area. That sub-area is confined to a spatial reach of 1.5 km, versus 10 km employed in the main study. As expected on purely mathematical and aerometric grounds, both the modeled Pb-B levels and associated magnitudes of IQ point loss are much greater at 1.5 km than for a total reach of 10 km. The corresponding numbers of children affected, of course, are less than for the larger spatial scenario.

The full health risk characterizations in this draft include use of the dose-response relationship that includes linearization of lower exposures (lower Pb-B values) in the dose-response. Draft #1 made the cut point arbitrarily high, setting a dose-response threshold, i.e., a -0- IQ point change below ~2.5 µg/dl Pb-B. The Panel for various reasons roundly criticized this step. For example, there is no evidence to conclude that the lowest measured Pb-Bs in the 2005 pooled analysis defined a theoretical threshold to the dose-IQ response relationship. Other factors affecting the range or size of risk values included the use of those alternative NAAQS values suggested by the Panel in its 3/27/07 letter, shorter averaging time, etc.

Chapter 3 of the draft, lead exposures and modeling, provides further argument in favor of the general urban approach being relatively as valid as, or more valid than, the two smelter case studies, i.e., the general urban case is not “hypothetical” (see the plots, Figure 3.2). These plots of different modeled Pb-B estimates versus the NHANES IV Pb-Bs, as the authors state, show that the general urban study plots bracket the NHANES IV curve, with the high-end exposure results quite close to the NHANES line.

In Chapter 4, the overall IQ point loss estimates from Pb-B increases (Tables 4-2 and 4-3) are seen to be higher overall with the general urban case approach than they are for estimates from the two smelter site illustrations.

RESPONSES TO CHARGE Qs: CHAPTER 3

Q. 1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis, including characterization of Pb concentrations in media, the modeling of...technically sound, appropriately balanced, and clearly communicated?

These questions are overly broad and ambiguous. Nonetheless, the technical soundness is uneven, having gaps in places, problematic interpretations in others, etc. The clarity of this Chapter could be markedly improved. See my previous comments for more specific responses.

Q. 2. Are the methods used to conduct the exposure analysis, including the modeling of population-level distributions of total Pb-B levels and pathway-apportionment at those blood lead levels, technically sound?

The question is overly broad, ambiguous and difficult to answer. See my previous comments. Also, for the indoor dust pathway, the apportionment of exposure between recent air and other sources seems OK.

Q. 3. Regarding the hybrid dust model for the general urban case study, what are views on the approach to link outdoor ambient air Pb with indoor dust Pb, and use of the empirically-based component to represent other sources of indoor dust (paint, outdoor soil/dust, and additional sources?

The approaches appear qualitatively reasonable but there's little in the way of assessing quantitative validity.

Q. 4. Specifically regarding blood lead modeling, what are the Panel's views regarding the decision...to use the IEUBK model?

This is the only Q focused enough to elicit something resembling a focused answer. There are several issues that arise with use of the biokinetic models. First, there is the validity of biokinetic modeling results for one or more site-specific case studies versus NHANES IV data that represent the nation as a whole or any of its multiple national socioeconomic strata. I noted reasons for caution about this at multiple places in past comments, and won't repeat them here.

The general urban case study is comparatively more appropriate for matching to the NHANES IV (interpolated) data than the two smelter site cases, in that the spatial reach of the former more closely captures a "national" reach than the other site-specific cases. This is borne out somewhat by comparing the modeled exposure curves in Figure 3.2.

In general, I have no problem with use of the IEUBK model rather than the Leggett model or even the concurrent use of both, for several reasons. In principle, it did not make sense to carry two models and their multiple later uses throughout the original pilot risk assessment. Things are complex enough in the documents without leaving the reader mulling over which of two models is the better one and which of the model-specific subsequent exposure numbers and IQ point loss values are the more valid ones.

The IEUBK model has been more extensively evaluated and calibrated than the Leggett, or the Leggett-Pounds incarnation, or the O'Flaherty model. I refer the authors to the detailed 2005 assessment by the National Academy of Sciences of the IEUBK model and its use for risk assessment in such exposure scenarios as Superfund site risk assessments.

<u>Ref:</u>

NAS/NRC 2005. Superfund and Mining Megasites. Lessons from the Coeur D'Alene River Basin. Ch. 6: Human Health Risk Assessment: Lead Exposure and Uptake-Use of the IEUBK Model, pp. 223-283, Washington, DC: National Academy Press.

The NAS/NRC committee concluded (p. 270) that "Design and functioning of the IEUBK blood lead prediction model are consistent with current scientific knowledge.

I would, however, caution that the Leggett model generally is judged to have a positive biasing of ca. three-fold only by means of relative comparisons, if one sets aside for the moment the question of valid use of NHANES data to establish ground truth. A better comparison would be a complete comparison of all models using a single, empirically well-characterized human population-wide empirical data set for evaluation and calibration of all three models simultaneously. This has not been done. One generally finds in the literature comparisons of different empirical data sets vs. individual biokinetic models.

Our SAB Panel evaluating the development of an All-Ages Lead Model by EPA's NCEA provided a number of comments about the relative performance of Leggett, IEUBK and O'Flaherty. Staff can consult the detailed and summary comments provided from the review. Leggett overestimates Pb-B outputs, versus O'Flaherty and IEUBK, the overestimation arising from such factors as a more uniformly high use of uptake percentage for various media, e.g. dust, than the others, and differences in lead flow approaches. O'Flaherty is a much simpler model to use versus Leggett. On the other hand, O'Flaherty builds into its construction a more precipitous drop in lead uptake by children earlier in life than the others. That is, the collapse of elevated uptake rate in going from infancy to older childhood is steeper than the others.

Q. 5. What are the Panel's views on the staff interpretation of the various performance evaluations completed for the analysis (described in Section 4.3.3) with regard to...?

a. Modeled ambient air levels? Generally OK as to summary interpretation. See the sub-area analyses for the primary smelter case submitted 8/16/07 as well. As expected, this sub-area is closer to the stack(s) and will produce higher modeled air levels.

b. Outdoor soil and indoor dust Pb levels? I have no specific comments.

c. Estimation of blood lead levels? Generally OK. See also my previous comments on the biokinetic models.

ADDITIONAL PRE-MEETING COMMENTS FOR THE OAQPS DRAFT #2 FULL RISK ASSESSMENT FOR PB NAAQS REVIEW: COMMENTS ON APPENDICES C AND H

Reviewer: Paul Mushak, Ph.D.

I. APPENDIX C

The principal determinants of the full risk assessments (children's IQ point losses) in this draft, like those for the earlier pilot assessments, are the media-specific lead concentrations and media mass or volume intakes through the major exposure portals for impacted children. This applies as well to the overall apportioning of contributions to variability and uncertainty. By

comparison, the exposure (Pb-B simulations) and health risk modeling parts of the exercise are relatively less uncertain and are more stable and manageable.

This Appendix covers quantification of media lead levels for the general urban case study. By definition, the spatial scope of the "general urban" case study is larger than for the others and requires different approaches. The authors and the Appendix recognize and attempt to address this. As I noted in the previous comments on Draft #2, there are over 40 Metropolitan Statistical Areas in the U.S. which have populations over 1 Million and there is the question of how many are to be included.

Appendix C appropriately divides the most important media-specific lead sources and pathways into: C.1. Air, C.2. Soil, and C.3. Indoor Dust.

Clarity and Transparency

I found Appendix C difficult to read. It can be greatly simplified and clarified. Some parts appear to defy dissection and linearization of their logic by the outside interested reader. Statements are made in isolation. Computations are made with little reason given why they are made or what the assumptions are that go into their production.

For example, Exhibit C-2 presents the ratios of maximum monthly or quarterly average concentrations to annual average levels. Calculation of annual average air lead levels themselves might be viewed better for use in the steady-state IEUBK model, which requires stable air lead estimates. I assume this is also because the relative rates of dust lead loadings (eventually converted to concentrations) are more accurately integrated over annualized air lead deposition rates. However, it's not clear what the point of calculating the ratios is.

On the other hand, nowhere in the Appendix is there any discussion of how the higher Pb levels in maximum quarterly or monthly average values versus the annualized concentrations may underestimate shorter-term toxicity risk for exposed children. That would be of particular concern for the recent-air-only scenario. Shorter-term higher exposure detection by going from maximum quarterly to maximum monthly as recommended by the Panel in the 3/27/07 Letter clearly implies the need to be able to model short-term exposures. The Leggett Model would be more capable of computationally doing this than the IEUBK approach. Exhibit C-2 shows, as one would expect, that the maximum quarterly or monthly averages are considerably higher than the annualized estimates. In the 95th percentile case, the ratio is 7.6, and the mean max monthly to annual ratio is 4.0.

Other examples of data presentation with minimal context are Exhibits C-1, C-3 and C-4. It is not clear why the maximum monthly average for alternative NAAQS 0.5 $\mu g/m^3$ is used and not the maximum quarterly average in any of these tables for that air Pb level.

Appendix C needs smoothing for better transitions between topics. This is in the interest of transparency about what the authors did and why they did it.

C.1. Air

There are a number of examples of problematic transparency to the writing in this air section. For example, what exactly does the term "inhalation exposure" refer to? It is linked to values that are considerably below the static air Pb monitoring concentrations and appear to be used to take account of movements of the exposed child to spaces with different air lead values and other variables, e.g., inside vs. outside air Pb levels. Adjusted air lead concentrations are not the same as estimated inhalation exposures. Isn't there a better way to express the temporal and behavioral factors in children's contact with ambient air that require a distinction between just "ambient air" (Exhibit C-3) and time-and behavior-adjusted "ambient air" than using "inhalation exposure" (Exhibit C-4)? Inhalation exposure means something quite different to toxicologists and physiologists than the meaning employed by the authors in this appendix. I see what the authors are getting at but what's wrong with just using "adjusted" ambient air to differentiate what children are encountering from air lead monitoring data? I think I follow the computations of the ratios for the tabulated air quality scenarios in Exhibit C-4.

C.2. Soil

This section can be shortened considerably by limiting the size of Exhibit C-5 to the National Housing Survey soil lead levels and several others that have more useful data for national approximations. Numerous studies are tabulated but are of marginal relevance to the actual soil lead data used for the major urban case study. Ditto for the figure Exhibit C-6. It's not clear what the values in that figure actually tell us about use of the National Housing Survey soil lead levels chosen.

Use of the term "inhalation exposure" in its assumed context as an adjustment to ambient air from monitoring sites in Section C.1 has no adjustment parallels employed by or even accessible to the authors with the other relevant media-specific lead sources and pathways, even though there's good argument for adjustments to such media use.

For example, it is bulk soil lead levels that appear to have been measured for the National Housing Survey for Lead and many of the other studies tabulated by the authors in Exhibit C.5. Bulk sample lead concentrations are typically different, i.e., much lower, than lead levels in the more mobile and bioavailable small-particle fraction of soils. They are more mobile by attaching to children's hands and clothing, shoes, pets, etc. and more bioavailable through higher solubilization rates for smaller particle lead from the GI tract. Small-particle soil lead provides the source of mobilizable dust lead that can be reentrained through various mechanisms.

The authors selected a single mean value for each yard, i.e., used a yard-wide value, despite the availability of area-specific samplings from within a given yard in the data set used by the authors, the National Housing Survey data gathered in the late 1990s. This produces several layers of uncertainty. First, drip and entryway soil lead values principally reflect fractions of soil lead from lead paint versus air lead deposition. In the case of multi-unit residences, including high rises in densely urbanized city areas that are the focus of the "general urban" case study, the soil contribution to overall lead exposures is likely attenuated versus units with yards. However, both types of soil profiles are used in the national survey.

The National Housing Survey report noted that about 40% of the units evaluated had lead paint present and soil levels for a like fraction of housing may have lead paint input to soil. This can be taken to mean that 60% of all surveyed housing had no lead paint, at least from direct transfer from the unit with the yard. However, this general urban case study involves urban, arguably older areas of the U.S. where housing stock lead paint prevalence would be higher than this.

C.3. Indoor Dust

This section is extremely short, relying as it does on two general dust lead computational approaches described in Appendix G: the hybrid model, i.e., one giving a steady-state solution for a mechanistic model or an air-only regression model (Appendix G: G-7 to G-31) and plugging in the corresponding air Pb levels from elsewhere in the appendices. Its brevity glosses over the high level of uncertainty underpinning the dust lead estimates.

II. APPENDIX H

This Appendix is a reasonably clear and appropriate presentation of the approaches employing the various Pb-B models with references to lead inputs, biokinetic computations and various aspects of outputs. Two biokinetic models and one empirical model were presented and discussed. Pros and cons of the uses of each were presented.

Backgrounding material for the models was presented earlier in the pilot document by the authors, so it is not surprising that this draft is generally a somewhat more polished version of its predecessor in the pilot risk assessment reviewed earlier by the Panel. One improvement was a much better discussion of the innards of the biokinetic modules for the two mechanistic models. There was also a better discussion of how the authors dealt with the ranges of GSDs encountered empirically for small data sets (lower GSDs, more homogeneity) and more nationally representative population data sets (higher GSDs, less homogeneity).

The attachment to the appendix and its discussion for dealing with inhalation parameters is appropriate.

Dr. Michael Rabinowitz

Review by Michael Rabinowitz of
"Lead Human Exposure and Health Risks Assessments for Selected Case Studies
(Draft Report July 2007)"

August 20, 2007

Title
Can we change it to be "Human Lead Exposure..." instead of "Lead Human Exposure..."?

Chapter 1 Introduction
I liked sections 1.1 and 1.2. It outlines how existing data from different sites were used with the ISC models, which were then applied to these specific cases.

Page 1-4 line 13 suggest replace "reflecting" with "to mimic the well-know"
line 28 change "derived from" to "fit to "
line 29 change "levels" with "distributions"

Chapter 2 Design of Exposure and Risk Assessment
Figure 2.1 prompts me to ask: At what level of PbB does each of the endpoints kick in. From the text values for each of the 7 endpoints appear to me to be 3-5, 15, xx, <10, <10, >15, >30. My question is: Do you want to show these few numbers in this Figure. They would be the only numbers but may be useful for why you chose what you did.

Page 2-5, line 22-24. Consider omitting this paragraph. It does not help, and, in a sense, it is misleading. Lead in a child's environment causes the poisoning. SES does not cause it; lead in dust does. SES may be at most a marker for owner/renter, patterns of maternal behavior, and housekeeping practices. How does this paragraph help us in this analysis?

Page 2-6, section 2.1.3. Internal Disposition
I liked this section but suggest line 24 at the end of the sentence add "…among occupationally exposed adults." I believe that is correct and does not apply to children, or does it?

Page 2-7, line 2. You could add at the end of that paragraph, "Indeed, further strengthening this case are many studies where tooth or hair were collected, but the authors in their reporting convert these to an equivalent blood value for their readers."

Page 2-12, Section 2.1.4.5.
Line 5. Fix the number. It is 15 µg/dL for children. The threshold for lead to elevate FEP for both anemic and iron replete children is near 15 (see Sergio Piomelli's many articles, and one by Needleman and me). Maybe 30 for adults.

Page 2-13, line 28. About the shape of the curve: I share the sense of oddness to what has been suggested and apparently accepted. To me, I am uneasy, because presumably at progressively, lower levels, approaching a natural background, any lead-induced effects must be nil, so the shape of the curve there must be flat and then become very step. This greater sensitivity for lower doses reminds me too much of homeopathy. I suspect (and I could be wrong) that the non-linearity reported is an artifact of PbB being very well-measured but the other, stronger, confounders are not so well-measured, and when they are they are often categorized, before regression analysis, further degrading their apparent impact. This allows the lead variable to assume more of the variance. I am not sure if I am correct about this, but in any case, we need to take a practical approach to this. The choice of exact shape of the curve seems not to make any real difference in the subsequent analysis and conclusions (as noted on page 2-14, line 4, and elsewhere).

Page 2-14, line 26. Do you want to say "… for these IQ effects…"

Line 28. How about "…any of the neuro-toxic…"

Page 2-17, line 10. Would you consider "…are smelted and refined to produce marketable lead metal."

Line 11 Maybe a comma after "century" for a compound sentence.

Page 2-19. It may be noting the following:

> Primary lead smelters were sited to be close to the lead mines and close to a source of coal, which was needed to reduce the lead ore to lead metal. Lead metal was so valuable it could be economically shipped to the eventual markets. Hence smelters tended to be in more rural sites such as Idaho, Colorado, Utah, and SE Missouri. In contrast, secondary smelters were sited near populations, who supplied the raw material (old automobile batteries, mostly) and who were the customers for the product (new batteries).

Section 2.4. Analytical Approach. This section gave a necessary and useful and clear presentation of what is coming. I would only add on page 2-36, line12, in case I missed it elsewhere:

> Yet another source of uncertainty is difficult to estimate or calculate is that our current bio-kinetic models are based on transfer coefficient derived from a few adult males with PbBs about 18 or 20 µg/dL. How well these coefficients apply to children with lower blood leads is hard to quantify. To a first approximation they are likely OK, but they could be off by a factor of 2. I suppose Pb binding to RBCs is about the same for children or adults, but for reasons related to how body perfusion changes with age and development, how urinary clearances change with age, and how some enzyme systems may respond to changes in lead levels, the transfer rates we have available perhaps may be adequate only for approximate results.

Chapter 3 Exposure Assessment

CHARGE QUESTIONS- CHAPTER 3

1. This section was straight-forward and clear. My only comment is that I am satisfied with your choice of slope for the air/blood ratio (what had been called the beta in documents of long ago), and your rational. I am relieved you are not using the air/blood regression slope from the population's downward movement when lead was removed from gasoline circa 1975-85, because that coincided with lead's removal from food, hence confounding that ratio for that particular data base.

2. Yes, I found the conceptual approach appropriate, and the Tables 3-7 onward to present this clearly enough. For myself, looking at those tables, to analyze them further, I combined your first 2 columns (diet and water), the next 2 (outdoor detritus and indoor detritus of old origins), and the last 2 columns (current air ingested or inhaled).

Those 3 composite numbers are: Total Diet, Debris (indoor and out), and Airborne.

Then the last item (the one we are trying to regulate) can be compared with the sum of the other two (which are somewhat out of reach). I found this useful. These composite numbers do vary a lot going through the rows of the tables.

3. My views touch a few topics. Relevant to page 3-10, there exists a pool of dust lead in people's attics. For some this is minor, but others can be sizeable. A dramatic example can be found in dry, dusty places like Port Pirie and Picher OK. Above the ceilings of 1 story houses are accumulations of lead laden dust among the rafters, which required remediation, because they find ways to come down through the ceilings. Does your mechanistic model include such a pool of dust? Does it matter?

On another concern, for example on page 3-11 (and following in many examples) formula are presented with numerical values. Three significant figures are shown, but alas, no error terms or uncertainties, or ranges. We have no way of knowing if what you are telling us is any different from zero. So, using Table 3-2 as an example, please write out the hybrid coefficient not as 1.15 but 1.15 (0.23) (I made up that value; it might be: 1.15 (0.05) 05, 1.15(0.85)); or you could give us an 85% confidence range , maybe 1.10-1.20, or 1.14-1.16 or 0.8-1.9, so we have some sense of your confidence. Your readers can absorb the extra information, and I think you owe it to them. My fear is that the presentation as is makes the data look as solid as engineering data.

Looking at C-6, do your models recognize the seasonal nature of the variations in ventilation rates (opening windows…)? Also, there is a wealth of data on building penetration rates (indoor/outdoor ratios) for allergens and spores which you may have already tapped.

Exhibit C-5. You seem to have more than enough studies of urban soils, but if you need one more I offer some old work from Boston: Rabinowitz M, Leviton A, Needleman H, Bellinger D,

Waternaux C. Environmental correlates of infant blood lead levels in Boston. Environ Res. 1985;38(1):96-107

Exhibit C-7 and C-8 compare the 2 approaches. They are fairly different in results. For example, under current conditions, 122 and 24 ppm vs. 47 and 60 ppm. The orders are different and the values are off by about a factor of 2 or 3. This difference should alert us to the size of the uncertainties of just these 2 models, let alone differences between these models and what might be found.

C-7, line 25. Might you say, "…per gram of dry, sieved soil,…"?

On the graphs on Page G-23 Exhibit G-5 and G-12 and G-13, I was struck by the good agreement at air lead values below 0.4 or so, which is the area of interest. Although the Figures on page G-42 show higher values, they may not be relevant.

4. Yes, I concur that this model can be the primary method for blood lead predictions, but I do have a reservation. We really want to be able to predict "brain lead", but "blood lead" is taken as the metric, for very good reasons. But how well does blood lead serve in predicting brain accumulation of lead with brief repeated (for example, two hours, twice a day for 5 days) not chronic exposure? As lead enters the blood plasma for transport to the brain, and to other organs, it is rapidly taken up by the RBCs. These RBCs help clear the plasma, along with the kidneys. If the intake rate is very high, are the model projects adequate?

So, I would ask if the chosen model could generate a series of brain/blood ratios as exposure times are shortened to see if the model predicts what we suspect, that the ratio goes abruptly when dose times are very short.

5. In looking at the evaluations of the analysis, I was tempted to re-title Table 4-20 "Comparison of IQ losses for the highest and lowest risk models applied to three case studies." You can shorten it to "Comparison of IQ losses for the highest and lowest risk models".

Also, regarding styles of presentation, I would have the last columns be "magnitude of the Difference," not percent, which magnifies or leverages any discrepancy. So, the values for median and 99.5% would be:

4+	9
3+	5
2+	5

The values here indicate the uncertainties are such that for the General Urban Case the median discrepancy is 4-5 IQ points. This magnitude of uncertainty is comparable to the effect sizes we are examining. Hence, my unease.

Chapter 4 Risk Assessment

This Chapter was clearly written and presented. Yet, I would make a few suggestions.

Figure 4-1. Do you want to explain the 2 piece linear model looks like a straight line because your axis only goes to 10 and the hinge is near 13? You could expand the axis, or just note it, in anticipation of readers' queries.

Section 4.1, pages 4-3, 4, and 5. Can you please add the standard deviation to the slopes and cut-off values calculated for the 2 metrics for each of the 3 models? Then we can see how much the metrics differ among the 3 models.

For example, may I suggest a table? I made up the standard deviation.

THESE are only made up examples

		METRIC	
MODEL		concurrent	life time
2 piece	Hinge	11(3)	13(2)
	beta 1	.45 (.1)	.38 (.1)
	beta 2	.11(.1)	.1(.1)
log linear	Beta	.3 (.1)	.3 (.1)
	cut point	1 (.3)	1.5 (.3)
log linear & line	Beta	2.7(1.1)	3(2)
	Cut	1	1.5
	linear slope	2.7(1.4)	2.1(1.1)

This might be a helpful overview.

Overall, this is a necessary and sufficient exercise, looking at all these possibilities. But now maybe we can see how they choices of model do not effect any conclusions (or do they?)

Page 4-3. On the choice of models, using the log-linear model with a cut-point makes the most sense to me. By tying the lowest lead to a zero IQ effect, it avoids having to extrapolate below the cut-point.

Page 4-8. We could say that with the current NAAQS standard set at 1.5 we actually achieved 0.15, so maybe if we set the next standard for 0.2 we might get 0.02, which seems too good to be true.

Page 4-9. The Table 4-2 and others like it are very good displays to include. They show that as proposed standards are considered for both recent and "past air" how much, if at all, the IQ is affected. From these we can see that compared to current conditions, lowering the standard impacts IQ at the Primary smelter, but the secondary smelters and the urban case not at all, and considering "past air" in the standard (table 4-3) , lowering the air standard seems to have no effect on IQ loss projections.

The next set of Tables on page 4-12 to 4-31, Table 4-4 to 4-19 look at pathway sources of the lead that impact IQ among the 3 different sites under current conditions and under some suggested air standards looking at both lifetime and concurrent blood lead levels. I found them very useful. I collapsed them into 3 sources: combing diet and water (called food), and combining dust and soil from indoors and outdoors (called dirt), and the inhaled air (called air). In broad summary, regardless of the model or metric some conclusions may be obvious. I created this table:

Approximate Estimated Percentage via each pathway for 3 sites using 2 blood metrics : Concurrent and Lifetime

		Urban		Primary		Secondary	
Source	pathway	concurrent	lifetime	concurrent	lifetime	concurrent	lifetime
Total Diet	food+water	30	35	55	50	65	65
Detritus	dust+soil	70	65	45	45	35	35
Airborne	Inhaled lead and recent air in dust	1	1	1	1	1	1

We can see that choice of blood metric has little effect. Also, since most of the lead will come from food and water and old dust and soil, regulating current airborne lead will have a negligible effect.

There look like some instabilities, or irregularities, or surprises in Table 4-11 and 4-13 in the columns "indoor dust from recent air." For example, page 24, dust model with log-linear blood (lower table), next to last column, with decreasing IQ loss percentages from 99.9 to 99.5, 99, 95, 70, 75, and median go 1, 46, 2, 2, 26, 1 and 2. So I suspect something amiss.

Page 4-40, Table 4-20. I suggest a different Title for the table, "Comparison of predictions from the modeling approaches which yield the highest and lowest risks."

Also, in the last column, rather than look at percentages, which magnify errors, I'd just put in the magnitude of the effect. So, the spread in estimates for the general population: 4+ (median) 9 (99.5%), for the primary 3+ and 5, for the secondary 2+ and 3. In that these spreads in estimates are comparable to the effect size can see how large the uncertainty is.

Most of these suggestions are really only a matter of style, but might be helpful.

CHARGE QUESTIONS- CHAPTER 4

4. This is a key point. The degree to which the model fails to predict known data (goodness of fit), and the uncertainties in how well we know the value for the parameters in the model are surely part of the potential error in our projections. Additionally blood lead is the chosen metric, but how well it serve in predicting brain accumulation of lead with brief repeated (for example, two hours, twice a day for 5 days) not chronic exposure? As lead enters the blood plasma for

transport to the brain, and other organs, it is preferentially taken up by the RBCs. These RBCs help clear the plasma, along with the kidneys. If the intake rate is very high, are the model projects adequate? I have no idea, but there are people who I would trust [to find out.]

Appendix A Sources, Emissions, and Air Quality

I found this section clearly written and usefully organized. I only have 3 very minor points:

1. Regarding "Utility Boilers," as in Table A-1, pages A-2, would these be oil, gas or coal fired? I suspect it might be mostly coal.

2. Also, about Coke Ovens and Iron and Steel Foundries, 2 categories, aren't these ovens usually located at the iron and steel works? I suppose you could have 2 types of sources in one place.

3. Regarding glass making, don't you want the title to include the flat or plate glass, so change: "Pressed and Blown Glass and Glassware Manufacturing " to "Glass Making: Pressed, Blown, Molded, or Floated;" or "Pressed and Blown Glass, Glassware and Plate Glass Making."

Dr. Frank Speizer

Pre-meeting Comments on Lead Human Exposure and Health Risk Assessments for Selected Case Studies. Draft Report July 2007

Submitted by: Frank E. Speizer

August 21, 2007

Chapter 4: Risk Assessment

General Comment: I found the Chapter well written and concise. It follows a logical pattern of presenting the various concentration-response functions, with summary tables and plots for each of the population scenarios considered. The source of the parameter estimates are specified, although might be better described (see below). The risk estimates are presented and the uncertainty described (mostly). Missing seems to be an overall summary and conclusion; that probably needs to be here as well as part of an Executive summary at the beginning of the whole document.

Specific Comments:

Section 4.1.1.1-4.1.1.3 In each of these sections although references are given for where the slope and cut points come from, it might be useful to generate an overall table with these parameters, and a reference and comment section that documents the source. It may be present in the earlier chapters but when I looked for it I could not find it and in any case probably should be here.

Page 4.7 third bullet: Expanded set of risk estimates: Following on the previous bullet it is not clear if the policy-relevant background pathways of diet and drinking water are included here. If so needs to be specified and if not indicated again.

Page 4-33. An additional element of uncertainty not mentioned has to do with the population exposed. Children in different settings, different social class, different activity states are not likely to be exposed at random in different indoor/outdoor environments. Do not believe much can be done about it other than to mention it as another bullet leading to uncertainty.

Page 4-36, bullet on Blood Pb modeling: The issue here seems not that Leggett model is significantly over-predicting blood Pb in comparison to IEUBK but that it is measuring something else (short term rather than long term). This is clearly stated elsewhere but seems to be forgotten in this paragraph.

Other issues handled well.

Appendix K: This could use a summary. What I gather from the series of tables is to get the contribution of Pb from recent air below 1% for Urban Air Exposures one needs an Alternative NAAQS standard of 0.2µg/m^3. This is in sharp contrast to the Primary Pb Smelter Case study where it does not seem to matter from where you start or what level you pick. Ditto secondary Pb smelter case study.

Appendix L. Page L-7. Please spell out at the end of para. 1 whether the data and resource limitations that prevent a full-scale assessment at this time are time related, financial resource limited, personal capable of doing the job resource limited, or data limited. As far as I am concerned only the lack of data is an acceptable excuse.

Page L-7, para. 2, line 11-12. Spell out what the professional judgment was that resulted in specific selections. For example the next paragraph does an excellent job of telling us what was done to get alternative values for both outdoor soil/dust and the ambient air Pb concentrations explored.

Tables are quite useful in supplying background information for figures used in Chapter 4.

Appendix M: The presentation of the summary limitation contributing to uncertainty as qualitatively described in Table M1seems to have been a useful exercise. It certainly provides a thoughtful identification of the issues to be considered. The bolding of some and not others might be considered arbitrary by some and therefore I would suggest taking out the bolding.

Pre-meeting Comments

Submitted by Frank E. Speizer

Response to Charge Questions Chapter 4, 2nd Draft Risk Assessment Lead
(Questions paraphrased)

1. Do the presented concentration-response functions adequately span range? This was difficult to determine from the tables presented so I constructed an alternative way that might be considered for the final draft.

Predicted loss in IQ GSD=1.7 Dust model Air only (from tables 4.4-4.7)

Percentile affected	95	90	75	Med
Current Condition	5	4	3	2
Current NAAQS	9	7	5	4
NAAQS 0.5maxMO	5	4	3	2
NAAQS 0.2maxQua	4	4	3	2
NAAQS 0.2maxMO	4	3	2	2
NAAQS 05maxMO	4	3	2	2

Predicted loss IQ GSD=2.0 Dust Model Hybrid (Tables 4.4-4.7)

Percentile affected	95	90	75	Med
Current Condition	10	8	5	3
Current NAAQS	14	11	7	5
NAAQS 0.5maxMO	10	8	5	3
NAAQS 0.2maxQua	9	7	5	3
NAAQS 0.2maxMO	8	7	4	3
NAAQS .05maxMO	8	6	4	2

The suggestion from these data that there is a leveling off of effects as one gets to the lower levels of the new NAAQS, however, even at the lower levels there is considerable IQ loss for the upper 5-10% of the population.

2. Are methods to combine concentration-response function with predicted blood Pb to generate IQ loss technically sound? Without commenting specifically on the formulae that are presented the approach appears sound. There may be a typo in table 3.2. Note that coefficient for PbWIPE as presented is 0.921, but in the expanded version for combined hybrid equation is 0.931.

3. Methods for combining various analyses

    a. Does range of risk estimates provide reasonable indices of uncertainty?

    I have some trouble considering the range of estimates as an index of uncertainty. The range of estimates should be generated to provide an estimate of the variability of the potential exposures and responses. The uncertainty seems to me to be a different

phenomena that needs to be discussed (as it is in the text) because of the insufficiency of the data base from which the calculations are made as well as the unknown (and known) assumption that go into using any particular set of variables. The range therefore can not be used as a measure of that uncertainty.

b. Are modeling approaches biased in either direction? Probably not.

c. Same

d. Should one choose one approach over another?

YES, I suggest that only the urban population based approach provides useful information. (Look at the numbers of actual exposed in the Primary and Secondary Smelter approaches — very little information is obtained and no generalizability can be considered.

4. Are key elements of uncertainty identified and does one model do better than another? Yes, the key issues of uncertainty are well discussed, and yes the urban model is far better than the smelter models.

5. How well does the Urban case study and its assumptions allow for understanding risk?

    a. What are important statistical determinates of risk as related to population size? See table summarized above.

    b. What are limitations by use of spatial uniformity? No questions there are limitations of the data and not taking into consideration of both spatial variations as well as population variations add to uncertainty. But isn't this why the law is written with an adequate margin of safety? The uncertainty is not a statistical issue it is a policy issue that the politicians will have defend on judging how big or little to make the margin of safety.

General comment for Charge questions. The true risks have not been calculated as I read this document. Were is the section that translates the risk estimates of an 8-point decrease in IQ for 5% of the exposure population translated into number of children at risk by city, urban area, region, etc.? What are the economic cost of 5% of the population of 6-year-olds having 8 IQ points lower because of exposure and how does this 8 IQ points get partitioned between what is preventable and what is cumulated risk from past exposure? I thing the Chapter has all this data but the calculations seemingly have not been made yet? Does that come as part of the next paper that in the past would have been the staff paper? We will need to discuss if we have to make these calculation or will Staff do them.

Dr. Ian von Lindern

Post Meeting Comments of Ian von Lindern

Second Draft Lead Human Exposure and Health Risk Assessments for Selected Case Studies

September 2007

My comments are arranged in four general areas as follows:

Concerns regarding the Form, Function and Purpose of the Document – The Message to the Administrator does not convey the Uncertainties Associated with the Risk Assessment in a manner convenient to development of Policy.

Perhaps the most important function of this document is to convey the status of air lead regulation to the Administrator and those who will develop policy within the Agency. This document fails to convey the poor state of knowledge regarding lead in air, the Agency's role in allowing the knowledge base to deteriorate, the uncertainties resulting from that lack of knowledge, the likelihood that there is insufficient information available to consider deleting the NAAQS, and that large margins of safety will likely be required to assure protectiveness in the development of the revision.

The Agency has failed in the last fifteen years in one of its primary missions to characterize the extent and severity of this priority pollutant's exposures in the U.S. The deleterious effects of lead have always been underestimated and now, from a health perspective, there may be significant damage ongoing to susceptible populations. Consequently, the Agency must now reassess the significance of lead exposures without adequate data and greater uncertainty. Twenty years ago lead was among the best characterized priority pollutants in terms of exposure and response relationships. As a result, the current NAAQS was adopted with an atypically low margin of safety incorporated in the standard. Such a low margin cannot be justified by the state of knowledge today.

The current risk assessment should be designed and implemented to both address these shortcomings and provide the Administrator with information that allows the Agency to avoid the same mistake in this review. In that regard, there are both important technical analyses and candid discussions missing from the document.

With regard to these shortcomings, the discussions do not sufficiently convey the lack of information and unknowns associated with the knowledge base of lead in the air, as documented in the AQCD. This risk assessment is largely a modeling exercise utilizing numerous assumptions based on decades old data and empirical relationships derived from environmental regimes that no longer exist in the U.S. This approach is necessary and must be undertaken because the data do not exist to address this problem more appropriately. This must be

recognized, and conveyed to the Administrator that significant uncertainties are inherent in these analyses.

Appropriate safety margins should be incorporated into both the risk assessment and the NAAQS.

As a result, there is concern with this risk assessment that the message carried forward to the Administrator will not inform the policy makers of the consequences of the Agency's failure to adequately regulate lead in air over the last fifteen years, and could result in eliminating the NAAQS, or adopting a NAAQS that simply certifies current conditions.

Concerns regarding the Failure to Recognize Lessons Learned from the History of Lead Regulation and incorporating those into the presentation.

The history of lead regulation in the last thirty years shows a clear pattern of the federal regulating Agencies: (i) assembling the available scientific evidence regarding the public health risks associated with lead in the environment; (ii) conducting generally thorough, competent and balanced evaluation by staff and advisory panels; (iii) implementing rules and programmatic protections that largely fulfilled or exceeded the stated goals; and (iv) interestingly in each case, subsequently finding through human health and epidemiologic research that these safeguards were insufficient to protect public health.

Lead was always found to be more damaging than previously believed, and the policies developed to protect children's health fell short. Changes in policy, however, evolved only after statutes or Courts mandated reviews of the accumulating evidence. Consequently, these reviews resulted in compelling arguments to further reduce lead exposures and absorption in the nation's children and the cycle repeated.

In the 1980s, with regard to lead in air, the goals for the 1978 NAAQS were greatly exceeded, as noted in the AQCD discussions and this document. Most of the reduction in air was achieved through the virtual elimination of lead from gasoline, and the subsequent closure of primary and secondary smelters in the U.S. as demand declined. Curiously, the primary driver for lead reduction in gasoline 30 years ago was not public health per se, but the protection of catalytic converters in automobiles. However, when epidemiologists were able to identify the associated reduction in urban children's lead absorption, EPA was moved to sustain the unleaded gas rules and the prevalence of lead in the atmosphere decreased to well below the NAAQS. Significant reductions of lead in other media quickly followed. Lead producers curtailed or ceased operations that greatly reduced emissions and discharges to the environment, dietary contributions were substantially reduced through food chain effects, elimination of lead from food containers and consumer goods, and revised drinking water standards. Residual sources of lead in the environment were addressed through lead paint and contaminated soils cleanups in several programs.

The AQCD chronicles these achievements and provides a rationale for EPA's decision to not implement a reduced NAAQS, as recommended by the SAB in the early 1990s. Since that time,

science and society has, once again, learned that lead exposures were more deleterious than previous consensus, and further reducing exposures would better protect public health. The regulatory exercise that the EPA and CASAC are currently engaged in must evaluate the costs and benefits of the various alternative scenarios in addressing contemporary lead exposures. It is abundantly clear that this is a difficult and challenging task, largely due to lack of usable data regarding sources, transport mechanisms, and exposures throughout the nation. This lack of pertinent data can be traced to EPA's decision to not update the NAAQS in 1991 followed by suspension of monitoring and collection of characterization data.

In hindsight, it is clear that EPA would have been well-advised to have lowered the NAAQS to a protective level consistent with observed ambient concentrations. In that case the monitoring programs would, likely, have not been abandoned.

Technical Concerns

There are four major Technical Concerns in the document as presented in the August draft. Two of those concerns represent shortcomings in fundamental steps in comprehensive risk assessment. This document dwells on characterizing current typical exposures and risk characterization. The reviewer of this document is not presented with (i) a clearly defined Reasonable Maximum Exposure (RME or worst-case). The exposures presented may underestimate worst-case conditions and potential impacts by a factor of 2. (ii) an indication of the benefits of reducing exposures to levels substantially below those observed today. No substantive improvement scenarios are presented. iii) The partitioning of effects by media within both the media concentration models and the blood lead contributions may be internally inconsistent. iv) The selection of an appropriate level of protectiveness (95^{th} percentile) should be accomplished in concert with the GSD. It is most important that the components of that overall GSD that represent variance be distinguished from the portions that are due to uncertainty arising from lack of knowledge. Policymakers will need to understand the distinction as the former should be accounted for in the form of the standard, and the latter in the margin of safety adopted.

(a) The Pilot Case Studies as presented do not convey the Worst-case Scenarios possible under the various candidate NAAQS.

Worst-case conditions: Urban Case Study: The current model utilized assumes that the typical annual average concentration observed in the ambient air is approximately 33 % of the NAAQS. This is apparently based on its observed ratio to the maximum observed quarterly value. As a result, the typical child is exposed to an annual mean of 0.5 µg/m3. Also under this analysis, all children in the pilot urban area are assumed to be exposed uniformly at that concentration. Any variance in that exposure across the urban population then is accounted for in the application of the 2.1 GSD. The 2.1 GSD used throughout the risk assessment for various applications (and suggested as appropriate for use in standard development) already accounts for numerous other sources of variance, as discussed below. Simply including the geographic variance into the overall GSD results in the inherent assumption that that 0.5 µg/m3 is the worst-case annual average concentration any child will experience under the 1.5 µg/m3 NAAQS. This doesn't seem to make sense, as the obvious worst-case situation under the current NAAQS is 1.5 µg/m3

quarterly average. In order to achieve a mean annual concentration of 0.5 µg/m3, the mean value for the other 3 quarters must be (2.0-1.5)/3= 0.16 µg/dl. If true, this presumably would be due to a seasonal *meteorological* effect and it is difficult to believe that the peak quarter would be so much different than the remainder of the year.

If indeed the data from which the ratio was derived supports this extreme divergence, that data may well have been influenced by seasonal *source* behavior. Neither of these meteorological or source factors may apply across all U.S. urban areas. *As a result, the potential worst-case scenario and, ultimately, the estimated IQ decrement may be underestimated by a factor of two or more for the current NAAQS case.* This problem continues to occur in the alternate NAAQS analyses, as well, because these are apparently developed using rollback calculations based on the inherent relationships developed for the 1.5 µg/m3 case.

Primary Smelter Case Study: Similarly the maximum (or worst-case) situation developed in the primary smelter application seems to be about 0.76 µg/m3 annual average, or about 50% of the current NAAQS. This result seems to derive from a combination of the ratio between the quarterly and annual average, *and* the current geographic disposition of population relative to the smelter in Missouri. As opposed to the urban case, the worst-case for a child near this primary smelter is *not*1.5 µg/m3, as that compliance point apparently occurs where no one lives. However, the NAAQS implicitly requires that exposure, at least theoretically, be assessed at the compliance point. The ultimate result is that potentially worst-case exposures and estimated IQ decrement are possibly underestimated by a factor of, perhaps, 1.5 to 2 times.

Secondary Smelter Case: Due to the lack of real data available for the secondary facility, this analysis has evolved into nearly a pure modeling exercise based on professional judgment and informed estimates. The outcome seems to have produced exposures that are incredibly low for all media and predict unlikely low outcomes. The apparent worst-case scenario for living near a secondary smelter (that minimally complies with the current NAAQS) seems to be safer than the urban U.S. This outcome seems especially troublesome, as most point sources of this magnitude would usually be found in an urban air shed, thus (at an extreme logic) potentially implying that building a secondary smelter nearby would improve air quality. This analysis needs considerable revision, or should be abandoned in favor an alternate “small to medium level point source” case study.

(b) The document fails to investigate, analyze or assess the potential benefits of implementing a NAAQS at concentrations below current levels (i.e. a NAAQS that could significantly improve ambient air concentrations),

Best Case Scenarios: A corollary shortcoming in the message that could be conveyed to the administrator is lack of information regarding “best case” scenarios, or what benefits could accrue from adopting a NAAQS well below current ambient levels. This would likely imply that significant actions be undertaken to curtail current emissions, or develop remedies addressing the residual effects of past emissions. The difficulty of developing this analysis is the lack of data regarding (i) current air lead levels in the U.S.; and ii) the relationship between air lead levels and resultant lead concentrations in other media. Nevertheless, it seems incumbent on the Agency to inform the Administrator regarding the potential benefits that substantial *improvement*

in air lead levels would bring, and the costs associated with achieving those levels. In the very least, the Administrator should be made aware that the Agency is severely hampered in performing this mandated function due to past failures to collect the relevant information, and recommend actions to remedy that problem before the next review.

(c) The analyses do not balance the expected exposure reductions across pertinent media sources.

Apportionment of Media-specific Impacts on Dust and Blood Lead Levels: There are two key steps in the risk assessment process that will eventually be utilized quantitatively to derive a new NAAQS. Those are the estimation of dust lead levels from air lead concentrations and the apportionment of blood lead increments to various media intakes. Both of these are largely accomplished based on empirical relationships developed several years ago, and mechanistic models with limited empirical support. Again, these analyses are hampered by lack of contemporaneous real-world observations of media lead content in the U.S. environment.

Several of the underlying analyses used are empirical, and rely on regression equations to assign weights to the various compartments. For example, lead in dust is partitioned into components derived from soil, air and a background contribution. Blood lead, similarly, is partitioned into components deriving from soil, diet, dust, and air. The relative amount of total lead is partitioned into these compartments by the coefficients selected. Various calculations are then performed to determine the relative amounts attributed to these compartments. However, a problem that often arises in these types of analyses is that unknown amounts are, by default, assigned to the background component (or intercept coefficients in regression equations). Great care must be taken to assure that effects, actually related to sources of interest, are not attributed to background (or intercept values). This erroneous assignment to background often follows in statistically-based approaches because these unknowns are, in reality, due to the imprecise and unaccounted for nature of the underlying data.

Because little data are available, these effects cannot easily be isolated through statistical techniques or assessed by comparison to contemporaneous empirical relationships. As a result, the interpretation of the analyses should pay special attention to the plausibility of the modeled outcomes. This can be accomplished by sequentially assessing the candidate NAAQS levels in the risk analyses — beginning with the current NAAQS, through current conditions, to potential or candidate NAAQS levels. This has been accomplished to some degree in the current draft, but *only* from the perspective of the effects of air lead reductions. These analyses have been oriented toward the sources and media that will be regulated under the NAAQS regime. This focus may have the unintended consequence of diminishing the importance of these sources in favor of assigning more weight to background and “other” sources. The analyses conducted for the air lead components should be equally applied from the perspective of the other compartments. Simultaneous consideration of the changes in relative impacts across the several NAAQS exposure levels could provide insight into the plausibility of these model results and assist in assessing the impact of unknown versus background effects.

(d) The Level of Protectiveness must be consistent with the treatment of variance and uncertainty in outcome predictions.

Protectiveness and the GSD – Protectiveness refers to the percent of population can be protected by adopting a reasonable standard. In 1978, EPA selected 99.5% (and a GSD of 1.3) in order to accommodate "special high risk groups within the overall population" that was estimated to exceed 20 million children at that time. This parameter is likely the most "policy" driven selection in the process, but there is little scientific reason to advocate a lessening of the degree of protectiveness today. However, there are, at least two additional considerations that are both related to why those children are observed in the "unprotected" portion of the tail of the blood lead distribution. The first consideration relates to factors that can't practicably be addressed by the NAAQS strategy. These children may show high blood lead levels due to exposures and sources unrelated to the air lead input, or may be the result of extreme co-factor considerations (*i.e.,* pica-type behavior or physiological pre-disposition). With regard to the second consideration, these children may be in extreme situations where several factors multiply to result in excessive blood lead levels. This is the special population that EPA elected to protect with a combination of the 99.5 % /GSD=1.3 combination in 1978. These two parameters must be selected in concert in order to encompass those groups that can be effectively addressed with the control strategy. EPA faced a similar decision in the last decade in other programs and elected to pursue a 95%/GSD=1.6 selection as a risk management goal in CERCLA programs. However, this GSD generally is applied to populations surrounding particular lead sources.

It is most important to select the GSD in concert with level of protectiveness. There are, at least, two particular problems in selecting an appropriate GSD. Those are: (i) defining what variance the GSD accommodates and (ii), the GSD for a similarly exposed population likely increases at lower blood lead levels. The GSD, as used by EPA, encompasses: (i) basic individual response variance for those with similar internal exposure; (ii) additional variance associated with absorption and excretion that can be due to differences in pre-disposition such as nutritional status, etc.; (iii) variance due to differences in intake associated with behavior; and (iv) differences due exposure gradients. The GSD found in typical population surveys includes to some degree all of these (and perhaps other effects), and uses this single parameter to account for several elements of variance in blood lead response. The 1.3 value selected in 1978 was reported for the epidemic at the Bunker Hill site in Idaho and likely represents an individual response differences and lowest reasonable value. This GSD is also reflective of the dose-response relationship and the mechanisms of absorption. Exposure, pre-disposition, behavioral and absorption mechanisms in these severely exposed children were possibly saturated, masking out several potential contributors to variance, leaving individual response. Slightly higher (1.4) GSDs were noted for highly (but not so severely) exposed children in urban environments (New York City) at that time. As these high exposures and blood lead levels decreased, GSDs were noted to increase, arriving at a consensus value of 1.6 for risk assessment in the 1990s, although this value is often disputed. More recent surveys, notably NHANES, show GSDs of 2.0 and greater. The latter values are more appropriate for the risk analyses conducted in this report and in the development of the NAAQS, as these address lower blood lead levels in a national context with large variations in exposure regimes.

These large surveys likely encompass several elements of overall variance and includes those "special populations" that should be considered (in or out) in selecting the level of protectiveness. GSDs are probably higher now than in 1978 (possibly excepting the individual variance), both because lower blood lead levels can show more variance as means decrease and fluctuations are relatively more significant; and because the dose-response relationship is "steeper" at lower blood leads. This becomes important in the selection of the slope factor relating blood lead levels to air lead levels below. The selection of the slope factor must consider the shape of the dose response relationship at current blood lead levels, as opposed to typical levels in 1978. As a result, the GSD must be selected in concert with both the level of protectiveness described above and the slope factor discussed below. The combined recommendation to utilize a level of protectiveness of 95%, a GSD of 2.1, and the steeper IQ blood lead slope below 7.5 µg/dl are internally consistent and appropriate for the analyses in this risk assessment.

Charge Questions:

1. To what extent is the presentation of evidence from the health studies assessed in the Pb AQCD and the integration of information from across the various health-related research areas drawn from the Pb AQCD technically sound, appropriately balanced, and clearly communicated?

The document is well developed and has effectively drawn the pertinent information from the AQCD in the Pilot Assessments. In some cases the AQCD information has been effectively augmented with data from additional sources and, for the most part, that information has been appropriately integrated into the analyses. The entire document and risk assessment, however, is weakened by the lack of contemporaneous exposure data (discussed in the several responses to charge questions below). These weaknesses translate to uncertainty in outcomes that should be more thoroughly discussed in the final draft. Some additional analyses should be conducted to convey the significance of these limitations in the development of a protective NAAQS.

2. What are the views of the Panel on the appropriateness of staff's discussion and conclusions on key issues related to quantitative interpretation of epidemiologic study results, including, particularly, the form of a blood Pb-response function for neurocognitive effects, and the form of the associated blood Pb metric?

The discussion or the various issues regarding the most appropriate blood lead metric and the methodology employed to quantitatively estimate neurocognitive outcome are well developed and balanced, indicating the Staff has a clear and fundamental understanding of the issues and the current limitations of the studies and underlying data. This was well laid out in the document and was underscored by the public discussions at the August 28-29 Panel meeting. The most effective metric and the appropriate model forms and effective ranges of the data were identified in those discussions and should be applied in the final document. There are, however, uncertainties in estimating media concentrations, subsequent exposures and the blood lead metric that are carried through to the quantification of the health effect that should be noted more directly.

3. What are the views of the Panel on the appropriateness, technical soundness and clarity of staff's presentation of the approaches used for aspects of the assessment such as air and surface soil/dust concentrations for alternate scenarios, the temporal aspects of each scenario, and the differentiation of blood Pb and risk estimates with regard to policy-relevant exposure pathways.

The exposure estimates are the weakest point in these analyses due to the lack of monitoring data available to assess current exposures in the U.S. or to support the modeling analyses relied on in the document. This weakness was identified in the AQCD and the Staff is making the best of a poorly characterized situation. There are insufficient data to characterize active emissions and emission rates, ambient concentrations and the degree, extent and severity of ongoing redistribution of residual lead in the nation's environment. As a result, the Staff is heavily relying on modeling and decades-old empirical relationships to quantify exposures. As a result, there are considerable uncertainties inherent in the overall process, albeit unavoidable. The Agency was remiss in failing to update the NAAQS, as recommended, in 1991. As a result, the Agency stopped monitoring concurrent ambient air lead emissions, emissions and related absorption metrics. Subsequently, as the AQCD points out, previously undetected deleterious health effects do occur at lower blood lead levels, and probably at the environmental concentrations the Agency failed to monitor. Unfortunately, the Agency is not now able to effectively estimate the extent of this damage in the general population; the relationship between air lead levels, emissions and absorption; a safe air lead concentration; or the number of citizens exposed to potentially dangerous levels.

These deficiencies should be emphasized in the uncertainty discussions forwarded to the Administrator. This is an especially important message to convey because the pilot study case histories and the methodologies selected to estimate representative exposures may tend to underestimate actual exposures in the U.S. today.

Exposure Assessment (Chapter 3):

1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis, including characterization of Pb concentrations in media, the modeling of 3 multi-pathway Pb exposure and application of biokinetic blood Pb models, as presented in Chapter 3 technically sound, appropriately balanced, and clearly communicated?

The exposure analyses and interpretations in Chapter 3 are technically sound and clearly communicated. The sequencing of analyses and model applications are appropriately applied and the assumptions made are sufficiently justified and discussed in the text.

The Staff Paper provides significantly more information in this regard than was presented in the AQCD and the information seems to be a good and fair representation of the state of knowledge in the U.S. at this time. It does indicate, however, that the database is a poor representation of true emissions and ambient air lead levels in the U.S. today. This lack of appropriate data should be dealt with in a more comprehensive sensitivity analyses and uncertainty discussions.

The dust lead concentrations developed in the pilot analyses may be underestimated, as discussed above, leading to underestimation of effects throughout the chain of analyses leading to IQ decrement.

2. Are the methods used to conduct the exposure analysis, including the modeling of population-level distributions of total blood Pb levels and the pathway-apportionment of those blood Pb levels (e.g., air-inhalation, versus soil-ingestion versus dust-ingestion, versus background) technically sound? Specifically, regarding the indoor dust ingestion pathway, is the apportionment of exposure between recent air-related and other sources (e.g., paint, outdoor soil/dust and additional sources), appropriate?

The logic and sequencing of the modules that have been combined to characterize lead from source to adverse health risk represents a thorough understanding of important factors and mechanisms that influence childhood lead poisoning and the consequent effects. However, for this case it is paramount that the hypotheses be clearly stated. The inherent nature of the structure of the models and the lack of precise input data regarding the sources of pollution, the receptor populations, and the specification of parameters and co-factors require that the problem be clearly stated and the uncertainty and sensitivity analyses address that hypothesis.

The report recognizes that there are several difficulties inherent in both estimating and quantifying outcomes within each step and in utilizing those results in subsequent steps. However, the report fails to recognize the effects are often multiplicative, meaning that, although uncertainty increases, the consequences of under-estimation or over-estimation can be dissimilar, and are dependent on the status of other variables in the formulations. There is concern that effects actually due to the air-lead to dust lead transport and interactions may be being accorded to background concentrations that are assumed to be beyond the NAAQS.

3. Regarding the hybrid indoor dust model developed for the general urban case study, what are the Panel's views on (a) the mechanistic component used to link outdoor ambient air Pb with indoor dust Pb concentrations and (b) the empirically-based component used to represent other sources of indoor dust Pb (e.g., paint, outdoor soil/dust and additional sources)? Does the performance evaluation of this model (along with the results of that evaluation) support application of this model in the Pb NAAQS risk assessment?

The hybrid model is appropriately developed and has considered the various issues that must be addressed in performing these analyses. However, because there are little empirical data available to support this model, the effects of varying all the main parameters should be evaluated and presented in sensitivity analyses.

4. Specifically regarding blood Pb modeling, what are the Panel's views regarding the decision, based on performance evaluation results, to use the IEUBK model as the primary method for predicting blood Pb levels in the analysis and use the Leggett model in the sensitivity analysis?

The IEUBK model is clearly the most appropriate for these analyses. Use of the Leggett model for sensitivity analysis adds information and is useful. However, in addition to assessing model performance, a key purpose of the sensitivity analysis is to assess the effect of modification of

input parameters and component models on outcome (ultimately IQ). As a result, the effect of modification of key IEUBK parameters should be assessed with respect to the effect on IQ deficit.

Possible underestimation of background soil levels for the primary and secondary point source communities, as discussed above, and ignoring the potential impacts of other soils in these communities could result in minimally low intake estimates for soil. As a result, soil concentrations should be varied as one of the sensitivity components.

The 18% absorption estimate is an absolute minimum value and should be regarded as the lower limit in the sensitivity analysis and uncertainty discussion. The upper limit for sensitivity analyses should likely be near 40%, reflecting the likely predominance of lead oxides and small particles associated with the smelters and paint contributions.

5. What are the Panel's views on the staff interpretation of the various performance evaluations completed for the analysis (described in Section 4.3.3) with regard to the representativeness of individual modeling steps completed for the analysis (e.g., characterization of ambient air Pb levels, outdoor soil and indoor dust Pb levels and the estimation of blood Pb levels for specific case studies)?

The Staff has done a good job in evaluating model performance with respect to the individual models. However, there is a concern that several assumptions made in individual steps in the methodology tend to possibly underestimate the policy relevant background contributions to intake, ignore the potential impacts of aggregate soils in the community, and underestimate dust lead concentrations. The result of combining low or consensus estimates in all component models could results in minimally low lead intake estimates for soil and dust that translate to low blood lead predictions and health impacts in the primary and secondary smelter situations. These, in turn can result in false negative findings for children near the threshold concentrations. Additionally, the lack of air quality data may dictate that determining an appropriate standard will likely rely on modeled concentrations, emphasizing the importance of conducting appropriate sensitivity analyses.

Perhaps, most critical is the estimation of dust lead concentrations. It is interesting that despite the immense amount of work accomplished with lead dust in the last sixteen years, the same relationships developed for the 1990 analyses are employed today. These models are likely appropriate in the absence of any data, as they relate to estimating the increases in dust lead concentrations associated with increasing air and soil concentrations. These models however, at least in this reviewers experience, tend to underestimate dust lead concentrations at the much lower air and soil levels seen today. The empirical relationships were developed from databases in which the dust concentrations were overwhelmed by high air and soil lead levels associated with the point sources. Although lead associated with other sources, such as paint, was always present, it seldom showed up as significant variable in these equations. This resulted in intercept terms that failed to reflect "policy-related background sources" and under-prediction in dust lead levels as soil and air lead approached background concentrations. At the primary smelter site in Idaho (Bunker Hill Site) that underwent a similar remediation to the Missouri site, dust lead levels reached a plateau at 300-350 µg/g after completion of soil cleanup (mean soil

concentrations were less than <200 µg/g post-remediation with no significant air lead remaining after smelter demolition). Dust lead surveys conducted in socio-economically similar communities with no history of lead emissions showed house dust lead levels of about 50 µg/g in homes built after 1978 and 200 µg/g in older homes. It is unlikely, at the Bunker Hill site that mean dust lead levels will decrease to below 200-300 µg/g due to residual soil and dust contamination and the age of the housing stock. The low dust lead levels estimated for both the active primary and secondary smelters seem low to use as the baseline, but might be appropriate for the lower limit sensitivity run. The Agency should consider adjusting the baseline run to reflect higher dust concentrations and a background component indicative of older housing, and then performing sensitivity analyses at higher and lower concentrations based on a review of available data.

Risk Assessment (Chapter 4):

1. In general, are the concentration-response functions and blood Pb metrics used in the analysis appropriate for this review? Given the uncertainty associated with the shape of the concentration-response relationship at low blood levels (i.e., < 5 µg/dL), does the suite of concentration-response functions included in the analysis adequately span the range of likely relationships?

The discussions and recommendations developed subsequent to the August CASAC meeting appropriately address this issue.

2. Are the methods used to conduct the health risk assessment, including the method used to combine the concentration-response functions with predicted blood Pb levels in order to generate IQ loss estimates for specific population percentiles, technically sound?

The discussions and recommendations developed subsequent to the August CASAC meeting appropriately address this issue.

3. In recognition of uncertainty in various key modeling elements, the analysis employs multiple modeling approaches, representing combinations of different models and input parameters, for each case study.

a. Does the resultant range of risk estimates for each case study provide a reasonable indication of the uncertainty in the modeling elements?

The potential range of impacts is probably captured in these analyses. Additional discussions should be provided to the readers and policymakers to aid them in utilizing these results in the context of appropriate safety margins for the eventual standard.

b. Is this set of modeling approaches biased, either towards low- or high-risk, or does it provide a reasonably balanced treatment of uncertainty?

The estimates are likely biased low due the reasons noted in the discussion above regarding Technical Concerns.

c. That is, does the range of risk results for a given population percentile provide an appropriate perspective on the range of potential risk, given key sources of uncertainty in the analysis?

The uncertainty analyses and discussions are incomplete. Reviewers and policymakers should be better informed by discussions and examples that indicate the likelihood of children being mis-classified with respect to potential outcomes. These discussions should be oriented toward developing a protective standard and appropriate margin of safety.

d. For each of the case studies, would the Panel place significantly greater confidence in results associated with any one (or more) of the modeling approaches, relative to the others?

Please see the comments above regarding the secondary smelter analysis.

4. To what extent does the uncertainty characterization, including the sensitivity analysis, adequately identify key sources of uncertainty and the nature of their impact on risk results?

Extensive sensitivity analyses should be conducted by (as indicated above) identifying an appropriate baseline or typical model run, and then varying key components while holding others constant. However, the sensitivity runs should be accomplished using both lower and upper limit analyses. It would enhance the uncertainty analyses by framing the discussions in terms of potential false positive and false negative errors for the final risk assessment.

Determination of recontamination rates and equilibrium media concentrations for active source areas remains a major shortcoming in the overall analysis.

Perhaps, most critical is the estimation made in the entire methodology is dust lead concentration from soil and air lead concentrations. Possible underestimation of background dust lead levels and ignoring the potential impacts of other soils in the community likely results in minimally low input variables for soil contribution to dust. Additionally, the air variable that determines this route's contribution to dust has particular uncertainties. As a result, it will be important not only to assess the effect of dust on outcome health indices, but the agency should also consider additional sensitivity analyses of the soil and air selections on dust lead.

5. Given the hypothetical nature of the general urban case study, and associated simplifications and generalities, as well as the use of multiple modeling approaches, what is the panel's view with regard to how various aspects of the results might be used to inform our understanding of risk in U.S. in U.S. urban areas?

The discussions and recommendations developed subsequent to the August CASAC meeting appropriately address this issue. See the discussion above.

a. What are important considerations in identifying different modeling approach results that may be more appropriate to some urban areas, and associated populations, than others? For example, would risk results generated using the lower GSD be more informative for smaller U.S. urban areas (e.g., neighborhoods), and risk results generated using the larger GSD more informative to larger urban areas?

See the discussion above.

b. What limitations and uncertainties does the use of spatially uniform media concentrations carry into interpretations of the case study results with regard to U.S. urban areas? For example, would less uncertainty be associated with consideration of small, as compared to larger areas, where a greater spatial variability in media concentrations might be expected?

See the discussion above.

Ambient Pb information and analyses (Appendix A):

To what extent are the emissions and air quality characterizations and analyses clearly communicated, appropriately characterized, and relevant to the context for the risk assessment?

See the discussions above regarding worst and best case conditions. The current analyses rely on average emission rates that may be out of date. In several instances there are no site-specific data and there is reliance on surrogate data, which also seems in short supply. It may be that the current document does provide the best baseline values that can be developed. However, the range of potential emissions associated with the types of facilities assessed in the Pilot should be developed. The sensitivity analysis should then be conducted using some reasonable maximum estimates and similar reduced rates.

The overall process envisioned in the Pilot Assessment suggests that the final runs will be accomplished at candidate NAAQS revision concentrations. This will be a convenient method to examine prospective standard levels. However, due the lack of pertinent air quality data at critical exposure locations, these analyses ultimately rely on modeled data. Because the accuracy of model predictions is (at best) a factor of two this will present "special" uncertainty questions. The predictions drive much of the internal concentrations derived in the model will ultimately be used to determine risk. However, those same model predictions will presumably be used to determine the protective concentration for the most exposed location and then at the property line. This value will, presumably be a driver with an appropriate margin of safety as the proposed NAAQS, if adopted. As a result, it may be advisable to conduct a "last" uncertainty discussion relative to the selection of the appropriate air concentration value.

This problem is compounded by the possible need to convert from a TSP-based risk assessment analysis to a proposed PM_{10} standard. If it is proposed to develop a PM_{10} standard, the sensitivity to the conversion equation assumptions should be tested in a sensitivity analysis.

Dr. Barbara Zielinska

Barbara Zielinska

Comments on Chapter 3 of the 2nd Draft Technical Report: Lead Human Exposure and Health Risk Assessments for Selected Case Studies.

Chapter 3 of the second OAQPS draft report is generally well written and clearly presented. My overall concern regarding this report is that the general urban case study is somewhat limited in scope and can hardly represent the general urban population in the whole U.S.. It would be desirable to estimate children exposures in several representative urban areas to support the establishment of the new Lead NAAQS. The case studies of both primary and secondary lead smelter sites represent rather limited locations and thus are less relevant to the national Pb standard. Although TSP lead data are rather sparse, the general population exposure is probably better characterized by PM_{10} lead (especially close to roadways) and I believe that there are data available (beside Houston) that can support such evaluation.

To answer the specific charge questions:

Q1. To what extent are the assessment, interpretation, and presentation of the results of the exposure analysis, including characterization of Pb concentrations in media, the modeling of multi-pathway Pb exposure and application of biokinetic blood Pb models as presented in Chapter 3 technically sound, appropriately balanced, and clearly communicated?

This is a very broad question, but in my opinion, the chapter is technically sound and clearly communicated.

Q2. Are the methods used to conduct the exposure analysis, including the modeling of population-level distributions of total Pb-B levels and pathway-apportionment at those blood lead levels, technically sound?

I found this analysis presented in Tables 3-11 to 3-22 very informative, especially with respect to alternative Pb NAAQS. The analysis seems to me to be technically sound. However, I think that it would be useful to emphasize the exposures related to ambient air (i.e., recent air in indoor dust - ingestion, and recent air inhalation) as relevant to the alternative Pb standard, since the level of the standard influences the Pb blood level not only via inhalation, but also via ingestion of settled dust. The influence of the Pb standard level on the inhalation of the recent air doesn't look very impressive, but the sum of the two last columns (recent air - ingestion and inhalation) shows the real difference.

Q3. Regarding the hybrid dust model for the general urban case study, what are views on the approach to link outdoor ambient air Pb with indoor dust Pb, and use of the empirically-based component to represent other sources of indoor dust (paint, outdoor soil/dust, and additional sources?

I think that this approach of linking outdoor air Pb with indoor dust Pb and the use of the empirically-based component is justifiable. However, it is difficult to judge the uncertainties of this approach.

Q4. Specifically regarding blood lead modeling, what are the Panel's views regarding the decision based on performance evaluation results to use the IEUBK model as the primary method for predicting blood Pb levels in the analysis and use the Leggett model in the sensitivity analysis?

I agree with the use of the IEUBK model as the primary method for predicting blood Pb levels in the analysis. The IEUBK model has been more extensively evaluated and calibrated than the Leggett's model and seems to be adequate for longer exposure times. I'm not sure why the Leggett model was included in the sensitivity analysis (the chapter does not explain this), since it consistently over-predict the Pb blood levels, as shown in a comparison with measured blood levels (page 3-47).

Q5. What are the Panel's views on the staff interpretation of the various performance evaluations completed for the analysis (described in Section 4.3.3) with regard to the representatives of individual modeling steps completed for the analysis?

In general, I find the analysis of the uncertainties associated with various modeling steps very useful and the interpretation of the results reasonable.