Clean Air Scientific Advisory Committee (CASAC)

A Federal Advisory Committee to the U.S. Environmental Protection Agency

September 18, 2023

EPA-CASAC-23-003

The Honorable Michael S. Regan
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: CASAC Review of the EPA's *Integrated Science Assessment (ISA) for Lead (External Review Draft – March 2023)*

Dear Administrator Regan:

The 2021 Clean Air Scientific Advisory Committee (CASAC) Lead Review Panel, hereafter referred to as the Panel, met on April 11, 2023, June 13-14, 2023, and August 23-24, 2023, to peer review the EPA's *Integrated Science Assessment (ISA) for Lead (External Review Draft – March 2023)*, hereafter referred to as the ISA. The Chartered CASAC approved the Panel's report on August 24, 2023. The CASAC's consensus responses to the charge questions and the individual review comments from the Panel are enclosed.

Overall, the CASAC finds the ISA to be a comprehensive assessment of the available science relevant to understanding the health and welfare effects of lead (Pb). The CASAC has recommendations for strengthening and improving the document as highlighted below and detailed in the consensus responses. The CASAC believes that with these recommended changes, the document will serve as a scientifically-sound foundation for the agency's review of the National Ambient Air Quality Standards (NAAQS) for Pb.

Appendix 1 is well written, and documents Pb emission sources and transport through the media (e.g., air, water, soil, sediment, and diet) that are most relevant to environmental exposures. The CASAC notes several areas that could be improved such as describing source sector categories consistently, providing greater discussion of aviation gasoline emissions and approaches to reducing Pb emissions from aircraft engines, specifying the criteria for monitoring site elimination, and listing Pb nonattainment areas (e.g., site location, design values). Development of the companion document "Overview of Lead (Pb) Air Quality in the United States," is a good approach to supplement the ISA with more up-to-date information. However, many of the figures and tables in that document should be included in the ISA. The CASAC recommends a re-evaluation of the current monitoring criteria, expansion of non-compliance monitoring, and Pb analysis of archived particulate matter ($PM_{2.5}$ and

PM_{10}) Federal Reference Method (FRM) filters to assess population exposures below the current NAAQS.

Overall, Appendix 2, documenting exposure, toxicokinetics, and biomarkers, is remarkably comprehensive, clearly written, and well organized. It would be useful to include call-out boxes to summarize key points for the major sections, as well as a diagram that shows the relationship between particle size distribution range and exposure pathways (ingestion versus inhalation). There is overlap and connection between topics, both within this appendix and across the ISA; additional cross-references could be added for greater clarity. The CASAC recommends adding a discussion of the relevant sources and parameter estimates of childhood soil ingestion rates, highlighting the relevance of the central tendency exposure values for purposes of Pb exposure modeling. The appendix nicely integrates concepts of age-specific differences in blood Pb kinetics, and age-specific differences in exposure and physiology. The expected shifts in blood Pb and bone Pb following changes in exposure are well described, but the slow decline of blood Pb due to high bone Pb stores could be better illustrated. Regarding empirical models of the relationship between air Pb and blood Pb, the CASAC recommends including key insights from occupational studies that examine particle size-related air-blood Pb relationships and adding empirical models from occupational studies that provide some indicator or measure of particle size as a modifier of the relationship between air and blood Pb. The appendix should include additional discussion on alternative metrics such as absolute blood Pb versus delta blood Pb. The appendix should also add a summary of the influence of particle size of inhaled Pb on respiratory tract deposition, clearance, transfer into the gastrointestinal tract and uptake into blood.

The eight appendices on health effects are generally well-written and provide a comprehensive and detailed review of an enormous body of literature. The authors are to be commended for this impressive work. However, a number of areas would benefit from clarification/correction related to characterization of study limitations. Some of these include: mischaracterization of exposure timing based on tooth Pb levels; overstating concerns about the potential for past exposures to impact current blood Pb levels in adults; misidentification of confounders; overstating limitations of cross-sectional study designs and not adequately considering limitations that can still occur in longitudinal study designs; and mistakenly identifying studies as having limitations if potential interactions among components of the mixture were not included in study models. Further, the text summarizing findings often does not include sample size; this is important in interpretation of the findings.

Regarding causality determinations, some discussion of the inferences possible from different Pb biomarkers is needed. For example, the relative importance of long-term (i.e., bone Pb levels) versus short- term (i.e., blood Pb levels) exposure biomarkers is not clearly acknowledged as a source of variation in study findings that does not necessarily undermine causal inference. Some health outcomes are more consistently associated with long-term exposures, which may reflect differences in pathophysiology and critical exposure periods, not merely inconsistencies in the literature. It is not appropriate to expect consistent findings across all Pb exposure biomarkers but the ISA often implies this expectation. The relative weight in causal determinations placed on consistency of effect estimates across studies versus the precision of effect estimates varies by health outcome. A more consistent approach to balancing these two considerations is needed. The relative importance of different study designs (e.g., cross-sectional versus longitudinal) and the strength of study designs in causal determinations needs better consideration. It is unclear how health categories with multiple distinct health outcome measures are considered. Specifically, when findings vary by outcome measure, it is unclear how the different measures are prioritized for causal determination. Per the EPA weight-of-evidence approach, a very strong association with only one outcome or coherence of findings across multiple outcome measures within an outcome category can contribute to a causal determination.

However, for several health measures, it is not clear how this approach is operationalized. The CASAC suggests several possible approaches for the EPA's consideration in its consensus responses. Studies that consider Pb as part of a chemical mixture or that consider effect modification should not be excluded. The CASAC suggests that the EPA rely on consistency of evidence from human studies for causality determination if toxicologic data are absent, inconsistent, or restricted to higher doses. The ISA needs more text regarding the relative role of the drivers of the causal determination for each outcome assessed – for example, an explanation such as: "for health effect 'x,' causality determinations were based on longitudinal study findings confirmed by cross-sectional studies with contributory animal evidence." Consistent use of forest plots (or other data visualization techniques) could help demonstrate the role of different exposures, outcomes, and study findings in the causal determinations.

Regarding specific causality determinations, the consensus comments include a summary table of the CASAC's comments. Based on the available literature, including compelling evidence that cumulative Pb exposure is associated with cognitive decline in several longitudinal studies of adults, the CASAC finds the relationship of Pb with decrements in adult cognitive function to be "causal" and does not agree with the ISA's designation of this as a "likely causal" association. For cardiovascular effects, the CASAC agrees with having a single causality determination for cardiovascular effects as "causal" and that it appropriately reflects the strengths and limitations of mechanistic studies and the extensive epidemiologic literature examining the association between Pb exposure and cardiovascular outcomes. For renal effects, the CASAC concurs with the determination of "causal" and suggests inclusion of specific additional evidence (detailed in the consensus response) to strengthen support for this determination. For effects on the immune system, the CASAC recommends better justification. The CASAC also finds that better justification is needed for the downgrading of the causality determination related to sensitization and allergic response. The CASAC agrees with the causal determination for hematologic effects. The CASAC recommends that the literature for preterm birth and pre-eclampsia be more carefully reviewed, which may lead to an upgrading in the causality determinations for these outcomes. The CASAC recommends revisiting the causal determination for pregnancy and birth outcomes for possible upgrading of this outcome category once the more careful review of preterm birth and pre-eclampsia is complete. The CASAC further suggests that the EPA consider adding separate outcome categories for these outcome measures or adding a more detailed summary of how discordant causal determinations for health outcomes within this outcome category are reconciled (e.g., evidence supporting causal versus suggestive determinations). The CASAC finds that the literature supports a "likely causal" relationship between Pb and earlier age at menopause and thus recommends that the female reproductive function category be upgraded to "likely causal." For effects on other organ systems and mortality, overall, the causality determinations seem largely appropriate given the strengths and limitations described. The CASAC recommends adding cardiovascular mortality, determined to be "causal," as a subcategory of total mortality and both could possibly be included in a new appendix. For cancer, the ISA conclusion should be corrected to remove epidemiologic evidence as the basis for the causality determination and to appropriately acknowledge the strong animal evidence.

Appendix 11 provides an excellent synopsis of the available toxicity data and the approach used in the Pb ISA assessment of welfare effects; it includes a good, balanced discussion of concepts, models, and approaches. The organization of the document based on endpoints (plants/ invertebrates/vertebrates) followed by biological complexity (suborganism/organism level) is logical and appropriately addresses uncertainty with the various organizational levels. However, it makes the document difficult to read and follow; thus the CASAC suggests presenting the data in tabular form. Additional discussion/justification for the application of the literature cut-off values may be useful and warranted. Section 11.1.4 makes a very important point, regarding the difficulty of attributing observed effects in the field to Pb (or even more difficult, atmospheric Pb), which must be stressed in the later sections. Section 11.1.6 provides a

good overview summary of the importance of Toxicity Modifying Factors (TMF) in the various matrices and this reflects the current "state-of-the-science" for evaluating bioavailability and predicting toxicity. This is a major improvement over past evaluations and reflects the recognition of all the research that has been developed in the past decade and the value of EPA/Industry cooperative research agreements. There needs to be a clear discussion of environmental fate and source apportionment for Pb, which would be important and necessary to establish air quality criteria.

The discussion of terrestrial ecosystems is comprehensive and highlights several important advances since the last ISA. Available new data support previous observations and support the decision to leave the causality determinations for terrestrial ecosystems unchanged. Regarding freshwater systems, Section 11.3 provides an excellent description of the freshwater research that has been conducted since the 2013 ISA and does a good job of summarizing the new findings in the context of previous causality determinations. New research refines some key observations and nearly always supports the 2013 determinations. However, in several places, reference is made to studies that are based on "nominal" or unmeasured concentrations; it is suggested that these studies do not meet the current standards for data acceptability and should be rejected and not further considered. For neurobehavioral effects in aquatic invertebrates, the CASAC suggests that the EPA reconsider the causality determination with the potential to upgrade the determination to "causal." Regarding the marine environment, Section 11.4 provides an excellent synopsis of the available ecotoxicity information for Pb. The CASAC believes there is sufficient information to modify some of the 2013 causality determinations that were judged "inadequate" at the time, to "suggestive" or "causal" based on current data, e.g., the determination for survival, growth, and reproductive effects. The ISA does a good job of describing the environmental "effects" side of the risk characterization paradigm and the importance of considering bioavailablility to address site-specific toxicity; however, there is no discussion of the relationship between atmospheric concentration and waterborne (or soil) concentration. Some discussion of Pb speciation and environmental fate processes affecting Pb in marine waters would be beneficial in assessing the effects of atmospheric Pb in the marine environment.

The overall approach used for ISA development is clearly outlined in Appendix 12. The CASAC has some comments on improving the study selection/inclusion criteria. For section 12.4.2 (Exposure, Toxicokinetics, and Biomarkers) the CASAC is concerned about strict geographic inclusion criteria (as the CASAC has noted previously) for exposure, limited to US, Canada, western Europe, and Australia. In general, for Population, Exposure, Comparison, Outcome, and Study Design (PECOS) used in Section 12.4.3 (Health), and possibly for Population, Intervention, Comparison, Outcome, and Context (PICOC) and Level of Biological Organization, Exposure, Comparison, Endpoint, and Study Design (LECES) also, the CASAC notes that the scope of the research screening tools may have omitted some relevant studies that did not fall into the predetermined categories. Section 12.4.3.2 (Epidemiological Studies) notes that review articles are generally excluded, but they can be useful in summarizing important findings from older literature. In Section 12.4.4.(Welfare), geographic exclusion criteria are problematic for welfare exposures, possibly excluding some relevant sources such as mine tailings. This section should clarify how high toxicological study concentrations were screened for inclusion, since most are greater than 10 µg/L. Section 12.5 (Literature Search and Study Selection) notes a literature cutoff of June 2022 for the ISA. This may miss more recent literature given the expected completion dates for a final ISA and the final rule. In Table 12-5, it notes that studies with more than two pollutants are excluded from consideration owing to co-pollutant collinearities. Some of these multipollutant studies may contain useful information, and a hard rule to exclude all of them may not be appropriate.

The Executive Summary (ES) is very well written and provides a high-level overview of the ISA findings. Figure ES-1 is a diagram of the conceptual model of multimedia Pb exposure, but is somewhat

difficult to follow. The summaries of 2017 National Emissions Inventory (NEI) data should be updated to use the 2020 NEI data. The Integrated Synthesis (IS) is well written and organized, and appropriately references the appendices instead of citing individual literature. For each causal category, each section ends with a very useful comparison between evidence from the 2013 ISA and this ISA that details what evidence is new in this review. Overall, it is a very useful and effective approach to present the detailed findings of the ISA. In the discussion of blood pressure effects in Section 7.4.2.1, it states these racial differences may also reflect the history of greater exposure to Pb among non-Hispanic Black populations. This seems to contradict the consistent findings of greater effects in African Americans at similar exposure concentrations. This is a key point that needs to be clarified. Note that greater effect in one group suggests a steeper dose-response relation *at the same levels of the exposure*. One group having worse effects because they experience higher exposures is a different issue. Race or ethnicity is not a biological issue, and this should be made clear throughout the document. Additional comments on ways to improve the clarity of the IS are detailed in the consensus responses.

The CASAC appreciates the opportunity to provide advice on the ISA and looks forward to the agency's response.

Sincerely,

/s/

Dr. Elizabeth A. (Lianne) Sheppard, Chair
Clean Air Scientific Advisory Committee

Enclosures

NOTICE

The Clean Air Scientific Advisory Committee (CASAC) is a chartered federal advisory committee, operating under the Federal Advisory Committee Act (FACA; 5 U.S.C. § 10). The committee provides advice to the Administrator of the U.S. Environmental Protection Agency on the scientific and technical bases of the National Ambient Air Quality Standards. The findings and recommendations of the committee do not represent the views of the Agency, and this document does not represent information approved or disseminated by EPA. The CASAC reports are posted on the EPA website at: https://casac.epa.gov.

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee

CHAIR

Dr. Elizabeth A. (Lianne) Sheppard, Rohm and Haas Professor in Public Health Sciences, Department of Environmental & Occupational Health Sciences and Department of Biostatistics, Hans Rosling Center for Population Health, University of Washington, Seattle, WA

MEMBERS

Dr. Michelle Bell, Mary E. Pinchot Professor of Environmental Health, Yale University School of the Environment, New Haven, CT

Dr. James Boylan, Chief, Air Protection Branch, Environmental Protection Division, Georgia Department of Natural Resources, Atlanta, GA

Dr. Judith C. Chow, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and Research Professor, Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

Dr. Mark W. Frampton, Professor Emeritus of Medicine, Pulmonary and Critical Care, University of Rochester Medical Center, Rochester, NY

Dr. Christina H. Fuller, Associate Professor, School of Environmental, Civil, Agricultural and Mechanical (ECAM) Engineering, University of Georgia College of Engineering, Athens, GA

Dr. Alexandra Ponette-González, Associate Professor, Department of City & Metropolitan Planning, University of Utah, Salt Lake City, UT

SCIENCE ADVISORY BOARD STAFF

Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board Staff Office, Washington, DC

U.S. Environmental Protection Agency
Clean Air Scientific Advisory Committee
Lead Review Panel (2021)

CHAIR

Dr. Elizabeth A. (Lianne) Sheppard, Rohm and Haas Professor in Public Health Sciences, Department of Environmental & Occupational Health Sciences and Department of Biostatistics, Hans Rosling Center for Population Health, University of Washington, Seattle, WA

MEMBERS

Mr. George A. Allen, Chief Scientist, Northeast States for Coordinated Air Use Management (NESCAUM), Boston, MA

Dr. James Boylan, Chief, Air Protection Branch, Environmental Protection Division, Georgia Department of Natural Resources, Atlanta, GA

Dr. Judith C. Chow, Nazir and Mary Ansari Chair in Entrepreneurialism and Science and Research Professor, Division of Atmospheric Sciences, Desert Research Institute, Reno, NV

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

Dr. Christina H. Fuller, Associate Professor, School of Environmental, Civil, Agricultural and Mechanical (ECAM) Engineering, University of Georgia College of Engineering, Athens, GA

Dr. Philip E. Goodrum, Principal Toxicologist, GSI Environmental Inc., Fayetteville, NY

Mr. Perry Gottesfeld, Executive Director, Occupational Knowledge International, San Francisco, CA

Dr. Daven Henze, Professor, Mechanical Engineering, University of Colorado Boulder, Boulder, CO

Dr. Howard Hu, Flora L. Thornton Chair of the Department of Preventive Medicine, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA

Dr. Chris Johnson, Professor, Department of Civil and Environmental Engineering, Syracuse University, Syracuse, NY

Dr. Susan Korrick, Assistant Professor of Medicine, Department of Medicine, Brigham and Women's Hospital, Channing Division of Network Medicine, Harvard Medical School, Boston, MA

Dr. Bruce Lanphear, Professor, Children's Environmental Health, Faculty of Health Sciences, Simon Fraser University, Vancouver, BC,

Dr. Joel G. Pounds, Retired, Linden, MI

Dr. Brisa Sánchez, Dornsife Endowed Professor of Biostatistics, Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, PA

Dr. Brian Schwartz, Professor, Environmental Health and Engineering, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

Dr. William Stubblefield, Professor, Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis, OR

Dr. Kathleen Vork, Staff Toxicologist, Office of Environmental Health Hazard Assessment, Air Toxicology and Risk Assessment Section, California Environmental Protection Agency, Oakland, CA

Dr. Marc Weisskopf, Cecil K. and Philip Drinker Professor of Environmental Epidemiology and Physiology, Department of Environmental Health and Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA

SCIENCE ADVISORY BOARD STAFF

Mr. Aaron Yeow, Designated Federal Officer, U.S. Environmental Protection Agency, Science Advisory Board Staff Office, Washington, DC

Consensus Responses to Charge Questions on the EPA's
Integrated Science Assessment (ISA) for Lead (External Review Draft – March 2023)

Appendix 1 – Source to Concentration

Appendix 1 is well written, and documents lead (Pb) emission sources and transport through the media (e.g., air, water, soil, sediment, and diet) that are most relevant to environmental exposures. There are several areas, as noted in the following responses, that merit additional clarification.

The EPA should be commended for its companion document "Overview of Lead (Pb) Air Quality in the United States" (U.S.EPA, 2023a). This is a good approach to supplement the ISA with more up-to-date information. It would be helpful to include the current version of the figures (Figures 1-9) and tables (Tables 1 and 2) from this document in Appendix 1. This will allow important information to be presented and it will document the data that was available at the time the ISA was released to the public. The ISA should acknowledge that newer data and updated figures and tables may be available in the stand-alone document.

To what extent are Pb sources and emissions appropriately captured and described in Section 1.2? Are the relative importance and uncertainties in comparing contemporary vs. legacy Pb sources adequately explained?

Section 1.2 documents the major contemporary and legacy Pb sources. It would be informative to discuss uncertainties associated with emissions from each source. Source sector categories should be made consistent among the supporting documents: e.g., "Non-road Mobile and Highway Vehicles" is used in U.S. EPA (2023a) and "Mobile Aircraft" is used in U.S. EPA (2021a). Changes in emissions during COVID-19 shutdowns could be examined to evaluate emission inventory and ambient concentration accuracies.

Aviation gasoline (AvGas) accounts for ~70% of U.S. atmospheric Pb emissions. A discussion of AvGas emissions related to airplane idling, take off, cruising, landing, and taxiing, and their potential contributions to background concentrations and long-range transport, would be useful to better understand exposures near civil aviation airports. Zahran et al. (2023) found increased blood Pb levels for children residing near and downwind of Reid-Hillview County airport in San Jose, CA. The ISA should include a description of this study and compare it with Pb concentrations near the study area for 2008-2010.

In October 2022, EPA proposed an "Endangerment Finding for Lead Emissions from Aircraft Engines that Operate on Leaded Fuel" (U.S. EPA, 2022) that is scheduled to be finalized this year. Information regarding the approaches to reducing Pb emissions from aircraft engines should be discussed. Unleaded AvGas that is suitable for most general aviation (piston aircraft) is available now, and this should be acknowledged (AVWeb, 2019, 2021; SwiftFuels, 2022).

Air emissions in the 2020 National Emissions Inventory (NEI) (U.S. EPA, 2023b) should be compared with emissions reported in the Toxics Release Inventory (TRI) (U.S. EPA, 2023c). The largest contributor of industrial Pb emissions derives from battery manufacturing and recycling plants in the TRI database. Discussion of soil contamination from U.S. industrial sources can benefit from studies of

Pb battery recycling and/or manufacturing facilities. An aggregate estimate of soil suspension can be used to estimate source contributions from legacy Pb contamination.

Further elaboration is needed on potential emissions and impacts from "Other Sources" such as ammunition production and use (indoor/outdoor firing ranges and hunting), waste incineration, construction dust, use of Pb chromate and other Pb-pigmented industrial paints on roadways, bridges, and steel structures, as well as abrasive blasting of Pb paint.

As particulate matter (PM) Pb concentrations differ between days with and without fire emissions, it would be helpful to tabulate fire events (e.g., location, duration, acreage), document nearby Pb concentrations, and compare the ratio of Pb levels before, during, and after fires to evaluate their environmental impacts.

To what extent are Pb fate and transport adequately covered in Section 1.3 with appropriate detail and balance across media?

For Pb transport into and within soil columns, it may be helpful to relate soil horizons to airborne dust suspension. Pb levels at upper depths and their potential transport distances (e.g., neighborhood versus longer range) need to be clarified. In addition, soil transport with organic forms of Pb (e.g., tetraethyl or tetramethyl Pb) from leaks or spills from aviation fuel and old paint also needs to be noted.

To what extent are recent advances in the development of measurement methods and method performance adequately covered in Section 1.4?

Nationwide, there are ~164 air compliance monitors for total suspended particulates (TSP) or PM_{10} Pb. The 2006-2008 CASAC Lead Panel recommended a sizeable increase in Pb monitoring, including measures of smaller particle sizes. U.S. EPA (2023a) states that "At a minimum, there must be one source-oriented site located to measure the maximum Pb concentration in ambient air resulting from each non-airport Pb source estimated to emit 0.5 or more tons of Pb per year and from each airport estimated to emit 1.0 or more tons of Pb per year." These monitoring criteria were based on the 2002 and 2005 NEIs, resulting in low and medium emission thresholds of 0.42 and 0.64 tons/yr, respectively (75 FR 81126, December 27, 2010). The adequacy of ambient Pb monitoring criteria that took effect in 2011 (40 CFR Appendix D to Part 58) to protect vulnerable populations needs to be re-evaluated (e.g., Bui et al., 2022; Richmond-Bryant et al., 2014; Zahran et al., 2023).

The EPA should re-evaluate the number of air monitors needed to bolster research that identifies particle sizes that contribute to blood Pb levels and the lowest air quality levels associated with health effects. Since the number of monitors is decreasing, criteria for site elimination (e.g., those with design values less than 50% of the National Ambient Air Quality Standards) should be specified in the ISA. The deployment of additional research (non-compliance) monitors to assess population exposures for several size fractions (TSP, PM_{10}, and $PM_{2.5}$) should include measurements to evaluate airborne Pb exposures below the current National Ambient Air Quality Standards (NAAQS).

The ISA provides a good overview of the Federal Reference Methods (FRMs) currently approved for Pb measurement. It is well known that TSP samplers do not have size-specific cut points (~30-60 μm). Size-selective PM inlets are needed to obtain representative PM mass and Pb concentrations. More effort needs to be made to conduct collocated TSP and PM_{10} Pb measurements.

To what extent do presentation of recent concentration trends and size distribution observations in Section 1.5 adequately reflect the recent literature?

Pb nonattainment areas (e.g., site location, design values) should be listed in the ISA. Discrepancies between current nonattainment areas (U.S.EPA, 2023d) and those shown in Figure 6 of the air quality companion document (U.S.EPA, 2023a) with 2019-2021 design values >0.15 $\mu g/m^3$ should also be explained.

Although over 75% of sites in the U.S. reported decreasing trends in Pb over the last decades, this does not represent the recent high Pb concentrations at some individual sites (e.g., in Indiana and Ohio). Additional neighborhood-scale (population-orientated) monitoring (40 CFR Appendix D to Part 58) to target communities in proximity to Pb emission sources is warranted to ensure safe exposure levels. Whereas Pb monitoring sites use TSP for NAAQS compliance monitoring, or PM_{10} (for sites not expected to exceed the current NAAQS of 0.15 $\mu g/m^3$), there are many more sites monitoring PM_{10} and $PM_{2.5}$ mass for compliance with the PM NAAQS. One low-cost option recommended by the CASAC is to provide nationwide Pb exposure data for future large-scale population-based studies (e.g., at Pb levels in the 10s of ng/m^3, not for compliance use). This database would include Pb analysis of existing PM_{10} and $PM_{2.5}$ FRM filters by x-ray fluorescence (XRF, which also quantifies many other elements). Most of the Pb is expected to be in the $PM_{2.5}$ fraction, with the exception of re-entrained soil-Pb which is dominant in the coarse size fraction. If a measurement is to represent exposures at the neighborhood to urban spatial scale, $PM_{2.5}$ is a useful measurement since larger particles typically don't travel beyond 0.5 to 1 km. This should be done retroactively, since state and local air agencies retain FRM filters for at least seven years. These data, along with existing Pb data from the Chemical Speciation Network and the Interagency Monitoring of Protected Visual Environments (IMPROVE) network, would enable studies of Pb health effects at levels well below the current NAAQS using National Health and Nutrition Examination Survey (NHANES), Medicare data, and other large cohort studies. This approach can also inform Pb health endpoints that are currently not causal for future NAAQS reviews.

The topics covered in Appendix 1 were selected to provide useful context ranging from sources to concentrations of Pb relevant to understanding health and ecological effects covered in the ISA. Please identify any missing or incomplete areas of research and provide specific additional studies which would fill any gaps in this appendix. Are any research areas overemphasized and is there specific text that could be reduced in length to remedy this?

To better understand the extent to which various sources contribute to environmental Pb concentrations at or near the industrial sites, Pb-source attribution studies or modeling should be initiated. Pb isotopes are useful markers for various sources. Characterization of particle size distributions, including ultrafine particles, in ambient air and emission sources is warranted to better understand community exposures.

Appendix 2 – Exposure, Toxicokinetics, and Biomarkers

Please comment on the choice and emphasis of topics for providing useful context for the evaluation of human health effects of Pb in the ISA. Is the current organization of the appendix clear and logical?

Overall, Appendix 2 is remarkably comprehensive, clearly written, and well organized. The general structure of methods, example applications, and implications of findings works very well. The level of

detail on each element is appropriate. The sections on Pb toxicokinetics and the summary are very well written.

The EPA should consider including call-out boxes to summarize key points for major sections, as well as a diagram that shows the relationship between particle size distribution range and exposure pathways (ingestion versus inhalation).

There is overlap and connections between topics, both within Appendix 2 and across the ISA document. Additional cross-references could be added for topics such as whether sieved soil fractions may contribute to variability in study findings, the extent to which particle size fraction is considered when distinguishing between ingestion and inhalation exposure pathways.

The CASAC notes that if blood lead levels (BLLs) have declined over the past two decades in general, across all subpopulations, one would expect the absolute differences in BLLs to also decline, but not necessarily the proportional differences. The ISA should clarify whether the reduction in the race/ethnicity gap described in Section 2.1.5.4 holds true for both the absolute difference in average (or geometric mean) BLLs, as is summarized, as well as for the proportional differences (e.g., ratio of average BLLs) between subpopulations. Evaluating both the absolute difference and proportional difference in BLLs over time provides a more complete picture of trends in exposure and risk disparities among various subpopulations defined by race/ethnicity categories, or other correlates of disparate exposure experience.

The CASAC recommends adding a discussion of the relevant sources and parameter estimates of childhood soil ingestion rates, highlighting the relevance of the central tendency exposure (CTE) values for purposes of Pb exposure modeling. (Please see the individual panel member comments from Dr. Philip Goodrum and Mr. Perry Gottesfeld for additional details and recommendations.)

The following areas could be reworded and reorganized for clarity, completeness, and accuracy: the CSM diagram (Figure 2-1), the role of soil XRF measurements, references to "primary Pb sources" when referring to elevated blood Pb levels, the extrapolation uncertainty associated with dietary Pb exposure from home-grown foods, and the current action limit for Pb in spices used by New York. Section 2.1.5.4 that describes race/ethnicity disparities in exposure should be reorganized because it currently conflates racial residential segregation and other characteristics of place/neighborhood with individual-level race ethnicity.

Conflicting evidence from personal air monitors and air monitoring stations could be introduced earlier in the overview of airborne Pb exposures. Please see the individual panel member comments from Dr. Goodrum and Mr. Gottesfeld for additional details and recommendations for including references to studies that point to additional sources of Pb exposure, contributions of emissions from waste to energy incinerator plants, a new Pb standard for addressing occupational Pb exposures by military personnel, and a study conducted by the California Department of Public Health at the Exide site.

Please provide any recommendations to integrate exposure and toxicokinetics information more clearly throughout the appendix.

The appendix nicely integrates concepts of age-specific differences in blood Pb kinetics, and age-specific differences in exposure and physiology. Please see the individual panel comments from Drs. Joel Pounds and Kathleen Vork for specific recommendations on improving the accuracy of describing

the functionality and limitations of specific biokinetic models (e.g., Leggett 1993). The CASAC recommends incorporating model runs with the All-Ages Lead Model (AALM) to address uncertainty associated short-term, episodic exposures.

Please see individual panel member comments from Drs. Pounds and Vork for recommendations to improve the clarity in discussion of inhalation exposures, including the influence of particle size distribution on deposition in the respiratory tract, biological factors that influence such deposition, and empirical data on absorption of inhaled particles. In addition, they provide specific recommendations regarding the organization of the topic of ICRP's classification system for inhaled materials.

The section on Occupational Exposures should be reviewed for clarity and accuracy. Please see individual panel member comments from Dr. Vork for specific recommendations regarding how empirical data from personal air samples is reviewed, and how the inhalation transfer coefficient used in the CalEPA model can be more accurately described.

Does the appendix adequately describe air-related and non-air related pathways of Pb exposure?

Yes, the appendix adequately describes air-related and non-air related pathways of Pb exposure. Please see individual panel member comments from Dr. Goodrum for a few minor suggestions to clarify the definition of atmospheric soil Pb and the CSM figure.

Please comment on how well Section 2.3 reflects the current state of knowledge of Pb biomarkers and their interpretation as it relates to exposure and dose? Is the focus on blood Pb and bone Pb appropriate given the epidemiologic literature largely assesses exposure through these two biomarkers?

Yes, the focus on blood Pb and bone Pb is appropriate. Other biomarkers are also discussed, and the document explains their limitations well. The difference between whole blood and plasma is covered well, including in the summary.

It would be useful to briefly summarize (or refer to elsewhere in the document) the analytical advantages, disadvantages, and challenges to measuring blood and bone Pb at ever decreasing levels using the typical analytical methods.

Figure 2.8 illustrates the projected levels of blood, bone, and body Pb with a high (120 µg/d) or very high (4,020 µg/d) Pb intake. This extremely high, daily Pb intake for thirty years, seems unlikely for the 21st century, is not very relevant, and thus detracts from the flow of Appendix 2. In addition, for clarity, the CASAC suggests that the EPA identify what specific bone tissue is represented in the figure.

Is there sufficient and accurate discussion of the relationship between blood Pb and bone Pb?

The expected shifts in blood Pb and bone Pb following changes in exposure are well described (e.g., Section 2.3.5.1). Discussions of half-life are accompanied by concepts of quasi-steady state, and continued exchanges between bone and blood (i.e., "prolonged terminal elimination phase"). Figure 2.8 illustrates, in part, the slow decline of blood Pb due to high bone Pb stores. This point is not well illustrated by the current figure. This point could be made more evident by (a) zooming in the plot to focus on the ages ~45-60 years, and/or (b) quantifying the rate of decline in blood Pb between ages 50-70 (the halftime can be easily quantified by fitting the blood, bone, total body curves to a two or three-term exponential equation).

Different kinetics for blood and bone compartments over time, and implications for potentially altering the interpretation of epidemiological study results, is well described (p. 2-75). However, a brief discussion should be added of the evidence in the occupational studies by Fleming et al. (1997, 1998) of a supra linear vs linear curve in the relationship between bone Pb (tibia and calcaneus) and blood Pb. The ISA should refer to factors that can affect the slope that are discussed in Fleming et al. (1997, 1998), including variability as a function of job tenure.

Variability in rates of bone remodeling with age and pregnancy/lactation periods, as well as differences between trabecular and cortical bone kinetics are well described.

Are relationships between blood Pb and Pb in soft tissues and urine Pb adequately described?

A summary should be added of the potential impact of disposition and elimination of Pb during initial days following exposure via inhalation, such as that described in Leggett (1993). Urine Pb is marginally useful as a biomarker, and the appendix makes this point well.

Sections 2.5.1 and 2.5.2 discuss empirical models of the relationship between air Pb and blood Pb from recent and older studies. Please comment on the effectiveness of this section to accurately reflect what is known about air Pb-blood Pb relationships. Please provide recommendations on any studies that should receive less or greater emphasis.

The introduction provides a good overview of numerous sources of potential bias (both positive and negative) in the air Pb – blood Pb relationship. It would be helpful to add a table that lists the various examples with columns to show sources of and potential direction(s) of bias.

The ISA should include key insights from occupational studies that examined particle size-related air-blood Pb relationships. Empirical models from occupational studies that provide some indicator or measure of particle size as a modifier of the relationship between air and blood Pb (see studies described in Vork et al., 2023) should be added. From the Meng et al. (2014) supplemental table – $PM_{2.5}/PM_{10}$, TSP differences in sample size should be pointed out as a potential factor in statistical significance. It appears that there is a statistically significant finding from exposure to Pb in $PM_{2.5}$ in the age 6-11 cohort. A summary of this finding should be added.

The ISA should revisit the study data used to evaluate the relationship between air Pb and blood Pb to address differences in the literature regarding exposure scenarios represented, various units used to report results, and statistical methods. The CASAC notes that disparities in the cited literature may preclude the aggregation of study findings in the manner presented in the ISA. Any corrections or adjustments that were applied to conduct the evaluation should be clearly presented so that readers can interpret, reproduce, and/or verify the findings.

In the discussion of the absorption fraction in relation to concentration, mass, and surface area, common metrics are needed to compare findings. For clarity and comparison purposes, those adjustments should be made in this document. Also, the relationship between bio-accessibility and surface area of small particles vs large particles should be explained.

The Appendix should spend additional time on alternative metrics – absolute blood Pb vs delta blood Pb. Empirical models are most easily applied to support a delta blood Pb risk metric, whereas mechanistic

models are generally more applicable to absolute blood Pb. Both metrics have relevance to past CASAC assessments of air quality criteria, and both have had utility in the context of risk management of contaminated sites. Both require assumptions and reliance on multimedia Pb sources, and source contributions to blood Pb. This discussion might precede the more detailed overviews of empirical models (Section 2.5) and biokinetic models (Section 2.6), to provide additional context.

Section 2.6 Biokinetic Models of Exposure-Blood Pb Relationships

A summary should be added of the influence of particle size of inhaled Pb on respiratory tract deposition, clearance, transfer into the GI tract and uptake into blood and/or cite summaries found elsewhere as a component of the Leggett model adapted for assessment of occupational exposure scenarios (e.g., Vork et al., 2023).

The Appendix refers to the Integrated Exposure Uptake Biokinetic Model for Lead in Children (IEUBK) and the AALM model, which is appropriate. The Adult Lead Model (ALM) is introduced much later (p. 2-110). It might be helpful to at least introduce the ALM and California's Leadspread (CDTSC, 2022) earlier, as additional mechanistic modeling tools used to model (predict) relationships between environmental Pb levels and age-specific blood Pb (or changes in blood Pb in the case of Leadspread).

Appendices 3-10 – Overarching Comments about Health Effects of Pb Exposure

Appendices 3-10 review published epidemiologic and toxicologic studies relevant to the assessment of Pb's health effects including Nervous System Effects (Appendix 3), Cardiovascular Effects (Appendix 4), Renal Effects (Appendix 5), Immune System Effects (Appendix 6), Hematologic Effects (Appendix 7), Reproductive and Developmental Effects (Appendix 8), Effects on Other Systems and Mortality (Appendix 9), and Cancer (Appendix 10).

These eight appendices are generally well-written and provide a comprehensive and detailed review of an enormous body of literature. The authors are to be commended for this impressive work. However, a number of areas would benefit from clarification/correction to address the three charge questions applicable to all eight appendices.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

Characterization of study limitations:

1. Human exposure assessment and exposure biomarkers

    a. Tooth Pb levels are sometimes mis-characterized as reflecting exposures that occurred at the time of their collection (e.g., ages 6-8 years). The timing reflected by Pb in teeth depends on what part of the tooth is being analyzed. While circumpulpal measures can reflect exposures around the time of loss, more typical analyses of enamel and dentin reflect the time those parts of the tooth formed that are more prenatal and early postnatal exposures.

b. Concerns about the potential for past exposures to impact current blood Pb levels in adults are repeatedly invoked as a study limitation, particularly regarding estimation of relevant concentration-response (C-R) relationships. That is, adult blood Pb levels may reflect both recent, short-term exposures and Pb released into blood from bone that reflects past, longer-term exposures. Depending on the exposure period of interest, this is not always a limitation. For example, it is not relevant to the interpretation of cord blood Pb levels which are an estimate of fetal Pb exposure in late pregnancy; instead, the text mis-identifies maternal past Pb exposures as a potential limitation to studies that use cord blood Pb as a prenatal exposure biomarker.

2. Confounding

    a. Confounding is repeatedly invoked as a study limitation (where uncontrolled or residual confounding is possible) or strength (where presumed confounders are considered). However, the CASAC identified several instances where determinants of outcome or potential causal intermediates are mis-identified as confounders. These can be important factors to consider in analyses but, failure to include them does not introduce bias or undermine the validity of study findings. Similarly, inclusion of such factors (e.g., causal intermediates) may alter the interpretation of primary effect estimates in ways that are unrelated to confounding. Examples include adjusting child health outcomes for birth outcomes when considering Pb exposure before birth and adjusting adult psychological outcomes for psychotropic medication use when considering earlier Pb exposure.

3. Study design

    a. Cross-sectional study designs are repeatedly identified as having limitations, including uncertainty regarding the directionality of associations. This limitation is applied to studies in which bone Pb is used as a biomarker of exposure. However, because bone Pb levels reflect cumulative exposures that occurred over previous years to decades, uncertainty regarding directionality of association is unlikely to be relevant, with the possible exception of studies of bone Pb and kidney disease, for which it is theoretically possible that kidney malfunction could result in reduced excretion of Pb (and therefore, higher body Pb burden). Studies that measured bone Pb, particularly tibia Pb, should perhaps be more accurately characterized as cross-sectional for the outcome but retrospective for the dose assessment.

    b. Studies assessing Pb as part of a chemical mixture (rather than as a single exposure) are often mistakenly identified as having limitations if potential interactions among components of the mixture were not included in study models. In many cases, such studies did formal testing for interaction (e.g., using mixtures methods such as Bayesian Kernel Machine Regression) and found no evidence of interactions. In addition, if interactions exist, not including them in the model often does not create a spurious association with Pb, but rather would weaken the overall association relative to the association in the more vulnerable group (i.e. the interaction would imply differences in the effect of Pb by some other factor and the overall association would be less than that in the group with the stronger effect).

4. Effect Modification

    Studies including effect modification results are often included, however clarity is needed on: the variation in the language used to describe effect modification, because the variation in language makes it difficult for the reader to follow this topic; the rationale for the inclusion of effect modification results in the ISA; the weight given to effect modification analyses that were done *post hoc*. These issues with effect modification apply to several health appendices but are especially salient in Appendix 4 (see further details in the consensus response to charge questions on Appendix 4).

The text summarizing findings often does not include study sample size – this information is important to understanding whether imprecise estimates of association (large confidence intervals) may be related to limitations in study power versus other factors.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

All of the health appendices do an excellent job integrating mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

<u>Determination of causality</u>:

1. Discussion of the inferences possible from different Pb biomarkers is needed. For example, the relative importance of long-term (i.e., bone Pb levels) versus short-term (i.e., blood Pb levels) exposure biomarkers are not clearly acknowledged as a source of variation in study findings that does not necessarily undermine causal inference. Some health outcomes are more consistently associated with long-term exposures (e.g., hypertension is more consistently associated with bone, but not blood, Pb) which may reflect differences in pathophysiology and critical exposure periods, not inconsistencies in the literature.

2. Apropos of the above, it is not appropriate to expect consistent findings across all Pb exposure biomarkers but the ISA often implies this expectation. Instead, judgements regarding consistency of findings across studies should consider how well a given study characterized the critical exposure window/time frame of exposure for a given outcome and how well the chosen biomarker reflected that exposure. In addition, greater consideration of the implications of Pb exposure across the life course is needed to assess the true population health impact of Pb. Lastly, there is emerging evidence that new approaches to estimating cumulative Pb exposure that do not require bone Pb measurements (e.g., use of an empirically-derived DNA methylation signature) are more sensitive predictors of Pb's impact on hypertension (Park et al., 2009; Wang et al., 2022) and risk of Parkinson's disease (Paul et al., 2021). This highlights the potential for the specific properties of a Pb biomarker to be an important determinant of the strength (and likely causality) of observed associations.

3. The relative importance of different study designs (e.g., cross-sectional versus longitudinal) and the strength of study designs in causal determinations needs better consideration. In addition,

inclusion of at least some occupational studies may be useful as these can assess exposure before disease onset. When appropriate, causality determinations may benefit from greater emphasis on well designed and executed longitudinal studies. This issue is particularly problematic for some outcomes (e.g., renal effects) where inferences possible from cross-sectional study designs are severely limited by a strong likelihood of reverse causation (see specific comments for Appendix 5).

4. The relative importance of consistency of effect estimates versus precision of effect estimates in causal determinations is unclear. Some appendices focus on effect estimates with limited, if any, acknowledgement of their imprecision (e.g., 95% CIs include the null) whereas other appendices don't even summarize the effect estimates for which confidence intervals include the null.

5. It is unclear how outcome categories with multiple distinct health outcome measures are considered. Specifically, when findings vary across the multiple health outcome measures within an outcome category, it is unclear how the different health outcome measures are prioritized for causal determination. Per the EPA weight-of-evidence approach, a very strong association with only one health outcome or coherence of findings across multiple health outcome measures can contribute to a causal determination. However, for a number of outcome categories, it is not clear how this approach is operationalized. The CASAC acknowledges that this is an especially complex and challenging issue to address. The CASAC suggests several possible approaches for consideration: (1) use finer outcome categories for causal determination where there is clear divergence of causality determinations for the health outcomes within an outcome category or information regarding important causal associations are not acknowledged because the outcome category's overall determination is less than causal; or (2) for each outcome category, include a summary describing the strongest lines of evidence that contributed to the final causal determination as well as an explicit discussion of how evidence supporting discordant causal determinations for health outcomes within an outcome category was considered and reconciled. Although not exclusively an issue for summaries of several outcome categories within Pb's "Reproductive and Developmental Effects," uncertainty regarding causality determinations was particularly evident for "Pregnancy and Birth Outcomes" and "Female Reproductive Function" (see specific comments and discussion on preterm birth and pre-eclampsia health outcomes within the "Pregnancy and Birth Outcomes" category in the Appendix 8 consensus response).

6. In determination of causality, the contribution of many recent studies (e.g., those considering Pb as part of a chemical exposure mixture or considering modification of Pb's effects by sex, psychosocial, or other factors) were excluded reportedly because their design and findings were not amenable to the causal framework. Pb exposure rarely occurs in isolation, so just because a study did not measure the other components of what were likely mixed exposures does not mean that co-exposures were not present. In many cases, exposure mixtures studies provide estimates of Pb's effects, adjusted for potential confounding by other chemicals. Variation in confounding adjustment is inherent in the literature so it's not clear why this precludes consideration of such studies. Similarly, if there were consistent difference in Pb's effect by a modifying factor, the association of Pb with an outcome in a high-risk stratum could contribute to causal determination. In fact, in some cases, exposure associations may only be ascertainable among the most vulnerable subset of the population; absent assessing effect modification, associations would be mis-specified as null. This is already effectively being done in the causal determinations that assess Pb-associated outcome risk by age groups (e.g., "Nervous System

Effects Ascertained During Childhood, Adolescent and Young Adult Lifestages" versus "Nervous System Effects Ascertained During Adult Lifestages").

7. The CASAC suggests the EPA rely on consistency of evidence from human studies for causality determinations if toxicologic data are absent, inconsistent, or restricted to higher doses. This is relevant for the causal determination of pregnancy outcomes, specifically preterm birth, and preeclampsia.

8. The ISA needs more text regarding the relative role of the drivers of causal determination for each outcome assessed – for example, an explanation such as: "for health effect 'x', causality determinations were based on longitudinal study findings confirmed by cross-sectional studies with contributory animal evidence."

9. The ISA includes a lot of repetition making it difficult to understand how causal determinations were made. Consistent use of forest plots (or other data visualization techniques) could be used to demonstrate the role of different exposures, outcomes, and study findings in the causal determinations.

The CASAC has comments on several of the causality determinations in the ISA, which are summarized in the following table. Further details on these recommendations are provided in the consensus responses for each specific appendix.

Outcome	**ISA Determination**	**CASAC Comments**
Cognitive function decrements in adults	Likely causal	Recommend changing to "Causal" due to strength of evidence
Renal effects	Causal	Suggest including additional evidence (Lin et al., 1999, 2003) to strengthen the support for this determination
Immunosuppression	Likely causal	Recommend better justification
Sensitization and Allergic Response	Suggestive	Recommend better justification
Pregnancy and Birth Outcomes	Suggestive	Recommend revisiting the causal determination for possible upgrading once more careful review of the separate health outcomes is complete. Consider adding separate outcome categories or adding a more detailed summary of how discordant causal determinations for health outcomes within this outcome category are reconciled (e.g., evidence supporting causal versus suggestive determinations)
Pre-term birth	Part of Pregnancy and Birth Outcomes – Suggestive	Recommend more careful review of the literature which may lead to upgrading the causality determination
Pre-eclampsia	Part of Pregnancy and Birth Outcomes – Suggestive	Recommend more careful review of the literature which may lead to upgrading the causality determination

Outcome	ISA Determination	CASAC Comments
Female Reproductive Function	Suggestive	Recommend changing to "Likely causal" based on the age at menopause outcome
Total (nonaccidental) mortality	Causal	Recommend adding cardiovascular mortality, determined to be "Causal," as a subcategory of total mortality. Both could possibly move to a new appendix. The cardiovascular appendix still needs to appropriately consider cardiovascular mortality because it informs the overall cardiovascular effects causal determination
Cancer	Likely causal	Recommend correcting the conclusion to indicate the causal determination is based on animal study evidence

Appendix 3 – Nervous System Effects

As noted in the overarching comments above, this appendix provides a comprehensive synthesis of an enormous body of literature and the authors are to be commended for this work. In addition to the general issues identified in the overarching comments on all eight health appendices, Appendix 3 would benefit from clarification and correction of its content in several areas.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

1. Reframing/clarification would improve Appendix 3's discussion of the strengths and limitations of the use of cognitive and behavioral assessments in epidemiologic studies:

    a. Behavioral checklists are described as less valid and more prone to measurement error than standard psychometric tests in children. The cited studies use standardized and validated behavioral checklists which include indicators to flag potential biased responses. Assuming the latter are taken into account, these instruments are robust indicators of the functional implications of a range of cognitive and behavioral issues. Studies using such instruments should not be relegated to having lesser value.

    b. The linguistic and cultural appropriateness of using psychometric tests in populations for which the tests were not originally designed should be considered in assessing study limitations.

2. In considering the childhood/adolescent/young adult literature, the strengths, limitations, and causal implications of those studies assessing crime should be considered separately from studies assessing externalizing behavior. This is justified based on the multi-factorial nature of criminal activity and the multiple risk factors related to its occurrence which are distinct from externalizing behaviors. Of note, the EPA IRIS Office and collaborators are currently engaged in a systematic review of the evidence associating Pb with both externalizing behavior and violence/crime (Shaffer et al., 2022).

3. Despite extensive discussion of the limitations of the epidemiologic literature, there is limited, if any, discussion of the limitations of the toxicologic literature. For example, water maze escape times can reflect motor and/or sensory deficits (independent of cognitive skills) – controlling for these factors or failure to acknowledge these factors would both be important study limitations.

4. For toxicologic literature, there needs to be more emphasis on inhalational exposure studies because of the potential for direct central nervous system (CNS) exposure (via this route) that is not necessarily reflected in blood Pb levels.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 3 generally does an excellent job integrating evidence from mechanistic studies to inform conclusions on biological plausibility. However, reconciling dosing differences between human and toxicologic studies needs more consideration and discussion (e.g., toxicologic literature exposure levels seem high, at the equivalent of up to 30 µg/dL in blood).

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

1. There is good scientific rationale for sex-specific neurologic effects – more information regarding this issue is needed, as some sections of the ISA imply differential effects by sex are evidence of inconsistency and, therefore, a weakness in determining causality.

2. It is difficult for the reader to know which studies are critical in causality determinations. For example, not all new studies reviewed since the 2013 ISA are applicable to reaching a causality determination. Where new studies are considered, an explanation of whether these enhance, reinforce, or repudiate the prior literature (and related causal conclusions) is needed. Being more precise regarding what information is critical to causality determinations, including information from older studies, would be valuable both for understanding causality and examining the dose-effect relations of Pb with child cognition. For example, demonstrating concordance of the more recent literature with key past studies of child cognitive function (e.g., Hu et al., 2006; Wasserman et al., 2003) as well as inclusion of studies of academic abilities (Evens et al., 2015) would be valuable.

3. There is inherent variability across domains of cognitive function in adults. Discordant Pb associations across these domains reflect the inherent biologic variability in the outcomes, so variability of findings is expected. That is, divergent findings are likely a reflection of differences in outcome pathophysiology, not an indicator of inconsistency in the literature, although this is mistakenly a conclusion in the ISA. Instead, it is useful to consider the relative consistency of findings within a specific domain of cognitive function, or across an overall summary measure of the different cognitive tests (such as a summary z-score). In the latter case, though, attention must be paid to whether the individual tests that make up the summary score are the same across studies.

4. However, for specific "Neurodegenerative Diseases," for future reviews, the CASAC recommends that the EPA consider each disorder separately (e.g., Alzheimer's disease, Parkinson's disease). This may be useful for causality determinations given the fundamental differences in the pathophysiology of these diseases and variability in the strength of evidence available for each. Evidence is mounting for Pb associations with some of the outcomes and development of separate causal determinations merits further consideration in future reviews.

5. Based on the available literature, including: (1) compelling evidence that cumulative Pb exposure (e.g., tibia bone Pb) is associated with cognitive decline in several longitudinal studies of adults and (2) presumed countervailing studies in which the EPA-stated study limitations are not all relevant to the validity of study inferences (e.g., as above, inherent outcome variability should not be interpreted as a reflection of inconsistencies in the literature), the CASAC finds the relationship of Pb with decrements in adult cognitive function to be "causal" and does not agree with the ISA's designation of this as a "likely causal" association.

Appendix 4 – Cardiovascular Effects

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

The CASAC recognizes the large amount of work that went into selecting studies for inclusion in the ISA, and the effort to summarize them. A recurrent theme of the discussion is the CASAC's desire for the document to elaborate on strengths and limitations of studies summarized. The following comments highlight the need to further contextualize the importance of cited studies, including their strengths and limitations.

Contextualizing studies cited in the ISA is especially important when considering previous studies with strong designs. It may not be necessary to summarize newer studies that do not have stronger designs than previously cited studies; or, if they are included, state their limitations compared to prior work.

In comparison to other health appendices, Appendix 4 summarizes many effect modification results. Thus, below are several issues regarding effect modification that apply specifically to this appendix as well as other health appendices that include effect modification results:

1. Effect modification is described using many different terms that make it challenging for the reader to ascertain the key point being made in the effect modification section of the appendix. Given that some of the findings on effect modification are used later to describe at risk populations, the CASAC recommends that a single term/phrase be selected and used throughout. The section in the appendix could include a formal definition of effect modification, for example, and explicitly acknowledge that various studies used other terms for effect modification.

2. Effect modification section(s) should explicitly acknowledge the role played by findings about effect modification. For example, do they play a role in causality determinations, in identification of at-risk groups, or in establishing biological plausibility or mechanistic pathways? The CASAC notes that not all effect modification studies are useful for the purposes of the appendix.

3. The CASAC notes that effect modification analyses are sometimes done post hoc. Acknowledging the limitations of post hoc analyses is important. It would be advantageous for the appendix to be focused on literature where: (a) effect modification analyses are robustly justified and planned a priori; (b) where the study's inclusion fits the role that effect modification results play in the appendix (see comment above); and/or (c) where the modifier is well defined and does not include cardiovascular disease-related outcomes as part of its definition.

Additional comments on effect modification in cardiovascular disease are as follows:

1. The CASAC highlights that stress and allostatic load are potentially ill-defined modifiers: stress can be measured in many different ways and clarity on the specific definition is needed. Allostatic load includes cardiovascular disease-related outcomes as part of its definition, introducing circularity in the modifier and the outcome.
2. The ISA reports that no effect modification was observed by Menke et al. (2006), but does not report on effect modification results based on *a priori* hypotheses that were reported in Lanphear et al. (2018), who relied on the same cohort with longer follow-up. Lanphear et al. (2018) reported: "Examination of effect modification of the relation between concentration of Pb in blood and key characteristics showed that HRs for participants younger than 50 years were significantly larger than were those for participants aged 50 years or older, for all-cause mortality (HR 2·24, 95% CI 1.50–3.34 *vs* 1.53, 1.18–1.98; p=0.003 for interaction), cardiovascular disease mortality (2.93, 1.60–5.36 *vs* 2.08, 1.35–3.19; p=0.01), and ischaemic heart disease mortality (4.68, 2.42–9.05 *vs* 2.46, 1.51–4.01; p=0·02). The HR for cardiovascular disease mortality was significantly larger for non-smokers than smokers (HR 2.19, 95% CI 1.47–3.26 *vs* 1.32, 0.86–2.05; p=0.03 for interaction)."

There is a new study that reports on the effects of Pb on blebs and tears of endothelial cells that is not included in the Appendix (van Strijp et al., 2023). Although this paper may have been published after the cutoff date for the review, the CASAC recommends that it be included because the publication strongly supports the causality determination, and a mechanistic understanding of Pb's effects on initiation of atherosclerosis.

The strengths and limitations of cited articles are not always clear. While the conventional understanding is that longitudinal studies have a stronger design when compared to cross-sectional studies, longitudinal studies with large loss-to-follow up may have strong limitations that relate to bias. Similarly, studies where the exposures are extremely high may underestimate the "per µg/L" difference in outcomes for the general population if the underlying dose-response curve exhibits ceiling effects. Use of terms such as "increasing" or "decreasing" when referring to exposures or outcomes in cross-sectional studies overstates the strengths of such studies; the terms should be reserved for longitudinal studies where within-person increases/decreases can be observed within the study population, as opposed to higher/lower between-person exposure differences.

The CASAC notes a need to acknowledge the exposure window reflected in exposure biomarkers (short- or long-term exposure). In turn, highlighting the strengths and/or limitations of studies using a particular biomarker's association with a particular outcome is needed. For example, blood pressure at one point in time may be more strongly associated to blood Pb (both being potentially transient measures), whereas hypertension may be more strongly associated with bone Pb (i.e., a chronic health condition linked to cumulative exposure). A study examining blood Pb in relation to hypertension may be limited by not using the more appropriate biomarker for that outcome.

The CASAC acknowledges that Pb in blood at any given point in time may be the result of new exposures as well as mobilization of Pb from bone. This may sometimes cloud interpretations of associations of blood Pb with health outcomes. It is noted that the weight of the evidence of observational studies of blood Pb and cardiovascular disease (CVD) outcomes and the validity and reliability of blood Pb as a biomarker is bolstered by evidence from a recent study that followed the National Health and Nutrition Examination Survey (NHANES) participants for mortality to assess the association between blood Pb levels and cardiovascular mortality, while controlling for other risk factors (Ruiz-Hernandez et al., 2017). Thus, it is important to describe the interpretations of blood Pb more precisely as a biomarker, given the studies that demonstrate its validity and reliability.

The Weisskopf et al. (2009) study on mortality had to be corrected to adjust for selection bias. Weisskopf et al. (2015) should be cited instead.

The CASAC identified potential areas for future inclusion or research, such as: 1) estimating the population attributable fraction of cardiovascular outcomes attributable to Pb; and 2) inclusion of studies that use a natural experiment design or a randomized control trial (RCT) design, as these studies can further bolster causality determinations. Examples of these include, respectively, examination of health effects associated with the removal of leaded gasoline from National Association for Stock Car Auto Racing (NASCAR) races (Hollingsworth and Rudick, 2021); and results of RCTs evaluating chelation treatment on CVD outcomes such as the Trial to assess chelation therapy (TACT) study (Lamas et al., 2014; Lamas et al., 2016; and Escolar et al., 2014), and especially TACT2 RCT results where heavy metals were assessed.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Mechanistic studies are well integrated in this appendix.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

The causality determination appropriately reflects the strengths and limitations of mechanistic studies, given that such studies are well suited to identifying specific pathways.

The causality determination appropriately reflects the extensive epidemiologic literature examining the association between Pb exposure and CVD outcomes.

Please comment on the current organization of this appendix and the decision to develop a single causality determination for cardiovascular effects.

Cardiovascular mortality and all-cause mortality are artificially separated. It may be useful to create a single appendix on mortality and discuss findings for cause-specific mortality, specifically cardiovascular mortality, within that. For additional comments on mortality, please see the consensus response to the Appendix 9 charge questions. For Appendix 4, the CASAC finds that the literature on cardiovascular mortality should continue to be discussed in that it contributes to the causality determination for cardiovascular effects overall. However, a determination of causality specifically for cardiovascular mortality should be included in the appendix that discusses mortality. The CASAC agrees with having a single causality determination for cardiovascular effects as "causal." This is consistent

with the observation that cardiovascular effects are inter-related, as well as with a single cardiovascular causal determination in the 2019 Particulate Matter ISA and 2020 Ozone ISA.

Appendix 5: Renal Effects

This appendix provides a comprehensive synthesis of an extensive literature, and the authors are to be commended for this work. In addition to the general issues identified in the overview of all eight health appendices, Appendix 5 would benefit from clarification indicated below.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

Although several new longitudinal studies contributed to upgrading the Pb-renal effect association to "causal," this is an example where longitudinal studies are not all robust designs. In particular, newly cited studies assessing Pb exposure and renal function with short-term follow up (e.g., across two menstrual cycles) or among cohorts with chronic renal disease may still be impacted by reverse causation and this needs better acknowledgement. Similarly, the importance of reverse causation in the interpretation of newer cross-sectional studies needs to be more clearly acknowledged. That said, studies of cumulative Pb dose that measured tibia Pb, for example, would be much less likely subject to reverse-causality concerns than studies using blood Pb. Longer follow-up and starting with people without kidney dysfunction would be better still.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 5 does an excellent job integrating evidence from mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

Causality determinations would be strengthened by inclusion of key publications vis-à-vis study design utility. For example, studies by Lin et al. (1999 and 2003) identified patients with chronic renal insufficiency and a moderately elevated Pb body burden as estimated using provocative chelation with ethylenediaminetetraacetic acid (EDTA) by Pb excretion in the following 24-72 hours. Elevated Pb burden was defined differently in the different studies but was generally based on Pb excretion > 80-150 µg but < 600 µg in the collection period. Patients were randomized to receive chelation therapy or a placebo and followed longitudinally. Renal function declined more slowly in persons randomized to receive chelation therapy. The CASAC believes that evidence from this randomized controlled trial that associated longitudinal declines in renal function with Pb exposure estimated using a more health relevant Pb pool, is a useful contribution to the evidence of Pb's renal effects that can strengthen the support for the causality determination.

Appendix 6: Immune System Effects

This appendix provides a comprehensive synthesis of an extensive literature, and the authors are to be commended for this work. In addition to the general issues identified in the overview of all eight health appendices, Appendix 6 would benefit from clarification as provided below.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

1. Key limitations of the epidemiologic literature (e.g., limited power due to small numbers of cases of asthma or atopic disorders) are well described.

2. There is substantial diversity of outcome measures in this appendix, but their relative importance (strengths/limitations) is not clear. Functional measures of immune system processes are identified as preferrable, but the appendix includes review of a substantial body of research focused on non-functional outcomes (e.g., serum immunoglobulin levels).

3. There is no discussion of the relative value of cross-sectional versus longitudinal study designs, especially in relation to atopy and asthma, for which timing of exposure relative to outcome may be particularly important.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 6 does an excellent job of integrating evidence from mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

1. Causality for "Sensitization and Allergic Response" is determined to be "suggestive" in the current ISA but "likely causal" in the 2013 ISA. The logic behind this decision is not always clear but appears to be driven by less support from recent cross-sectional studies and persistently null longitudinal studies. In this context, some discussion would be helpful regarding: (i) whether the design of more recent studies, as compared to previously reviewed studies, explains the divergence of findings; and (ii) whether more recent studies were better designed and therefore more important in causality determinations than previously reviewed research.

2. The CASAC recommends that the EPA provide better justification for the immunosuppression causality determination.

Appendix 7: Hematological Effects

This appendix provides a comprehensive synthesis of an extensive literature, and the authors are to be commended for this work. The approach to causal determination used in this appendix is reasonable, although it appears to differ from approaches used for other health outcomes.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

Most of this literature is cross-sectional and limitations related to that design are appropriately reviewed as well as strengths of newer studies vis-à-vis more comprehensive adjustment for potential confounders than was done in earlier studies. However, the CASAC recognizes that Pb's effect, for example, in the heme synthetic pathway, the specific enzymes that are inhibited and at what levels, genetic polymorphisms that influence these effects, and the health parameters that are affected, are some of the best understood and documented of all of Pb's specific organ system effects.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 7 does an excellent job of integrating evidence from mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

In contrast to other health outcomes, the approach to causal determination for hematologic effects seems largely motivated by strong toxicologic data and historical epidemiologic evidence (including both population-based and occupational cohorts) among populations with higher levels of Pb exposure than are common today. In that these data sources are the strongest evidence supporting the causal determination, this approach seems appropriate. Regardless of how evidence was assembled and weighted for the appendix, the causal determination is seemingly irrefutable.

Appendix 8: Reproductive and Developmental Effects

This appendix provides a comprehensive synthesis of an extensive literature, and the authors are to be commended for this work. In addition to the general issues identified in the overview of all eight health appendices, this appendix would benefit from clarification as provided below.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

1. Studies using bone Pb as an exposure biomarker are categorized as being limited by cross-sectional design. As described previously in the consensus response of overarching issues relevant to all health outcomes, "Cross-sectional study designs are repeatedly identified as having limitations, including uncertainty regarding the directionality of associations. This limitation is applied to studies in which bone Pb is used as a biomarker of exposure. However, because bone Pb levels reflect cumulative exposures that occurred over previous years to decades, uncertainty regarding directionality of association is unlikely to be relevant." It may be more appropriate to recognize this misinterpretation by being explicit and explaining that such studies are cross-sectional in relation to outcomes but retrospective and cumulative in respect to Pb dose.

2. One of the more informative study designs cited in this appendix includes several studies assessing women and men seeking treatment for infertility/subfertility with Assisted Reproductive Technology (ART). This design has many strengths including prospective exposure assessment and ascertainment of early fetal loss. None of these strengths are clearly acknowledged. Instead, such studies are considered limited because of concerns regarding generalizability and potential participation bias. Generalizability is not a threat to study validity and the potential for participation bias (e.g., those seeking fertility treatment are more likely to be Pb exposed than those who don't seek treatment) seems unlikely. This interpretation risks underestimating the importance of such studies to causal determinations.

3. In several examples, confounding is inappropriately identified as a limitation (or explanation) for findings. For example, an observed attenuation of associations of blood Pb with age at menarche as blood Pb levels decline is attributed to potential confounding by body weight/adiposity. This assertion is problematic as many of the reviewed studies adjusted for body weight or adiposity measures to account for confounding.

4. This appendix would benefit from greater emphasis on the findings from a study with a unique quasi-experimental design in which birth outcomes among residents living near a racing track were observed before and after removal of Pb from racing car fuel and compared to residents further away from the track using difference-in-difference models (Bui et al., 2002). Given the robustness of this design, its findings should be considered particularly relevant to causal determinations (see causal determination recommendations below).

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 8 does an excellent job of integrating evidence from mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

1. The diverse health outcomes assessed in this appendix have wide variation in the strength of evidence available for causal determinations, particularly for measures subsumed under "Pregnancy and Birth Outcomes," a category of substantial public health importance. "Pregnancy and Birth Outcomes" are assigned an overall "suggestive" causal determination. Because of the wide variation in the strength of evidence, the CASAC recommends revisiting the causal determinations for possible upgrading once more careful review is done for each health outcome component of the outcome category. This recommendation is particularly applicable to "preterm birth" and "pre-eclampsia," health outcomes for which the CASAC finds that the literature supports a "causal" relationship with Pb. The EPA should consider adding separate outcome categories for these health measures, rather than subsuming them in a single overall determination for "Pregnancy and Birth Outcomes." Alternatively, the EPA could add a more detailed summary explicitly describing how discordant causal determinations for health outcomes within the "Pregnancy and Birth Outcomes" outcome category (e.g., evidence supporting "causal" versus "likely causal" or "suggestive") were reconciled.

2. The causality determination of "causal" for "Male Reproductive Function" seems largely driven by strong associations of blood Pb with semen quality, among multiple measures of male reproductive health. Similarly, there is strong data supporting associations of Pb with earlier menopause, among multiple measures of female reproductive health but causation for "Female Reproductive Function" is determined to be only "suggestive." This seeming discrepancy in the approach to causal determination for these parallel outcomes needs to be justified. Specifically, if prioritization of health outcomes is implicit in the causal determination (e.g., a particular outcome's public health importance is used to weight its impact on the overall causal determination for multiple outcome measures) then the apparent choice to prioritize semen quality but not age at menopause needs justification. The CASAC finds that the literature reviewed supports a "likely causal" relationship between Pb and earlier age at menopause, an important indicator of female reproductive health with additional substantial health ramifications well beyond reproduction (CVD risk, osteoporosis risk, etc.).

Appendix 9 – Effects on Other Organ Systems and Mortality

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

In general, Appendix 9 does a good job describing the health studies and they are placed in context well. While there is, in general, appropriate commentary on the strengths and limitations of the epidemiologic studies, there are some places where a little more nuance is needed. Specific points raised that relate to this are:

1. In many places it is stated that cross-sectional studies have several limitations, in particular, related to difficulties in being able to determine temporality. This is correct and appropriate, but it is important to keep in mind that this is more problematic for a biomarker like blood Pb that has a relatively short clearance half-time. Pb in bone has a much longer clearance half-time and as such reflects much longer-term cumulative exposure more than does blood Pb. As such the concerns about temporality are different when this exposure is measured at the same time as an outcome. This is worth stating explicitly.

2. In many places it is stated that "…with BLL [blood lead level], it is difficult to characterize the specific timing, duration, frequency, and level of Pb exposure that contributed to associations…." This is true, but it should be noted that this is also an issue with bone Pb since bone Pb only provides an estimate of cumulative exposure and so "specific timing, duration, frequency, and level of Pb exposure" is also not something that can be teased apart.

3. Regarding elevated fasting glucose, it is stated in several places that "…the small sample size (n = 150) in this study reduces statistical power, as well as the likelihood that an observed result reflects a true effect." The first part is correct, the second part, however, is a mischaracterization. With a continuous outcome, that the observed result might not reflect a true effect is captured by the confidence limits, which is captured by the first part of the sentence. There is really no reason beyond that to doubt that the confidence interval from a small study fails to cover a true effect any more than it would in a larger study. The situation is a little different for dichotomous outcomes, because when the sample size is very small, then getting just one or two individuals

wrong can dramatically affect the point estimate, but this is not the same with a continuous outcome.

4. Better adjustment for thyroid medication use is touted as a strength on p. 9-30, line 6, but it is not clear that it is a confounder (why is it related to Pb?). Instead, it would seem, if anything, it might be an effect modifier. This same issue arises on p. 9-39, line 5, where hormone therapy use is likely an issue of effect modification rather than confounding.

5. As a somewhat more general issue, there are many places (including other appendices) where in discussing findings from cross-sectional studies, language of "increase" and "decrease" of exposure and outcome is used. These imply longitudinal changes within a person, which these studies do not estimate. Instead, language of "higher" and "lower" for exposure and outcome is more appropriate.

6. It is noted that some health outcomes are not covered in Appendix 9, such as cataracts (Schaumberg et al., 2004) and dental caries (Wu et al., 2019; Yepes et al., 2020). The CASAC recommends that these outcomes be followed closely in the future for possible inclusion.

7. In section 9.1, the discussion of serum markers of liver function talks about the levels varying within normal ranges and so there is less clinical significance. But to the extent that Pb may affect levels even when those are largely in a normal range within a population, this can still have important population health implications (as opposed to individual health implications) and so should not be dismissed as less strong results—they are still meaningful findings. This applies generally to dimensional measures (as opposed to the binary disease or not situation).

8. Sometimes the text suggests that lack of testing for interactions is a limitation. However, not accounting for an interaction, if it was truly there, would most often tend to make the overall association appear weaker (relative to the association in the more vulnerable group). So, this should not necessarily be considered a major limitation.

9. As a more general point, more attention should be given to reporting sample sizes when describing different studies. Larger studies typically would carry more weight than smaller ones.

10. At times it seemed that null associations may have been interpreted as reflecting no effect of Pb when careful attention to the relevance of the time window of exposure reflected by the Pb biomarker for the biological effect in question may be an important consideration. For example, a null association with a single blood Pb measure may be irrelevant if the critical exposure window for biological effects is many past years of exposure. BLL may correlate with that somewhat (if conditions are right), but null associations could easily occur with BLL simply because it does not correlate well enough with long-term cumulative past exposure. Thus, consideration of exposure time windows should be incorporated into the synthesis of evidence used for causality determinations.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Overall, the integration of evidence to inform conclusions on biological plausibility is appropriate.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

Overall, the causality determinations seem largely appropriate given the strengths and limitations described in the appendix. The following additional specific comments are provided:

In sections 9.5.3.1 & 2, the epidemiologic studies of the bone mineral density and osteoarthritis outcomes seem particularly subject to reverse causation because most studies are cross sectional. This is stated at the end of 9.5.3.2, but not 9.5.3.1. In both sections, though, the concern is prominent and so the document could use more text about this. (This point is discussed a little in the summary, section 9.5.6). The CASAC notes, though, that the toxicologic data is quite strong for these outcomes and this evidence is summarized in the text, so the "likely causal" determination still is appropriate.

Additionally, for all-cause mortality, a follow-up study by Weisskopf et al. (2015) of an earlier study (Weisskopf et al., 2009) is not discussed. The 2009 paper has a correction to it, and the 2015 paper goes into issues of selection biases that can affect these kinds of studies and presents what are likely the most appropriate results to cite—specifically those in Table 2, model 4 after accounting for selection bias issues. The 2015 paper should be cited in Appendix 9 and added to Table 9-3.

The animal and toxicologic literature are often just accepted as is, without much consideration and discussion of the limitations of these types of studies. Such discussion should be included, as well as discussion of whether appropriate experimental controls were done or discussed in the various papers.

While the CASAC generally agrees with the final causality determinations, the CASAC also suggests that some more explicit text on what factors/studies are most driving the causality determination for each outcome would be a valuable addition.

Please comment on the current organization of this appendix and the decision to incorporate a separate causality determination for exposure to Pb and total (nonaccidental) mortality.

Overall, the organization of this appendix is appropriate. The CASAC agrees that including a section on all-cause mortality is important. Within this section, in addition to a causality determination for overall mortality, the ISA should also provide a causality determination specifically for cardiovascular mortality. All-cause mortality may integrate across many effects (many of which may not have sufficient evidence on their own for separate effects) that act synergistically to have effects beyond just the sum of the independent effects. Alternatively, cause-specific mortality studies may observe stronger associations because unaffected outcomes are not being combined with affected outcomes, and findings can be linked to specific mechanisms (e.g., blood Pb related to blood pressure, which is linked to cardiovascular mortality). For these reasons, separate causality determinations for all-cause mortality and cardiovascular mortality should be provided in this appendix. The CASAC thinks that in the future, separating out more specific causes of mortality will become more important as the evidence base grows.

Appendix 10: Cancer

This appendix provides a comprehensive synthesis of an extensive literature, and the authors are to be commended for this work.

Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and toxicologic studies.

This appendix appropriately identifies important limitations to the human epidemiologic literature assessing Pb and cancer including the use of cross-sectional or ecologic designs (particularly problematic for an outcome with a long latency period between exposure and disease development) and the absence of biomarkers of exposure and/or confounding adjustment (e.g., smoking status) in many studies, including occupational cohort studies.

What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

Appendix 10 does an excellent job of integrating evidence from mechanistic studies to inform conclusions on biological plausibility.

To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

The causality determination is largely based on strong animal evidence available at the time of the 2013 ISA as there have been essentially no new Population, Exposure, Comparison, Outcome, and Study Design (PECOS) relevant toxicology studies since then. Given this approach, the final conclusion in the ISA indicating, "Overall, the collective evidence is sufficient to conclude that there is likely to be a causal relationship between Pb exposure and cancer incidence and mortality" needs to be corrected as the causal determination is not based on human cancer mortality.

Further, the EPA's judgment that the epidemiologic evidence for cancer mortality is inconsistent needs to be corrected. (See Dr. Sheppard's individual comments for details.) While the CASAC finds that the epidemiologic studies reviewed in the document are consistent, other aspects of this evidence, such as the use of blood Pb with a clearance half-time of 30 days and the potential for reverse causality, provide sufficient rationale to downweight this evidence stream in the causal determination.

Appendix 11 – Welfare Effects of Pb Exposure

Introduction (Section 11.1) - The welfare effects appendix has an introductory section that includes concepts and tools for evaluating Pb effects on organisms and ecosystems. To what extent do the choice and emphasis of topics in the introduction provide adequate context for the evaluation of ecological effects of Pb in the ISA?

The Introduction to Appendix 11 provides an excellent synopsis of the available toxicity data and the approach used in the ISA. It includes a good, balanced discussion of concepts, models, and approaches. The document is well written, well organized, and does an adequate job of summarizing the "new" published ecotoxicolgic data (post-2013). The discussion of biotic ligand models (BLMs) and multiple linear regression (MLR) models is especially useful. The organization of the document based on endpoints and biological complexity is logical and appropriate, and addresses the levels of uncertainty with the various organizational levels.

The LECES approach to culling the literature is fine, but the exposure cutoffs for soil (230 mg/kg) and especially for water (10 mg/L) seem to be quite low. Additional discussion/justification for the application of the literature cut-off values is warranted. In some cases, the cut off values are very close to the current criteria/standards. For example, the aqueous freshwater cutoff is 10 µg/L while the current AWQC is 2.5 µg/L at 100 hardness. Section 11.1.4 provides a good description of how sub-organismal, organismal, and ecosystem level endpoints relate to each other and how they are addressed in the assessment. This section also makes a very important point, i.e., the difficulty of attributing observed effects in the field to Pb (or even more difficult, atmospheric Pb). This point must be stressed in later Appendix 11 sections. Section 11.1.6 provides a good overview summary of the importance of Toxicity Modifying Factors (TMF) in the various matrices and this reflects the current "state-of-the-science" for evaluating bioavailability and predicting toxicity. This is a major improvement over past evaluations and reflects the recognition of all the research that has been developed in the past decade and the value of EPA/Industry cooperative research agreements, like the Metals Cooperative Research and Development Agreement (CRADA). Section 11.1.7 provides an excellent synopsis of criteria/standards for each matrix from different locations around the world and reflects the state of the science for the jurisdictions. It is important to note that much of the recent research and toxicity model development have come from Europe, Australia, and Canada rather than the U.S.; perhaps there could be benefit to international research cooperative arrangements.

There does not seem to be a clear discussion of environmental fate and source apportionment for Pb. This would seem to be important and necessary if the ultimate goal is to establish air quality criteria. Much of the Pb in the environment today no doubt comes as a consequence of terrestrial runoff, direct discharge from industry or publicly-owned treatment facilities, and/or transport from non-point sources into freshwater systems. Atmospheric input is important and will influence sources like terrestrial runoff, but some discussion about the environmental fate processes and the contribution of atmospheric sources should be addressed in the document.

Equation (2) on page 11-24 is not correct in its current form in the text. The formula should properly be written as: FCV = CF [$e^{1.273\ (\ln(\text{hardness})-4.705}$] for the Final Chronic Value (FCV) but the conversion factor (CF) for Pb is hardness based as is calculated as CF=1.46203-[ln(hardness)(0.145712)]. This section should be clarified/corrected.

Terrestrial (Section 11.2) - Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.2 provides a good summary of the available data and new data (post-2013). The discussion of terrestrial ecosystems is comprehensive and highlights several important advances since the last ISA. The text is clear and very well-written. Available new data seem to support previous observations and provides no basis for heightened concerns, thus the justification is adequate and appropriate for the conclusion that there are no changes to the previous causality determinations. All in all, this section of the ISA supports the causality determinations in Table 11-2. The literature review and analysis are remarkably consistent with the causality determinations from 2013, supporting the decision to leave them unchanged.

Regarding the question about the relationship between Pb exposure and effects on individual organisms and ecosystems, the statement in Section 11.2.1 summarizes the situation quite well: "Although evidence for effects on growth, reproduction, and survival at the individual organism level and in simple trophic interactions makes the existence of effects at higher levels of organization likely, direct evidence is relatively sparse and difficult to quantify. The presence of multiple stressors, especially including other metals, continues to introduce uncertainties in attributing causality to Pb at higher levels of organization." The conduct and interpretation of community-level toxicological studies is extremely difficult due to the number of experimental variables. This is especially challenging when conducting tests in a real-world system (non-laboratory) or an uncontrolled field-based assessment. The characterization of "exposures" is particularly difficult when the goal is to ascribe observed environmental changes to a single contaminant (Pb; there are almost always multiple metals in real-world exposures) and it is even more challenging to apportion those exposures into atmospheric Pb sources versus waterborne or other fate pathways.

The complex relationship between atmospheric Pb and the deposition and subsequent reactions that can affect Pb speciation, bioavailability, and toxicity is an area that needs more investigation. As was noted in the text, investigations using Pb-spiked soils are difficult to interpret and may ultimately contribute very little to real-world exposures because their methods may not duplicate natural processes. The significance of Pb "aging" on the interpretation of study results is interesting and a useful addition. One must remember, though, that an atom of Pb is an atom of Pb. It does not "age" in any meaningful way; rather, what changes are the complexes with which the Pb is associated – e.g., natural organic matter, soil/sediment minerals, speciation, etc., all of which affect Pb bioavailability and toxicity. Similar complexities have been observed in aquatic sediment investigations with other metals (i.e., Ni) and test procedures have been developed to artificially reproduce those processes that occur under natural conditions, making it possible to conduct meaningful laboratory metal-spiking investigations (Besser et al., 2011).

The importance of certain types of mycorrhizal fungi in mediating uptake of Pb by plants is a welcome and important observation in this section of the ISA. It is also encouraging to see a growing literature on the influence, or lack thereof, of Pb on microbial community structure and function.

The CASAC is not currently aware of any additional terrestrial studies that have been conducted that should be considered at this time.

Freshwater (Section 11.3) - Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.3 provides an excellent description of the aquatic research that has been conducted since the 2013 ISA effort. It is well-organized and comprehensive. This section does a good job of summarizing the new findings in the context of previous causality determinations. New research refines some key observations and nearly always supports the 2013 determinations. Overall, the CASAC appreciates the presentation of the data followed by the summarization of the salient findings. Substantial new data are identified and included. The authors should be praised for the inclusion of the toxicity modifying factors and bioavailability section (11.3.2). A large amount of new data has been developed since the 2013 ISA

effort and it was appropriately captured in the section. The CASAC is not currently aware of additional freshwater aquatic studies that should be considered at this time.

The authors are to be commended for the level of effort that must have been required to produce these sections. That said, the report is in many places so detailed and inclusive that the reader can easily get lost and forget the reason for the discussion. In many ways this document is too detailed for the purpose, i.e., to address salient new data that affects the conclusion of causality. Sections like 11.2.2.1.10 "Parasites" or 11.2.2.1.11 "Bioturbation/association with sediment" are interesting, but have little to do with the issue at hand. Many of the sections, like most of 11.3.4, are very useful and might well be better placed as a contribution to an update to the 1984 Ambient Water Quality Criteria (AWQC) document rather than in this ISA. Some of the sections discuss large amounts of data that ultimately are of little importance, for example, while Section 11.3.4.1, "Effects on freshwater microbes" (an area not currently considered in developing AWQC or state standards), provides three pages of text, it is ultimately summarized by concluding that the data are negative or uninterpretable. Thus the EPA should consider removing some of the material or relegating it to an appendix to Appendix 11. The CASAC further notes that Section 12.6.1 states that one of the questions to be considered for papers included in the ISA is: "*Are the welfare effect measurements meaningful, valid, and reliable?"* The EPA might also consider adding *interpretable* to that list.

Although the CASAC appreciates the general organization of Appendix 11 based on organism type (plants/ invertebrates/vertebrates) followed by biological complexity (suborganism/organism level), it makes the document difficult to read and follow. Thus the CASAC suggests that presenting the data in tabular form, like Table 11-5, without the detailed discussion, would help. In several places, reference is made to studies that are based on "nominal" or unmeasured concentrations; the CASAC suggests that these studies do not meet the current standards for data acceptability and thus should be rejected and not further considered. In Appendix 12, Section 12.6.1 "Individual Study Quality," it clearly states that in evaluating the scientific quality of studies, one of the questions to be considered is: *"Do the analytical methods provide adequate sensitivity and precision to support conclusions?"* The CASAC believes that it is unlikely that a study based on nominal (i.e., unmeasured) concentrations would meet this criterion. The CASAC also notes that the scientific standard for data acceptability has changed over time and data that were acceptable in deriving the 1984 Pb AWQC would probably not be acceptable today, especially those based on nominal test concentrations. Minimum criteria for acceptability of a study should be integral to the discussion of data. There should at least be a reference in Appendix 11 to the description in Appendix 12. These should be included in this document and referenced as the standard by which the review was conducted. Some minimum standards for relevance and reliability should be documented and applied; publication in a peer-reviewed journal is not generally a *de facto* standard for data acceptability. The CASAC suggests that the guidelines for study acceptability applied in the ISA could employ those used by the EPA Office of Pesticide Programs and the Office of Water.

The CASAC is encouraged to see more studies that are trying to discern the relative contributions of dietary uptake and water column uptake to the total Pb burden in aquatic organisms.

Based on the new studies cited, the CASAC finds that the causality determination for neurobehavioral effects on aquatic invertebrates could be upgraded to "causal."

Section 11.3.6, Community and Ecosystem Effects, does a good job of identifying new studies that were not previously considered (post-2013). No clear new findings result from these studies, as they "reported either negative, positive, or null associations between sediment or porewater Pb concentration and

community and ecosystem effects." No mention is made as to how the study results compared to single-species tests or to extant AWQC or state standards. These types of tests are theoretically ideal for "validating" regulatory values.

The revised Criterion Maximum Concentration (CMC) and Criterion Continuous Concentration (CCC) values for low-hardness, low-pH, low-dissolved organic carbon (DOC) freshwaters are low enough to be in the range of natural waters with elevated Pb concentrations. This may be useful to consider in the possible development of a secondary standard. However, significant negative effects in freshwater systems have typically been observed at Pb concentrations well above typical ambient concentrations. Furthermore, the confounding effects of DOC, pH and hardness in regulating aqueous Pb concentrations make it very difficult to envision a workable model relating air-Pb to aqueous-Pb.

4d. Saltwater (Section 11.4) - Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.4 provides an excellent synopsis of the available ecotoxicity information for Pb in the marine environment. In past assessments, a paucity of studies and data from saltwater systems made it difficult to assess the significance of Pb on individual organisms and marine communities. This situation appears to be changing somewhat, resulting in a few new causality determinations. Although there is not as much information on saltwater organisms as is available for freshwater organisms, the CASAC believes that there is sufficient information to modify some of the 2013 causality determinations that were judged "inadequate" at the time, to "suggestive" or "causal" based on current data. In fact, there is now sufficient data to permit the calculation of regulatory marine values (HC5) using species sensitivity distribution approaches, e.g., Church et al. (2017). The CASAC suggests that the EPA examine the references in the supplemental information contained in Church et al. (2017) to identify data that would help define the causal relationship for survival, growth, or reproduction effects. This suggests that a great deal of research has been developed to characterize the response of marine organisms to Pb since the 2013 ISA and the information would be useful in updating the 1984 AWQC. However, the knowledge of Pb cycling and toxicity in these ecosystems still lags behind terrestrial and freshwater systems. All of these studies characterize the effects of Pb in the marine environment and are no doubt an important contribution to the understanding of the potential effects of Pb in the marine environment. The ISA does a good job of describing the environmental "effects" side of the risk characterization paradigm and the importance of considering bioavailablility to address site-specific toxicity; however, no discussion of the relationship between atmospheric concentration and waterborne (or soil) concentration is provided. No discussion of source apportionment is included, so the relationship between concentrations in air and concentrations in water or soils is missing.

The environmental fate processes affecting Pb in the marine environment are decidedly different from those that govern in the freshwater environment, but this is not addressed. For example, $PbCl^{+}$ is the dominant species over the normal pH range in saline waters, while in freshwater, Pb^{+2} is the predominant species. This is important in that $PbCl_2$ is poorly soluble and will tend to precipitate and be found in bottom sediments. Thus, some discussion of Pb speciation and environmental fate in marine waters would be beneficial in assessing the effects of atmospheric Pb in the marine environment. The generally high pH and salinity of marine systems, and in some cases high DOC as well in estuarine

waters, create conditions in which the percentage of free Pb^{2+} tends to be very, very low. Due diligence requires looking for new insights about Pb in saltwater environments, but it seems much more likely that terrestrial and/or freshwater systems will guide policy regarding environmental Pb toxicity. The EPA should discuss whether the contribution of direct atmospheric Pb to marine waters is significant or whether Pb in marine waters is primarily from freshwater sources. Pb in marine waters may come primarily from terrestrial runoff, direct discharge from industry or publicly-owned treatment facilities, and/or transport from non-point sources into freshwater systems, although this is not discussed in Appendix 11. So, the key questions are (1) are atmospheric contributions of Pb (either direct or via runoff from freshwaters) likely to be an environmental concern in marine environments; and (2) how high would loads have to be (on a local level) to result in marine concerns?

Section 11.4.6 addresses new information addressing saltwater communities and ecosystem effects. This is a nice synopsis of the available data and it highlights that adverse effects were noted in some studies, but not in others. It fails to address the important question: are adverse effects noted at Pb concentrations that are not expected to cause impact? How do reported effect concentrations in community-level studies compare to single-species tests or to derived criteria/standards, i.e., do they validate the methods used for developing regulatory protective values?

The CASAC is not currently aware of additional marine studies that should be considered at this time.

Appendix 12 – Process for Developing the Integrated Science Assessment for Pb

Please comment on the clarity with which Appendix 12 communicates the process undertaken to develop the Draft Pb ISA.

This appendix clearly outlines the overall approach used for ISA development. Figure 12-1, a diagram of the general process for ISA development, is a helpful overview. The first two sections (literature search and evaluation of study quality) are critical steps given the mountain of literature on Pb that has to be parsed for this review, and if they are not done well, the review document will not be adequate for its intended purpose. These two steps make up all but 4 pages of this appendix, with relatively brief mentions of the other important process steps, which include "Evaluation, Synthesis, and Integration of Evidence" and "Development of Scientific Conclusions and Causal Determinations."

Section 12.4 (Relevance and Scope) is useful, noting that we now have three different study inclusion criteria: Population, Intervention, Comparison, Outcome, and Context (PICOC) for environmental research [atmospheric science], PECOS for health studies, and Level of Biological Organization, Exposure, Comparison, Endpoint, and Study Design (LECES) for welfare effects.

For section 12.4.2 (Exposure, Toxicokinetics, and Biomarkers) the CASAC is concerned about the strict geographic inclusion criteria (as the CASAC has noted previously in other reviews) for exposure, limited to US, Canada, western Europe, and Australia.

In general, for PECOS used in Section 12.4.3 (Health), and possibly for PICOC and LECES also, the CASAC notes that the scope of the research screening tools may have omitted some relevant studies that did not fall into the predetermined categories. Section 12.4.3.2 (Epidemiological Studies) notes review articles are generally excluded, but they can be useful in summarizing important findings from older literature. The range of exposures noted in footnotes for Tables 12-3 and 12-4 should be brought into the

text to clarify levels of exposure cut-offs. In Section 12.4.4 (Welfare), geographic exclusion criteria are problematic for welfare exposures, possibly excluding some relevant sources such as mine tailings etc. This section should clarify how high toxicological study concentrations were screened for inclusion, since most are greater than 10 µg/L.

Section 12.5 (Literature Search and Study Selection) notes a literature cutoff of June 2022 for the ISA. This may miss more recent literature given the expected completion dates for a final ISA and the final rule. The EPA should consider extending the literature cutoff date before finalizing the ISA.

In Table 12-5, it should be clarified that race is primarily an indicator of social factors, not usually biological factors (as noted elsewhere in the ISA text). In Table 12-5, the second line notes that studies with more than two pollutants are excluded from consideration owing to co-pollutant collinearities. While these may involve too much uncertainty, this is not necessarily always the case. Importantly, confidence limits or the like will reflect this uncertainty and thus whether the uncertainty is too much can be directly judged. Some of these multipollutant studies may contain useful information, and this hard rule to exclude all of them may not be appropriate.

Executive Summary

Please comment on the clarity with which the ES communicates the key information from the Draft Pb ISA.

The Executive Summary (ES) is very well written and is accessible to the casual reader or those who want a quick high-level overview of the ISA findings. Context for findings based on the 2013 ISA versus studies since that ISA closed in 2011 is provided, and references to additional detail in the Integrated Synthesis (IS) are given. Figure ES-1 is a diagram of the conceptual model of multimedia Pb exposure, but is somewhat difficult to follow. The 6/14/23 EPA presentation on the IRP for the REA (page 7, "Pb-Related Residential Exposure Pathways Potentially Impacted by Ambient Air") has an exposure pathway figure that is easier to follow and could replace Figure ES-1, perhaps with modifications if needed.

Color-coded tables for the causality determinations (Figures ES-2 and ES-3) derived from tables in the IS (Table IS-1 and IS-14) are presented for both health and welfare effects, and important changes in categories are noted for both renal and total non-accidental mortality effects, now both causal. It should be noted in Figure ES-2, as it is in the IS summary of causal determination (Table IS-1 footnotes), that the cardiovascular effects category now combines four separate outcomes used in the 2013 ISA. The new category (marked with "+") of total nonaccidental mortality may be confusing to the reader. It is new only in the sense that this category consolidates previously separate categories, primarily (but not solely) cardiovascular mortality, as explained in Section IS 7.3.9, Mortality. This needs to be clearly noted in a footnote to Figure ES-2. The title for this figure is "Summary of causality determinations by exposure duration and health outcome." Should it be "by life stages and health outcome?" Section ES 6.1 might benefit from a brief discussion of thresholds for the various effects, similar to the 2012 National Toxicology Program monograph on Health Effects of Low-level Lead (https://ntp.niehs.nih.gov/whatwestudy/assessments/noncancer/completed/lead).

The summaries of 2017 NEI data presented in ES.2 will need to be updated to use the now available 2020 NEI data that now include wildfire Pb emissions (3% of total, the 4th highest category). Although

not in the inventories, PM from residential wood combustion also has Pb in it as noted in Section 1.2.3, and may be an important exposure pathway in areas with elevated woodsmoke PM in winter.

Section ES.5, "Exposure," clearly explains the important point that "These exposures are considered air-related if Pb passed through the air compartment at any point prior to plant, animal, or human contact." It discusses "primary" contribution to blood Pb levels without describing which population or subpopulation that this discussion refers to. The EPA should avoid reaching any conclusions on "primary" vs. secondary exposure sources as is correctly indicated in the same paragraph that such conclusions are "situation specific."

Section ES.6.1.1, "Effects of Pb Exposure on Health Outcomes Ascertained in Children, Adolescents, and Young Adults," page ES-9 lines 21-23, says "Notably, evidence suggests that some Pb-related cognitive effects may be irreversible and that the neurodevelopmental effects of Pb exposure may persist into adulthood (2013 ISA)." There is evidence for Pb-related cognitive effects persisting into adulthood. However, the CASAC is not aware of any studies examining irreversibility and therefore the EPA should use caution in using that term.

Another important finding is summarized in section ES 6.1.1, that recently available evidence does not provide evidence of a threshold for the observed neurodevelopmental effects across the range of blood Pb levels examined.

Section ES-7, "Key Aspects of Health and Welfare Effects Evidence," is a good summary of conclusions on key policy-relevant topics in this ISA. Section ES.7.1.2, Air-Pb–to–Blood-Pb Relationships, page ES-17 lines 6-10 states: "Although slope factors increase with decreasing air Pb concentration, it is possible that the contribution from non-air exposure pathways may lead to the higher slope factors at lower air concentrations. In other words, in older studies in which leaded gasoline or local sources were a major contributor to air Pb, there may be a greater likelihood of discerning the true effect of air Pb on blood Pb due to relatively less contribution from non-air exposure pathways." These statements are not correct. At lower air concentrations, higher air-Pb-to-blood-Pb slope factors are not due to other sources. When there are lower air concentrations (and lower exposure), there is more efficient uptake to red blood cells, which may be a biologic reason for higher slope factors.

In Section ES.7.1.3, "Concentration-Response Relationships for Human Health Effects," page ES-17 lines 23-31, while discussing both the air-Pb to blood-Pb relationship and the blood Pb to cognitive response relationships the point is made that the slopes of these relationships tend to be greater at lower values of the independent variable. If the EPA is interested in examining the air-Pb, blood Pb-cognitive, and blood Pb-cardiovascular disease response relationships across a wider range of blood Pb concentrations (e.g., 0.01 $\mu g/m^3$ to 1.5 $\mu g/m^3$ for air-Pb and 10 $\mu g/L$ to 300 $\mu g/L$ for blood Pb), they should consider adopting a non-linear approach. Most Americans, however, are exposed to air Pb levels below 0.04 $\mu g/m^3$ (https://www.epa.gov/air-trends/lead-trends) and blood Pb concentrations below 35 $\mu g/L$ (https://www.cdc.gov/exposurereport/blood_lead_early_release.html). The relationships of air-Pb to blood-Pb, blood Pb to cognition, and blood Pb to cardiovascular disease are likely to be linear at these lower levels (Richmond-Bryant et al., 2014; Lanphear et al., 2005; Lanphear et al., 2018). Thus, it may be preferable to rely on simpler linear models. Still, the EPA should confirm whether the air-Pb to blood Pb relationship is linear or non-linear at air Pb concentrations below 0.04 $\mu g/m^3$.

Please provide recommendations on any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

There is a sentence in Section 12.4.2, page 12-6 that is a clear description of exposures that might be brought into the ES: "… studies containing Pb concentrations in other media (soil, dietary sources, consumer products, occupational sources, and ammunition) were included because cumulative body burden can occur as a result of contributions from multiple exposure pathways (i.e., ingestion of Pb-containing soil by children) and the origin of Pb can be difficult to determine as stemming from an air-related source."

The CASAC does not have any additional recommendations for any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

Integrated Synthesis

Please comment on the usefulness and effectiveness of the summary presentation in the IS and provide any recommendations or alternate text that may improve the synthesis of available information across subject areas and the communication of key findings.

The IS is well written and organized, and appropriately references ISA Appendices instead of citing individual literature. For each causal category, each section ends with a very useful comparison between evidence from the 2013 ISA and this ISA that details what evidence is new in this review. Overall, the IS represents a very useful and effective approach to present the detailed findings of the ISA. Additional clarifications that would improve the IS are provided below.

- Corrections to the overall conclusions box on page IS-1: This has total mortality under the likely causal category instead of causal. Neurodegenerative disease is listed here as likely causal but elsewhere is suggestive. The last bullet within the health effects should be more explicit about separating health effects that are due to higher exposure and populations that are at higher risk because of higher exposure versus higher sensitivity.

- Table IS-1 on page IS-24 groups total mortality under "Effects on Other Organ Systems" (consistent with IS.1.2.2 "Organization of the ISA", page IS-7), while Figure ES-2 lists it as a major separate category after the "Effects on Other Organ Systems." The organization needs to be consistent across these tables and with the organization of the ISA. As noted in the recommendations for total mortality in the consensus responses for Appendices 4 and 9, this health outcome should be listed as its own separate category everywhere, including Table IS-1.

- The new category of total nonaccidental mortality (marked with "+") may be confusing to the reader. It is new only in the sense that this category consolidates previously separate categories, primarily (but not solely) cardiovascular mortality, as explained in Section IS.7.3.9, Mortality: "… the 2013 Pb ISA evaluated studies of all-cause mortality together with studies examining cardiovascular mortality, and these studies were all included within the CVD chapter." This change in how total mortality is presented in these two tables needs to be noted in the table footnotes, as is done for cardiovascular mortality in Table IS-1 (footnote a).

- To the extent possible in the context of this review, the EPA should consider moving the discussion in Section IS.7.4 more towards cumulative risk and away from the one chemical at a time approach, possibly bringing forward material from the appendices. Some of these other risk factors may in fact enhance the effects of Pb.

- In section IS 1.2.4, at the top of page IS-8, the basis for evaluation is laid out in terms that clearly allude to the Bradford-Hill criteria. This is not inappropriate, but some additional text that alludes to more modern thinking about causal inference would be good to add. Perhaps this could be simply an indication regarding consideration of studies' attention to potential biases and how attempts to eliminate any such biases were factored into the evaluation process.

- Figure IS-1 shows the Pb in the soils, but there is no map showing any ambient air quality or the location of the monitors. Some of the figures in the appendix could be brought forward here to at least have one graphic showing the location of the ambient monitors and current concentrations. There is no mention of nonattainment areas; at a minimum a list of the current nonattainment areas should be added to the IS.

- Figure IS-2 is a model figure that illustrates potential health effects pathways at the organismal and population levels that is not relevant to the IS text since the rest of the IS does not include any of these diagrams. Further, the text in this section does not provide much specific detail. If this figure does belong in the IS, then further explanation is needed.

- The summary tables IS-2A to IS-11 provide essential information summarizing evidence from the current and 2013 ISA by outcome. The are easy to navigate along with the associated text.

- Section IS.7.3.2.1 raises the issue of possible residual confounding by age as something that weakens the evidence for adult cognitive effects. This seems overstated as the strong longitudinal studies that have examined this question with measures of cumulative Pb exposure (Pb in bone) have adjusted appropriately for age.

- The discussion of populations at increased risk in IS.7.4 does a good job in synthesizing that information. Some of the "populations" that are described there aren't populations per se. For example, nutrition, stress, other metals, maternal self-esteem, and cognitive reserve are factors that might amplify the effects in particular populations.

- Section IS.7.4.2.1, on blood pressure effects, states that racial differences may reflect the history of greater exposure to Pb among non-Hispanic black populations. This is confusing because it seems to contradict the conclusions presented of consistent findings of greater effect in African Americans. This is a key point that needs to be clarified. Note that a greater effect in one group suggests a steeper dose-response relation *at the same levels of the exposure*. One group having worse effects because they experience higher exposures is a different issue. Race or ethnicity is not a biological issue, and this should be made clear throughout the document.

- Section IS.8 provides a good discussion of the types of endpoints that should be considered and the importance of having sound linkages between observed effects at lower levels of biological complexity and effects at the population/community levels of complexity.

- Section IS.9, "Policy-Relevant Issues," addresses air Pb to blood Pb ratios and concentration response relationships. The summaries here are concise and well done, but it might be useful to have one or more graphs pulled from the appendices into these sections to give the reader a sense of where the data support the discussion of the nature of those relationships.

- In section IS.9.2, "Concentration-Response Relationships for Human Health Effects," the discussion of studies on page IS-101, lines 1-13, indicates "…recent studies generally include somewhat older children, or… [are] designed to answer relatively narrow research questions (e.g., the effect of joint exposure to Pb and other metals, or the effect of concurrent Pb exposure independent from prenatal exposure) and consequently do not have the attributes of the studies on which the conclusion of the 2013 Pb ISA was based (i.e., early childhood BLLs, consideration of peak BLLs, or concurrent BLLs in young children)…[and] recent studies that reflect the lower early childhood Pb exposures now more common in the United States…are generally lacking. Overall, the recently available studies were not designed, and may not have the sensitivity, to detect (Cooper et al., 2016) the effect or hazard at these very low BLLs, nor do they provide evidence of a threshold for effects…." The implication of this statement is that studies of later childhood exposures are not as relevant to characterizing Pb hazards as are previous studies of prenatal and early childhood (~ages 3-4 years) exposures. It is unclear why this literature does not contribute to the health effects assessment and clarification of the rationale would be useful. For example, prenatal or early life are sensitive exposure periods, but Pb has demonstrable adverse effects on cognitive function measures across the life course. Depending on study design, the potential for confounding of later childhood exposure effects by prenatal/early childhood exposures may be a concern, but that is not explicitly discussed.

The IS includes a summary of evidence related to concentration-response relationships for human health effects and the timing of Pb exposure contributing to nervous system effects. To what extent do these sections appropriately synthesize the available evidence? To what extent do the conclusions in these sections adequately reflect the strengths and limitations of the evidence?

To what extent does the IS appropriately synthesize the evidence for populations at increased risk of experiencing effects due to Pb exposures, including consideration of children's health? To what extent do at-risk conclusions adequately reflect the strengths and limitations of the evidence?

The IS includes a summary of evidence related to ecological effects of Pb. To what extent do these sections appropriately synthesize the available evidence for effects observed in terrestrial, freshwater and saltwater organisms?

The CASAC finds that the summaries in the IS adequately reflect the appendices.

References

AVWeb, (2019). Shell announces unleaded 100-octane fuel (Updated). AVWeb. https://www.avweb.com/news/shell-announces-unleaded-100-octane-fuel-updated/

AVWeb, (2021). GAMI awarded long-awaited STC for unleaded 100-octane Avgas. AVWeb. https://www.avweb.com/aviation-news/gami-awarded-long-awaited-stc-for-unleaded-100-octane-avgas/

Besser, J. M., Brumbaugh, W. G., Kemble, N. E., Ivey, C. D., Kunz, J. L., Ingersoll, C. G., & Rudel, D. (2011). Toxicity of nickel-spiked freshwater sediments to benthic invertebrates-spiking methodology, species sensitivity, and nickel bioavailability. U.S. Geological Survey Scientific Investigations Report. https://doi.org/10.3133/sir20115225

Bui, L. T. M., Shadbegian, R., Marquez, A., Klemick, H., & Guignet, D. (2022). Does short-term, airborne lead exposure during pregnancy affect birth outcomes? quasi-experimental evidence from NASCAR's deleading policy. *Environment International*, *166*, 107354. https://doi.org/10.1016/j.envint.2022.107354

CDTSC (2022). LeadSpread-9. California Department of Toxic Substance Control (CDTSC). https://dtsc.ca.gov/leadspread-9/

Church, B. G., Van Sprang, P. A., Chowdhury, M. J., & DeForest, D. K. (2017). Updated species sensitivity distribution evaluations for acute and chronic lead toxicity to saltwater aquatic life. *Environmental Toxicology and Chemistry*, *36*(11), 2974–2980. https://doi.org/10.1002/etc.3863

Cooper, G. S., Lunn, R. M., Ågerstrand, M., Glenn, B. S., Kraft, A. D., Luke, A. M., & Ratcliffe, J. M. (2016). Study sensitivity: Evaluating the ability to detect effects in systematic reviews of chemical exposures. *Environment International*, *92–93*, 605–610. https://doi.org/10.1016/j.envint.2016.03.017

Dartey, E., Berlinger, B., Thomassen, Y., Ellingsen, D. G., Odland, J., Nartey, V. K., Yeboah, F. A., & Weinbruch, S. (2014). Bioaccessibility of lead in airborne particulates from car battery repair work. *Environ. Sci.: Processes Impacts*, *16*(12), 2782–2788. https://doi.org/10.1039/c4em00455h

Escolar, E., Lamas, G. A., Mark, D. B., Boineau, R., Goertz, C., Rosenberg, Y., Nahin, R. L., Ouyang, P., Rozema, T., Magaziner, A., Nahas, R., Lewis, E. F., Lindblad, L., & Lee, K. L. (2014). The effect of an EDTA-based chelation regimen on patients with diabetes mellitus and prior myocardial infarction in the trial to assess chelation therapy (TACT). *Circulation: Cardiovascular Quality and Outcomes*, *7*(1), 15–24. https://doi.org/10.1161/circoutcomes.113.000663

Evens, A., Hryhorczuk, D., Lanphear, B. P., Rankin, K. M., Lewis, D. A., Forst, L., & Rosenberg, D. (2015). The impact of low-level lead toxicity on school performance among children in the Chicago Public Schools: A population-based retrospective cohort study. *Environmental Health*, *14*(1). https://doi.org/10.1186/s12940-015-0008-9

Fleming, D. E., Boulay, D., Richard, N. S., Robin, J. P., Gordon, C. L., Webber, C. E., & Chettle, D. R. (1997). Accumulated body burden and endogenous release of lead in employees of a lead smelter. *Environmental Health Perspectives*, *105*(2), 224–233. https://doi.org/10.1289/ehp.97105224

Fleming, D. E. B., Chettle, D. R., Wetmur, J. G., Desnick, R. J., Robin, J.-P., Boulay, D., Richard, N. S., Gordon, C. L., & Webber, C. E. (1998). Effect of theδ-aminolevulinate dehydratase polymorphism on the accumulation of lead in Bone and blood in lead smelter workers. *Environmental Research*, *77*(1), 49–61. https://doi.org/10.1006/enrs.1997.3818

Hollingsworth, A., & Rudik, I. (2021). The effect of leaded gasoline on elderly mortality: Evidence from regulatory exemptions. *American Economic Journal: Economic Policy*, *13*(3), 345–373. https://doi.org/10.1257/pol.20190654

Hu, H., Téllez-Rojo, M. M., Bellinger, D., Smith, D., Ettinger, A. S., Lamadrid-Figueroa, H., Schwartz, J., Schnaas, L., Mercado-García, A., & Hernández-Avila, M. (2006). Fetal lead exposure at each stage of pregnancy as a predictor of infant mental development. *Environmental Health Perspectives*, *114*(11), 1730–1735. https://doi.org/10.1289/ehp.9067

Lamas, G. A., Boineau, R., Goertz, C., Mark, D. B., Rosenberg, Y., Stylianou, M., Rozema, T., Nahin, R. L., Terry Chappell, L., Lindblad, L., Lewis, E. F., Drisko, J., & Lee, K. L. (2014). EDTA chelation therapy alone and in combination with oral high-dose multivitamins and minerals for coronary disease: The factorial group results of the trial to assess chelation therapy. *American Heart Journal*, *168*(1). https://doi.org/10.1016/j.ahj.2014.02.012

Lamas, G. A., Navas-Acien, A., Mark, D. B., & Lee, K. L. (2016). Heavy metals, cardiovascular disease, and the unexpected benefits of chelation therapy. *Journal of the American College of Cardiology*, *67*(20), 2411–2418. https://doi.org/10.1016/j.jacc.2016.02.066

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., Rothenberg, S. J., Needleman, H. L., Schnaas, L., Wasserman, G., Graziano, J., & Roberts, R. (2005). Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, *113*(7), 894–899. https://doi.org/10.1289/ehp.7688 [erratum: *Environmental Health Perspectives,* 2019, *127*(9): https://doi.org/10.1289/ehp5685]

Lanphear, B. P., Rauch, S., Auinger, P., Allen, R. W., & Hornung, R. W. (2018). Low-level lead exposure and mortality in US adults: A population-based Cohort Study. *The Lancet Public Health*, *3*(4). https://doi.org/10.1016/s2468-2667(18)30025-2

Leggett, R. W. (1993). An age-specific kinetic model of lead metabolism in humans. *Environmental Health Perspectives*, *101*(7), 598–616. https://doi.org/10.1289/ehp.93101598

Lin, J.-L. (1999). Chelation therapy for patients with elevated body lead burden and progressive renal insufficiency. *Annals of Internal Medicine*, *130*(1), 7. https://doi.org/10.7326/0003-4819-130-1-199901050-00003

Lin, J.-L., Lin-Tan, D.-T., Hsu, K.-H., & Yu, C.-C. (2003). Environmental lead exposure and progression of chronic renal diseases in patients without diabetes. *New England Journal of Medicine*, *348*(4), 277–286. https://doi.org/10.1056/nejmoa021672

Meng, Q., Richmond-Bryant, J., Davis, J. A., Cohen, J., Svendsgaard, D., Brown, J. S., Tuttle, L., Hubbard, H., Rice, J., Vinikoor-Imler, L., Sacks, J. D., Kirrane, E., Kotchmar, D., Hines, E., & Ross, M. (2014). Contribution of particle-size-fractionated airborne lead to blood lead during the National Health and Nutrition Examination Survey, 1999–2008. *Environmental Science & Technology*, *48*(2), 1263–1270. https://doi.org/10.1021/es4039825

Menke, A., Muntner, P., Batuman, V., Silbergeld, E. K., & Guallar, E. (2006). Blood lead below 0.48 μmol/L (10 μg/DL) and mortality among US adults. *Circulation*, *114*(13), 1388–1394. https://doi.org/10.1161/circulationaha.106.628321

Park, S. K., Mukherjee, B., Xia, X., Sparrow, D., Weisskopf, M. G., Nie, H., & Hu, H. (2009). Bone lead level prediction models and their application to examine the relationship of lead exposure and hypertension in the Third national health and nutrition examination survey. *Journal of Occupational & Environmental Medicine*, *51*(12), 1422–1436. https://doi.org/10.1097/jom.0b013e3181bf6c8d

Paul, K. C., Horvath, S., Del Rosario, I., Bronstein, J. M., & Ritz, B. (2021). DNA methylation biomarker for cumulative lead exposure is associated with parkinson's disease. *Clinical Epigenetics*, *13*(1). https://doi.org/10.1186/s13148-021-01051-3

Pierre, F., Vallayer, C., Baruthio, F., Peltier, A., Pale, S., Rouyer, J., Goutet, P., Aubrège, B., Lecossois, C., Guillemin, C., Elcabache, J.-M., Verelle, B., & Fabriès, J.-F. (2002). Specific relationship between blood lead and air lead in the Crystal Industry. *International Archives of Occupational and Environmental Health*, *75*(4), 217–223. https://doi.org/10.1007/s00420-001-0303-3

Richmond-Bryant, J., Meng, Q., Davis, A., Cohen, J., Lu, S.-E., Svendsgaard, D., Brown, J. S., Tuttle, L., Hubbard, H., Rice, J., Kirrane, E., Vinikoor-Imler, L. C., Kotchmar, D., Hines, E. P., & Ross, M. (2014). The influence of declining air lead levels on blood lead–air lead slope factors in children. *Environmental Health Perspectives*, *122*(7), 754–760. https://doi.org/10.1289/ehp.1307072

Ruiz-Hernandez, A., Navas-Acien, A., Pastor-Barriuso, R., Crainiceanu, C. M., Redon, J., Guallar, E., & Tellez-Plaza, M. (2017). Declining exposures to lead and cadmium contribute to explaining the reduction of cardiovascular mortality in the US population, 1988–2004. *International Journal of Epidemiology*, *46*(6), 1903–1912. https://doi.org/10.1093/ije/dyx176

Shaffer, R. M., Forsyth, J. E., Ferraro, G., Till, C., Carlson, L. M., Hester, K., Haddock, A., Strawbridge, J., Lanfear, C. C., Hu, H., & Kirrane, E. (2022). Lead exposure and antisocial behavior: A systematic review protocol. *Environment International*, *168*, 107438. https://doi.org/10.1016/j.envint.2022.107438

SwiftFuels, (2022). Frequently asked questions. Swift Fuels, West Lafayette, IN. https://www.swiftfuelsavgas.com/faq

U.S.EPA, (2021a). 2017 National Emissions Inventory: January 2021 Updated Release, Technical Support Document. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/sites/default/files/2021-02/documents/nei2017_tsd_full_jan2021.pdf

U.S.EPA, (2021b). Advancing Pb Exposure and Biokinetic Modeling for U.S. EPA Regulatory Decisions and Site Assessments using Bunker Hill Mining and Metallurgical Complex Superfund Site Data. Office of Research and Development. EPA 600/R-21/017F. April.

U.S.EPA, (2022). EPA proposes endangerment finding for lead emissions from aircraft engines that operate on leaded fuel. U.S. Environmental Protection Agency, Washington DC. https://www.epa.gov/newsreleases/epa-proposes-endangerment-finding-lead-emissions-aircraft-engines-operate-leaded-fuel

U.S.EPA, (2023a). Overview of lead (Pb) air quality in the United States. U.S. Environmental Protection Agency, Washington DC. https://www.epa.gov/system/files/documents/2022-08/Pb_2021_0.pdf

U.S.EPA, (2023b). 2020 National Emissions Inventory (NEI) data. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-data

U.S.EPA, (2023c). Toxics Release Inventory (TRI) program. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/toxics-release-inventory-tri-program

U.S.EPA, (2023d). Nonattainment areas for criteria pollutants (Green Book). U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/green-book

Vork, K. L., Brown, J. P., & Carlisle, J. C. (2023). Evaluation and updates to the Leggett model for pharmacokinetic modeling of exposure to lead in the workplace – part II adjustments to the adult exposure model, confirmation of leggett+, and modeling of workplace exposure. *Journal of Occupational and Environmental Hygiene*, *20*(2), 55–83. https://doi.org/10.1080/15459624.2022.2150767

Wang, X., Bakulski, K. M., Mukherjee, B., Hu, H., & Park, S. K. (2023). Predicting cumulative lead (PB) exposure using the Super Learner algorithm. *Chemosphere*, *311*, 137125. https://doi.org/10.1016/j.chemosphere.2022.137125

Wasserman, G. A., Factor-Litvak, P., Liu, X., Todd, A. C., Kline, J. K., Slavkovich, V., Popovac, D., & Graziano, J. H. (2003). The relationship between Blood Lead, Bone Lead and child intelligence. *Child Neuropsychology*, *9*(1), 22–34. https://doi.org/10.1076/chin.9.1.22.14497

Weisskopf, M. G., Jain, N., Nie, H., Sparrow, D., Vokonas, P., Schwartz, J., & Hu, H. (2009). A prospective study of bone lead concentration and death from all causes, cardiovascular diseases, and cancer in the Department of Veterans Affairs Normative Aging Study. *Circulation*, *120*(12), 1056–1064. https://doi.org/10.1161/circulationaha.108.827121

Weisskopf, M. G., Sparrow, D., Hu, H., & Power, M. C. (2015). Biased exposure–health effect estimates from selection in Cohort Studies: Are Environmental Studies at particular risk? *Environmental Health Perspectives*, *123*(11), 1113–1122. https://doi.org/10.1289/ehp.1408888

Zahran, S., Keyes, C., & Lanphear, B. (2023). Leaded aviation gasoline exposure risk and child blood lead levels. *PNAS Nexus*, *2*(1). https://doi.org/10.1093/pnasnexus/pgac285

Appendix A

Individual Comments by the 2021 CASAC Lead Review Panel Members on the EPA's *Integrated Science Assessment (ISA) for Lead (External Review Draft – March 2023)*

Mr. George A. Allen **A-2**
Dr. James Boylan **A-8**
Dr. Judith Chow **A-10**
Dr. Deborah Cory-Slechta **A-16**
Dr. Christina H. Fuller **A-18**
Dr. Philip. E. Goodrum **A-20**
Mr. Perry Gottesfeld **A-24**
Dr. Daven Henze **A-30**
Dr. Howard Hu **A-32**
Dr. Chris E. Johnson **A-34**
Dr. Susan Korrick **A-37**
Dr. Bruce Lanphear **A-49**
Dr. Joel G. Pounds **A-55**
Dr. Brisa Sánchez **A-57**
Dr. Brian Schwartz **A-60**
Dr. Elizabeth A. (Lianne) Sheppard **A-71**
Dr. William Stubblefield **A-73**
Dr. Kathleen Vork **A-78**
Dr. Marc Weisskopf **A-81**

Mr. George A. Allen

Appendix 1 - Lead Source to Concentration

Section 1.2.3, page 1-14 cites a paper from the 2013 ISA for Pb emission factors from fireplaces and woodstoves. This cite is a very limited source (using European wood) and could be supplemented with data from a more robust NYSERDA/NESCAUM study of Pb woodstove emission factors that will be available this September, presented at the 2023 International Emissions Inventory Conference in Seattle. For communities with elevated winter woodsmoke PM, this could be the largest source of Pb exposure.

Section 1.4.3, Recent Advances in Sampling and Analysis, discusses recent ORD research on the existing HiVol FRM sampler and potential low-volume alternatives for TSP measurement. This work is very useful in understanding the size characteristics of these methods for particles larger than a few microns. The HiVol FRM sampler is now more than 80 years old (first developed during WW II). If it were proposed as a FRM sampler now, it would never be approved. While it is desirable to sample as much of the > 10 μm particles as possible from mechanically generated sources, industrial sources with large particles are now very few, making it harder to justify use of this antique method. Sampling particles larger than about 15 μm in ambient air is not practical. As reported in the Vanderpool et al. AS&T 2017 paper, the louvered inlet used in the PM10 low-vol FRM is a reasonable alternative to the HiVol FRM sampler; neither is ideal at 15 μm but both are good enough, and the low-vol TSP sampler is dramatically easier to run.

Appendix 12 (Process for ISA Development)

Charge Question 5a. Please comment on the clarity with which Appendix 12 communicates the process undertaken to develop the Draft Pb ISA.

This appendix clearly outlines the overall approach used for ISA development. Fig 2-1, a diagram of the general process for ISA development, is a helpful overview. The first two sections (literature search and evaluation of study quality) are critical steps given the mountain of literature on Pb that has to be parsed for this review, and if they are not done well the review document will not be adequate for its intended purpose. These two steps make up all but 4 pages of this appendix, with relatively brief mentions of the other process steps which include “Evaluation, Synthesis, and Integration of Evidence” and “Development of Scientific Conclusions and Causal Determinations” which are also important steps.

Section 12.4, relevance and scope, are useful, noting that we now have three different study inclusion criteria: Population, Intervention, Comparison, Outcome, and Context (PICOC -- “peacock”) for environmental research [atmospheric science], PECOS for health studies, and Level of Biological Organization, Exposure, Comparison, Endpoint, and Study design (LECES) for welfare effects.

For 12.4.2. [Exposure, Toxicokinetics, and Biomarkers] there is some concern about geographic inclusion criteria (as CASAC has noted previously) for exposure potentially being too strict: “Exposure studies pertaining to the U.S. population and U.S.-based Pb sources were preferred. Studies were included from outside the United States if these studies were judged to have important findings, with a focus on studies from Canada, western Europe, and Australia (i.e., areas with study populations and air quality characteristics most similar to the United States).”

Same for PECOS used in 12.4.3 (health effects), and possibly for PICOC and LECES also.

Executive Summary

Charge Question 6a. Please comment on the clarity with which the ES communicates the key information from the Draft Pb ISA.

The ISA executive summary is very well written, and is accessible to the casual reader or those who want a quick high-level overview of the ISA findings. Context for findings based on the 2013 ISA versus studies since that ISA closed in 2011 is provided, and references to additional detail in the Integrated Synthesis (IS) are given. Color coded tables for Causality Determinations based on tables in the IS are presented for both health and welfare effects, and important changes in categories are noted for both renal and total non-accidental mortality effects, now both causal. It should be noted here, as it is in the IS summary of causal determination (Table IS-1), that the cardiovascular effects category combines 4 separate outcomes used in the 2013 ISA. Figure ES-1 is a diagram of the conceptual model of multimedia Pb exposure pathways, but the version in the REA presentation of June 14 on page 7 is easier to follow. Section ES.5 clearly explains the important point that "These exposures are considered air-related if Pb passed through the air compartment at any point prior to plant, animal, or human contact." The summaries of 2017 NEI data presented here will need to be updated to use 2020 NEI data, now available, that now include wildfire Pb emissions (3% of total, the 4th highest category). Another important finding is summarized in section ES 6.1.1, that "recently available evidence does not provide evidence of a threshold for the observed neurodevelopmental effects across the range of blood Pb levels examined (ES.7.1.3)." Section ES-7, Key Aspects of Health and Welfare Effects Evidence, is a good summary of conclusions on key policy-relevant topics in this ISA.

Charge Question 6b. Please provide recommendations on any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

I do not have any recommendations for any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

Integrated Synthesis

7a. *Please comment on the usefulness and effectiveness of the summary presentation in the IS and provide any recommendations or alternate text that may improve the synthesis of available information across subject areas and the communication of key findings.*

Per EPA's ISA Charge Letter, the Executive Summary and Integrated Synthesis serve as the main body of the ISA. The Integrated Synthesis is "a more detailed synthesis of the ISA's most policy-relevant scientific findings". It is well written and organized, and at about 100 pages has more detail than the Executive Summary, but less than appendices. The IS mostly references ISA Appendices instead of individual literature cites, which is appropriate. For each causal category, each section ends with a very useful comparison between evidence from the 2013 ISA and this ISA that details what evidence is new in this review. Overall I find it a very useful and effective approach to present the detailed findings of the ISA.

The overall conclusions page, IS-1, has total mortality under the likely causal category.

Table IS-1 groups total mortality under "Effects on Other Organ Systems" (consistent with IS 1.2.2 "Organization of the ISA", page IS-7), while Figure ES-2 lists it as a major separate category after the "Effects on Other Organ Systems:

Table ES-2:

Respiratory Effects
Total (Non-Accidental) Mortality
Total Mortality +
Cancer

Is this a typo in Table ES-2 -- is it meant to be as listed in Table IS-1 (under "Effects on Other Organ Systems")?

Table IS-1:

Effects on Other Organ Systems and Mortality	Endocrine system effects	Inadequate
	Musculoskeletal effects	Likely to be causal
	Effects on ocular health	Inadequate
	Respiratory effects	Inadequate
	Total (nonaccidental) mortality	+Causal
Cancer	Cancer	Likely to be causal

The new category (marked with "+") of total nonaccidental mortality may be confusing to the reader. It is new only in the sense that this category consolidates previously separate categories, primarily (but not solely) cardiovascular mortality, as explained in Section IS 7.3.9, Mortality: "… the 2013 Pb ISA evaluated studies of all-cause mortality together with studies examining cardiovascular mortality, and these studies were all included within the CVD chapter." This change in how total mortality is presented in these two tables needs to be noted in the table footnotes, as is done for cardiovascular mortality (footnote a) in table IS-1.

ISA Appendix 1. Lead Source to Concentration.

Section 1.2.3, page 1-14 cites a paper from the 2013 ISA for Pb emission factors from fireplaces and woodstoves. This cite is a very limited source (using European wood) and could be supplemented with data from a more robust NYSERDA/NESCAUM study of Pb woodstove emission factors that will be available this September, to presented at the 2023 International Emissions Inventory Conference in Seattle. For communities with elevated winter woodsmoke PM, this could be the largest source of Pb exposure.

Section 1.4.3, Recent Advances in Sampling and Analysis, discusses recent ORD research on the existing HiVol FRM sampler and potential low-volume alternatives for TSP measurement. This work is very useful in understanding the size characteristics of these methods for particles larger than a few microns. The HiVol FRM sampler is now more than 80 years old (first developed during WW II). If it were proposed as a FRM sampler now, it would never be approved. While it is desirable to sample as much of the > 10 µm particles as possible to capture Pb from mechanically generated sources and re-entrained soil, industrial sources with large particles are now very few, making it harder to justify use of this antique method. Sampling particles larger than about 15 µm in ambient air is not practical. As reported in the Vanderpool et al. AS&T 2017 paper (https://www.tandfonline.com/doi/full/10.1080/02786826.2017.1386766), the louvered inlet used in the PM10 low-vol FRM is a reasonable alternative to the HiVol FRM sampler; neither is ideal at 15 µm but both are good enough, and the low-vol TSP sampler is dramatically easier and less expensive to maintain and run in state and local monitoring networks.

To expand the database of ambient Pb concentrations at levels well below the current NAAQS, this CASAC ISA review recommends that EPA analyze filters from all available PM2.5 and PM10 FRM samplers for Pb using XRF, a sensitive and relatively inexpensive analytical method that provides data of similar quality for Pb to the much more expensive ICPMS method. This would augment the limited XRF Pb data from the Chemical Speciation Network (CSN), a non-source oriented monitoring network operated by state and local air agencies at about 150 sites nationwide. The National Air Toxics Trends Station (NATTS) Network has an additional 24 sites with Pb analysis by ICPMS. However, there appear to be problems with XRF Pb analysis as presently performed for CSN sites which would impact the usefulness of an expanded analysis program. EPA contracts with laboratories for this analysis. In October 2016, the contract laboratory changed, resulting in dramatically degraded data quality. The time-series plots below show 3 examples of Pb data before and after the change in the XRF analytical laboratory, as well as some collocated data from ICPMS analysis for comparison. It is clear that XRF Pb data from the new laboratory is of much lower quality than the old laboratory, which makes analysis of Pb trends at these sites almost impossible. EPA needs to provide more QA oversight of this contract laboratory to ensure that data are of sufficient quality for the intended purposes.

Figure 1. Roxbury/Boston Pb, µg/m^3, AQS 25-025-0042. The blue line is collocated ICPMS data for comparison.

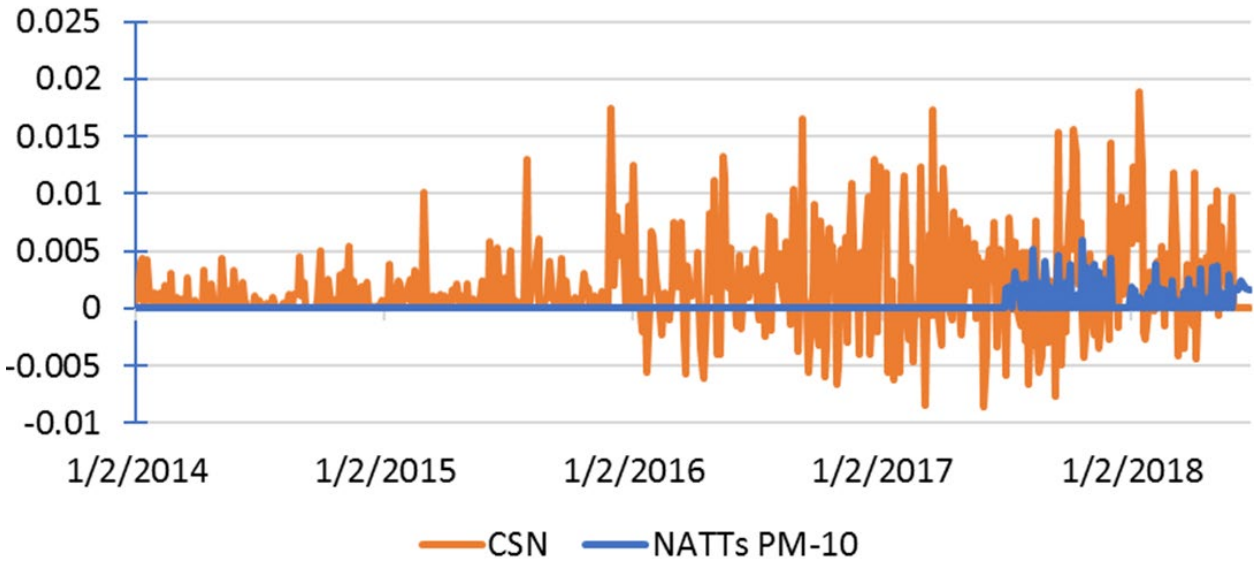


Figure 2. Rochester NY CSN Pb, µg/m^3.

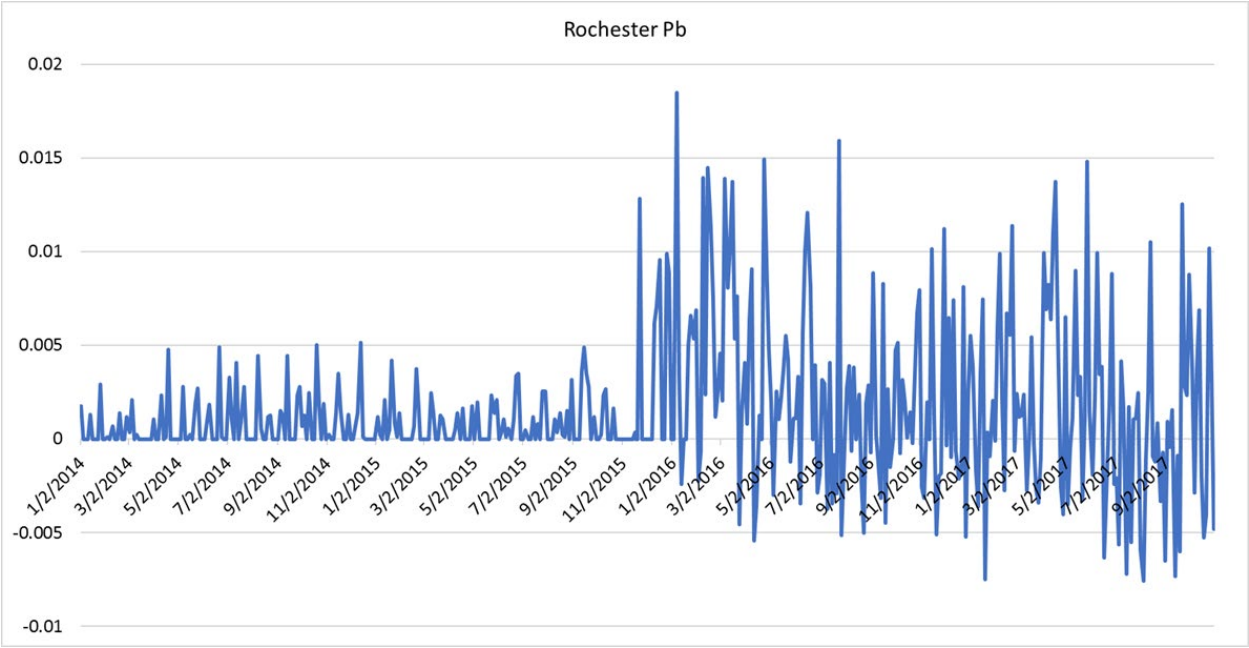
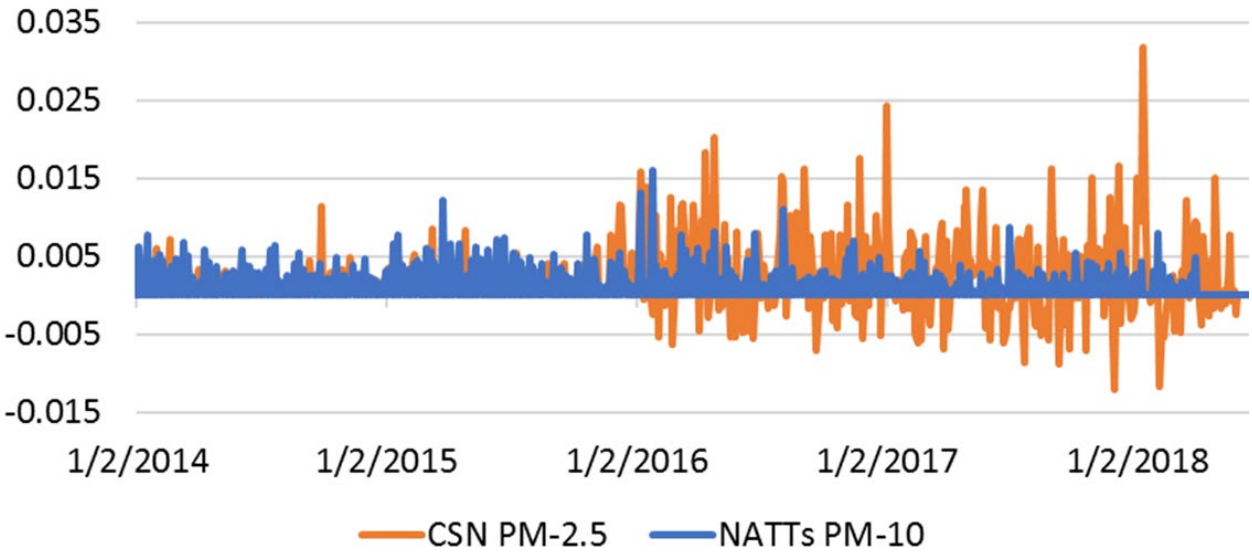


Figure 3. Bronx (NYC) Pb, µg/m^3, AQS 36-005-0010. The blue line is collocated ICPMS data for comparison.



Dr. James Boylan

Source to Concentration

1a. To what extent are Pb sources and emissions appropriately captured and described in Section 1.2? Are the relative importance and uncertainties in comparing contemporary vs. legacy Pb sources adequately explained?

Section 1.2 covers all the major sources of contemporary and legacy Pb sources and does a good job of discussing the relative importance of these sources; however, the uncertainties associated with these sources were not discussed in detail. It would be helpful to add additional details on the uncertainties associated with each source and provide an overall summary of the relative uncertainties across sources. It would be helpful to include Figures 1, 2, and 3 and the associated text from EPA's separate document "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)".

1b. To what extent are Pb fate and transport adequately covered in Section 1.3 with appropriate detail and balance across media?

Section 1.3 adequately covered the fate and transport of Pb in air, soil, water, and urban media.

1c. To what extent are recent advances in the development of measurement methods and method performance adequately covered in Section 1.4?

Section 1.4 gives a good overview of the Federal Reference Methods currently approved for Pb sampling. This section adequately covered recent advances in the development of measurement methods and method performance. It would be helpful to include Figures 4 and 5 and the associated text from EPA's separate document "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)". Also, it would be good to include this sentence in the ISA from the same document "At a minimum, there must be one source-oriented site located to measure the maximum Pb concentration in ambient air resulting from each non-airport Pb source estimated to emit 0.5 or more tons of Pb per year and from each airport estimated to emit 1.0 or more tons of Pb per year." Finally, it would be good to state the criteria that must be met for a lead monitor to be shutdown (i.e., the design value is less than 50% of the NAAQS).

1d. To what extent do presentation of recent concentration trends and size distribution observations in Section 1.5 adequately reflect the recent literature?

Section 1.5 gives a high-level overview of the recent concentration trends. It would be helpful to include Tables 1 and 2 and Figures 6, 7, 8, and 9 and the associated text from EPA's separate document "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)".

1e. The topics covered in Appendix 1 were selected to provide useful context ranging from sources to concentrations of Pb relevant to understanding health and ecological effects covered in the ISA. Please identify any missing or incomplete areas of research and provide specific additional studies which would fill any gaps in this appendix. Are any research areas overemphasized and is there specific text that could be reduced in length to remedy this?

I like that EPA has created a stand-alone document titled "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)". This document is outside the ISA and PA and can be updated as new data becomes available. It allows ORD and OAQPS to work off the same data, figures, and tables. However, I think it would be helpful to include the current version of the figures and tables from this document in the ISA and PA. This will make it easy for the reviewer to see this important information and it will document the data that was available at the time the ISA and PA are released to the public. A sentence could be added to the ISA and PA letting the reader know that newer data and updated figures and tables may be available in the stand-alone document.

This Appendix is missing a discussion on areas that are currently in nonattainment for the current lead NAAQS. EPA should provide a list of current nonattainment areas and a map with their locations. Also, EPA should explain why the list of current nonattainment areas looks very different than the map with 2019-2021 lead DVs over 0.15 $\mu g/m^3$ as shown in Figure 6 from "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)".

Integrated Synthesis

Page IS-1: The fourth bullet under "Human Health Effects" states, "Recent experimental and epidemiologic evidence supports *likely to be causal relationships* between Pb exposure and conduct disorders in children and young adults, internalizing behaviors in children and adolescents, motor function decrements in children, cognitive function decrements in adults, neurodegenerative diseases, immunosuppression, musculoskeletal effects, total (nonaccidental) mortality, and cancer." However, "total (nonaccidental) mortality" should be removed from this list and moved to another bullet since the draft ISA concludes a *causal relationship* with Pb exposure.

Page IS-12: Section IS.2.4 should include Figure 6 from EPA's separate document "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)". Also, this section should include a list of current nonattainment areas with their locations. EPA should explain why the list of current nonattainment areas looks very different than the map with 2019-2021 lead DVs over 0.15 $\mu g/m^3$ as shown in Figure 6 from "Overview of Lead (Pb) Air Quality in the United States (U.S. EPA, 2022)".

Dr. Judith Chow

Appendix 1: Lead Source to Concentration

Appendix 1 is well written and documents lead (Pb) emission sources and transport of Pb in multimedia (e.g., air, water, soil, sediment, and diet) that are most relevant to environmental concentrations of Pb. The following categories merit additional clarification.

1a. To what extent are Pb sources and emissions appropriately captured and described in Section 1.2? Are the relative importance and uncertainties in comparing contemporary vs. legacy Pb sources adequately explained?

To facilitate provision of the most current emissions and ambient concentrations, Appendix 1 refers to data presented in "Overview of Lead (Pb) Air Quality in United States" (U.S.EPA, 2023a). Figure 1 of this document shows large reduction of Pb emissions from 1990 to 2020. The "Non-Road Mobile" and "Highway Vehicles" categories in Figure 1 are entitled "Mobile-Aircraft" sector in Figure 2 for U.S. lead emissions by sector based on the 2017 National Emission Inventory (NEI, U.S.EPA, 2021). Consistent source sector categories should be used for cross comparisons.

Section 1.2.3 of Appendix 1 on "Fuel Combustion" notes that biomass combustion accounts for 16% of fuel combustion (~45 tons of Pb/year) or 7.2 tons/year (Lines 23-25, Page 1-13). This is ~60% less than the ~18 tons/year for wildfire emissions estimated by the 2020 NEI (U.S.EPA, 2020, 2023b). It is not clear if wildfire as well as wildfire burn infrastructure (e.g., anthropogenic structures and vehicles burned that mobilized Pb) are excluded from the biomass combustion category in the 2020 NEI. Definitions of "biomass combustion" and "fires" need to be explained. References should be given regarding the higher Pb emissions from smoldering compared to flaming phases (Lines 31-34, Page 1-14), as prolonged smoldering smoke exposure has been related to adverse respiratory effects.

Section 1.2.4 "Fires" summarizes several studies that have found elevated PM Pb concentrations in air, ash, and soil. As PM Pb concentrations varied between days with and without fire emissions, it will be helpful to tabulate these fire events (e.g., location, duration, acreage), document Pb concentrations, and compare the ratio of Pb levels before, during, and after fires to evaluate the environmental impacts of fires.

1b. To what extent are Pb fate and transport adequately covered in Section 1.3 with appropriate detail and balance across media?

Section 1.3.2 "Fate and Transport in Soil" documents Pb transport into and within soil columns that reveal Pb enrichment in different soil horizons. In most soil classification systems, horizons are associated with soil types. It will be helpful to add a footnote to explain geology horizons that define soil properties and depths from the surface soils in O, A, B, and C horizons. Section 1.3.2.4 "Summary" notes that "Once deposited onto soil, Pb is strongly retained in organic surface horizons with subsequent Pb retentions and distribution in soil…" (Lines 23-27, Page 1-30); does this refer to the O horizon for surface and the A horizon up to 10 inches from the surface? Unless the soil is remobilized, Pb particles in the A horizon are unlikely to become airborne. This needs to be clarified, especially for Pb levels at upper depths.

Section 1.3.3 “Fate and Transport in Water and Sediment” (Page 1-31) summarizes Pb levels in surface water, lake, and sediments at urban and non-urban environments. Different concentration units (e.g., mg Pb/kg, µg Pb/L, and pM) are given, making it difficult to follow. Dissolved Pb of 90-920 pM in ice, 88-550 pM in snow, and 30-120 pM in seawater are reported (Lines 24-36, Page 1-40); are these in units of picomole (one trillionth of a mole)? Relevant upper bound Pb levels in different media need to be provided as a guideline to evaluate the extent of Pb contamination.

To give some perspective on levels of severity, it might be helpful to discuss the U.S. guidelines for aquatic system such as 15 µg/L as EPA’s action level for Pb in water delivered to public drinking systems and the 1000 µg/L threshold given for aquatic toxicity.

1c. To what extent are recent advances in the development of measurement methods and method performance adequately covered in Section 1.4?

Section 1.4 “Monitoring of Pb in Ambient Air” (Page 1-53) notes that the 2008 Pb NAAQ revisions “… described conditions under which a Pb-PM_{10} FRM could be used as an alternative to the original Pb-TSP FRM.” (Lines 11-12, Page 1-56). However, not much effort has been made to conduct collocated comparisons between TSP and PM_{10} Pb.

For high-volume TSP, the cut point for 50% sampling efficiency varies from 30-60 µm, with lower efficiencies at the high end of the Pb particle size distribution (Chow, 1995; Watson et al., 1983). Vanderpool et al. (2018) found that none of the six commercially available low-volume (4 to 16.7 L/Min) sampling inlets tested in a wind tunnel produced ideal sampling performance. Collection efficiencies varied from 34 to 109% at 2, 8, and 24 km/hr wind speeds. This further emphasizes the importance of size-selective PM inlets to obtain representative PM mass and Pb concentrations.

1d. To what extent do presentation of recent concentration trends and size distribution observations in Section 1.5 adequately reflect the recent literature?

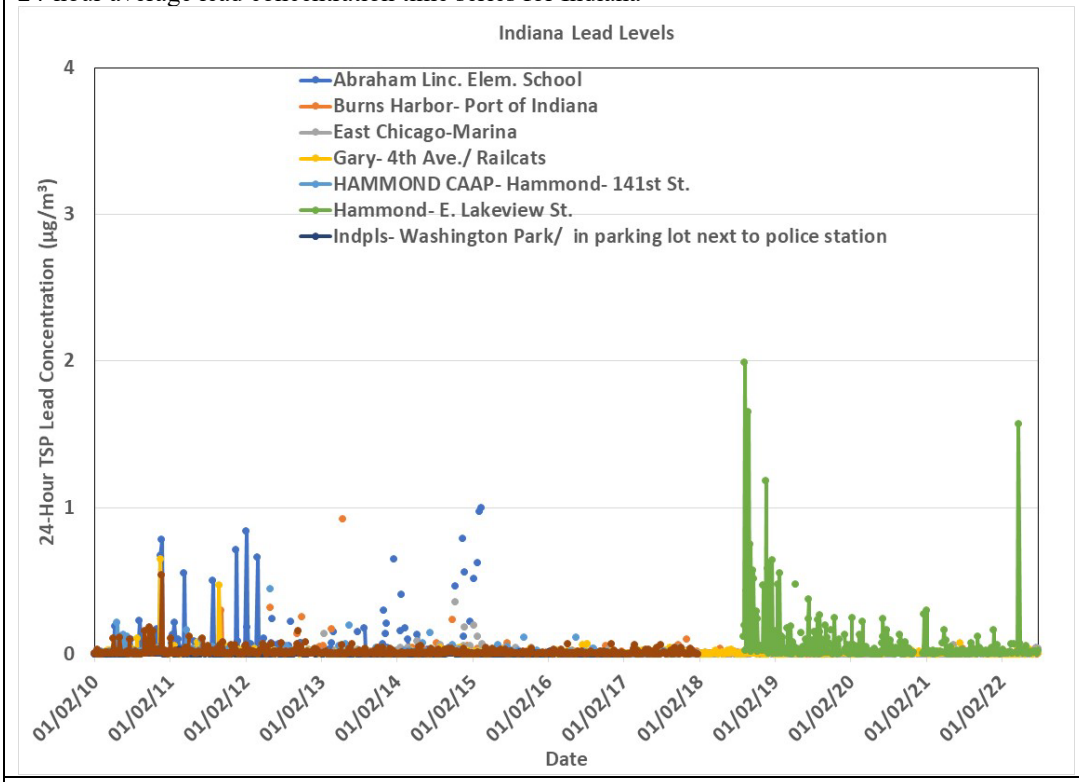
Table 1-1 in Section 1.5 “Ambient Pb concentration Trends” (page 1-58) assembles studies with seasonal variations in ambient Pb concentrations. The nine studies summarized in Table 1-1 found low levels of seasonal average Pb concentration from 0.2 to 42 ng/m^3, far below the NAAQS. Some illustrations on diurnal, seasonal, or annual trend of Pb concentration would be informative.

As of March 31, 2023, thirteen areas in six states (i.e., AZ, CA, KS, MN, MO, and PA) and Puerto Rico are designated as nonattainment for the Pb NAAQS (U.S.EPA, 2023c). These do not correspond to the five sites with Design Values exceeding NAAQS shown in Figure 6 of U.S.EPA (2023a). Temporal and spatial distributions of Pb concentrations at areas with Pb Design Values exceeding the current NAAQS need to be examined to better understand the community exposures.

Nationwide, there are ~164 TSP and PM_{10} monitoring sites as shown in Figure 4 (Page 4 of U.S.EPA, 2023a). Figure 7 (U.S.EPA, 2023a) shows that 28 out of 37 sites reported decreasing trends in Pb Design Values for the period of 2010 through 2021. Although ambient concentrations show large reductions in Pb concentrations at the national-scale, this does not represent the ambient Pb concentrations at individual sites. Downloading AQS data from the website, Figure A below shows frequent Pb concentration spikes in recent years at source-dominant environments in Indiana and Ohio with 24 hour average Pb concentration up to 3 µg/m^3, 20 times higher than the three-month average 0.15

$\mu g/m^3$ Pb NAAQS. Additional neighborhood monitoring at these communities is warranted to ensure safe exposure levels.

24-hour average lead concentration time series for Indiana



24-hour average lead concentration time series for Ohio

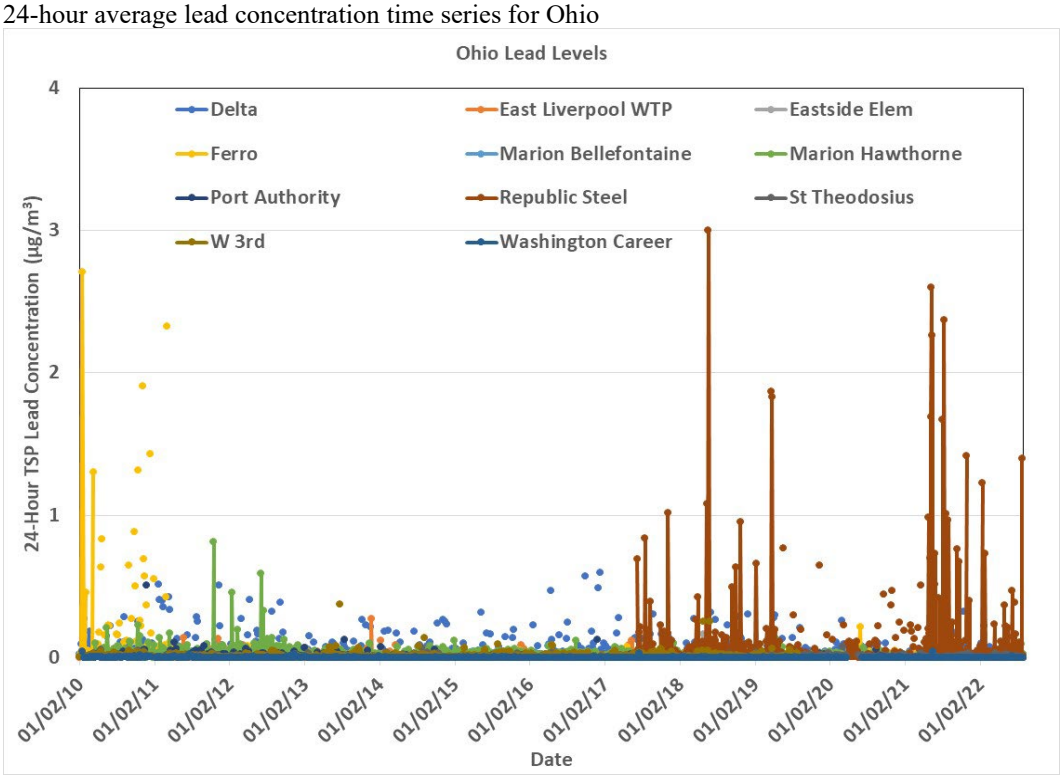


Figure A. Time series of 24-hour TSP lead concentrations in Indiana and Ohio (U.S.EPA, 2023d). The Hammond-E. Lakeview site in Indiana began monitoring on 08/3/2018 and the Republic Steel site in Ohio began monitoring on 6/6/2017.

1e. The topics covered in Appendix 1 were selected to provide useful context ranging from sources to concentrations of Pb relevant to understanding health and ecological effects covered in the ISA. Please identify any missing or incomplete areas of research and provide specific additional studies which would

fill any gaps in this appendix. Are any research areas overemphasized and is there specific text that could be reduced in length to remedy this?

To better understand the extent to which various sources contribute to environmental Pb concentrations at or near the industrial sites, lead-source attribution studies or modeling warrants additional research. Pb isotopic analysis has been found useful to distinguish lead sources (Wang et al., 2015). Research by the National Research Council (Bouwer et al., 2017) recommends the characterization of particle size distribution in ambient air and emission sources to better understand particle mobility and transport at locations upwind and downwind of emission source (Watson et al., 2010) to evaluate source contributions.

References

Bouwer, E.J., Barton, M.D., Betterton, E.A., Blum, J.D., Brantley, S.L., Chow, J.C., Fendorf, S.E., Hazen, R.E., Johnson, C.E., Manton, W.I., Miller, J.R., O'Day, P.A., Toll, J., White, W.H., Mantus, E.K., Wassel, R., Karalic-Loncarevic, M., Rose-Crawford, R., Dawson, T., (2017). Consensus study report: Investigative strategies for lead-source attribution at superfund sites associated with mining activities. The National Academies Press, Washington, DC. https://www.nap.edu/download/24898

Chow, J.C., (1995). Critical review: Measurement methods to determine compliance with ambient air quality standards for suspended particles. Journal of the Air & Waste Management Association, 45, 320-382. http://www.tandfonline.com/doi/pdf/10.1080/10473289.1995.10467369

U.S.EPA, (2020). 2020 NEI plan August 2020. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/sites/default/files/2020-08/documents/2020_nei_plan_final.pdf

U.S.EPA, (2021). 2017 National Emissions Inventory: January 2021 Updated Release, Technical Support Document. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/sites/default/files/2021-02/documents/nei2017_tsd_full_jan2021.pdf

U.S.EPA, (2023a). Overview of lead (Pb) air quality in the United States. U.S. Environmental Protection Agency, Washington DC. https://www.epa.gov/system/files/documents/2022-08/Pb_2021_0.pdf

U.S.EPA, (2023b). 2020 National Emissions Inventory (NEI) data. U.S. Environmental Protection Agency, Research Triangle Park, NC. https://www.epa.gov/air-emissions-inventories/2020-national-emissions-inventory-nei-data

U.S.EPA, (2023c). Lead (2008) designated area/state information. U.S. Environmental Protection Agency, Washington DC. https://www3.epa.gov/airquality/greenbook/mbtc.html

U.S.EPA, (2023d). Pre-generated data files. U.S. Envirionmental Protection Agency, Research Triangle Park, NC. https://aqs.epa.gov/aqsweb/airdata/download_files.html

Vanderpool, R.W., Krug, J.D., Kaushik, S., Gilberry, J., Dart, A., Witherspoon, C.L., (2018). Size-selective sampling performance of six low-volume "total" suspended particulate (TSP) inlets. Aerosol Science and Technology, 52, 98-113. 10.1080/02786826.2017.1386766.

Wang, X.L., Chow, J.C., Kohl, S.D., Yatavelli, R.L.N., Percy, K.E., Legge, A.H., Watson, J.G., (2015). Wind erosion potential for fugitive dust sources in the Athabasca Oil Sands Region. Aeolian Research, 18, 121-134. https://www.sciencedirect.com/science/article/abs/pii/S1875963715000658

Watson, J.G., Chow, J.C., Shah, J.J., Pace, T.G., (1983). The effect of sampling inlets on the PM_{10} and PM_{15} to TSP concentration ratios. Journal of the Air Pollution Control Association, 33, 114-119. http://www.tandfonline.com/doi/pdf/10.1080/00022470.1983.10465552

Watson, J.G., Chow, J.C., Chen, L.-W.A., Wang, X.L., (2010). Measurement system evaluation for fugitive dust emissions detection and quantification. Desert Research Institute, Reno, NV. https://www.researchgate.net/publication/235341860_Measurement_system_evaluation_for_fugitive_dust_emissions_detection_and_quantification?ev=prf_pub

Appendix 12: The Process for Developing the Pb Integrated Science Assessment

5a. Please comment on the clarity with which Appendix 12 communicates the process undertaken to develop the Draft Pb ISA.

Appendix 12 documents the process to develop the Integrated Science Assessment (ISA) including literature search and study selection, as well as synthetization, integration, and evaluation of evidence for developing scientific conclusions and causality determination.

EPA establishes a good framework to define the parameters and formulate PICOC (Population, Intervention, Comparison, Outcome, and Context) for identifying relevant atmospheric science studies; PECOS (Population, Exposure, Comparison, Outcome, and Study design) for experimental and epidemiologic studies; and LECES (Level of Biological Organization, Exposure, Comparison, Endpoint, and Study design) for ecological studies. However, the established criteria limit the scope of work and might omit useful research that does not fall into the predetermined categories.

As the ISA provides the scientific foundation for NAAQS review, a comprehensive literature search is needed to provide an unbiased review. Section 12.4.3.2 "Epidemiologic Studies" states that "… review articles (which typically present summaries or interpretations of existing studies rather than bring forward new information in the form of original research or new analysis…" (Lines 16-18, Page 12-8) are considered to fall outside the scope and were excluded in the ISA. Review articles synthesize a broad range of past studies, and should be considered as a starting point for further literature review. The footnote in Table 12-3 (Page 12-9) notes that studies that estimate Pb concentrations of $PM_{2.5}$ and PM_{10} are only considered for inclusion if they also include a relevant exposure biomarker. Examples of relevant biomarkers need to be provided (Note: the aerodynamic diameter for $PM_{2.5}$ and PM_{10} are 2.5 and 10 μm, not 2.5 and 10 μm^3).

Section 12.4.4 "Welfare- Effects on Terrestrial and Aquatic Ecosystems" notes the exclusion publications of site-specific studies at non-U.S locations and those near contaminated sites (e.g., mine tailings and industrial effluents) due to the high Pb concentrations and lack of connection to an air-related source or process (Lines 17-21, Page 12-10). It appears that the criteria set forth a priori exclude relevant publications that may be useful for science and policy assessment.
The rationale to select exposure cut-point concentration levels needs to be documented. Levels noted in the footnote of Tables 12-3 and 12-4 (Pages 12-9 and 12-12) are informative, but need to be stated upfront in the text.

Section 12.5 "Literature Search and Study Selection" includes the period from September 2011 to December 2020 (Lines 11-14, Page 12-13), with an update through June 2022. Given an anticipated final decision on Pb NAAQS in early 2026 and a final ISA in late 2023, a literature search cutoff date of June 2022 may miss some up-to-date scientific findings. Amendments or some special addition to the ISA need to be considered that allow for the update of scientific information.

Dr. Deborah Cory-Slechta

Health Effects of Pb Exposure

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

In general, the appendix accurately describes the collective strengths and limitations of the various types of health studies. However, some of the toxicological studies, particularly of animal behavior do not include potential limitations or uncertainties. For example, in studies using the water maze, there can be multiple reasons for delay in time to reach the escape platform that are not related to slower learning, e.g., deficits in motor function or in sensory capabilities required by the task. A study of swim speed is sometimes included as a control, but for potential differences in motor function, but the task is one that requires motor endurance which is not measured. Few studies adequately address these potential limitations. It may be that the authors do not report these controls. In tasks such as novel object recognition, data are only described for the ultimate outcome, the recognition index, i.e., whether the animals showed increased time allocation to a novel stimulus. But behavior during the preceding habituation session can also be critical in determining whether there are differences in the amount of activity per se, or perhaps side bias in response to the exposure. Again, these may not even be reported in the study per se but that should also be noted.

Another issue related to above that also increases confusion is the issue of blood Pb concentrations, as the studies considered in different health endpoints may involve different blood Pb values. It isn't clear how concentration-effect is being incorporated into the assessment of whether there are causal effects of Pb or not. It would be very helpful to include graphics or forest plots for all of the health endpoints; there are some shown for some of the health endpoints but this is not done routinely.

3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on the biological plausibility?

In general, the inclusion of evidence from mechanistic studies is useful in terms of informing biological plausibility.

3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

What is difficult to ascertain across the various health outcomes is how interpretations/conclusions were arrived at based on the collective evidence. This is particularly the case where the evidence is more inconsistent and where the interpretations/conclusions seem somewhat arbitrary. The IS lists the various criteria that were used, but no definitions or more refined categorization is provided for these, only a reference to a previous EPA document. Presumably the ultimate interpretation/conclusion relies on what is considered the most important of outcomes within a health category, but this is not always clear. This makes it difficult to determine how appropriate some of the conclusions are.

It is also difficult to ascertain how the interpretation of the collective evidence was arrived at because in some of the various outcomes included, there is a good deal of repetition of the same evidence both from

the 2013 ISA and the current evidence base. This makes it difficult for the reader to sift out what is the basis for the ultimate conclusions. One recommendation would be to include a specific italicized paragraph that is the ultimate basis for the conclusion for each of the health effects studied and to describe why it wasn't decided for the category above (if appropriate) or below it.

Executive Summary and Integrated Synthesis

6a. Please comment on the clarity with which the ES communicates the key information from the Draft Pb ISA.

In general, the ES communicates the overall and key information very well. The figures are easy to understand and provide a good overview of what is then communicated in more detail in the text.

6b. Please provide recommendations on any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

In general, it seems as if what is included in the ES is appropriate, and the inclusion of the key aspects of health and welfare evidence is important and belongs in the ES as these are key points in the assessment.

7a. Please comment on the usefulness and effectiveness of the summary presentation in the IS and provide any recommendations or alternate text that may improve the synthesis of available information across subject areas and the communication of key findings.

I'm assuming this refers to the text in the gray box at the beginning of the IS section. It is a very good summary of the information from the appendices. The IS is generally well written. In particular, the comparisons of the 2013 vs the 2023 ISA basis for conclusions is quite helpful and is further explained by the accompanying text.

7b. The IS includes a summary of evidence related to concentration-response relationships for human health effects and the timing of Pb exposure contributing to nervous system effects. To what extent do these sections appropriately synthesize the available evidence? To what extent do the conclusions in these sections adequately reflect the strengths and limitations of the evidence?
The IS appropriately synthesizes the available evidence for both concentration-response relationships for human health effects as well as for the timing of Pb exposure contributing to nervous system effects. The available evidence for each of these two topics is appropriately synthesized, and the conclusions accurately reflect both the strengths and limitations of what is currently known.

7c. To what extent does the IS appropriately synthesize the evidence for populations at increased risk of experiencing effects due to Pb exposures, including consideration of children's health? To what extent do at-risk conclusions adequately reflect the strengths and limitations of the evidence?

The IS appropriately synthesizes the evidence for populations at increased risk of experiencing effects due to Pb exposures, including consideration of children's health. These are an important addition to the ISA because they are relevant to the issue of cumulative risk in relation to Pb exposure and to the goal of protecting the most vulnerable populations. In addition, the at-risk conclusions do adequately reflect the collective evidence, both in terms of its strengths and limitations.

Dr. Christina H. Fuller

Executive Summary and Integrated Synthesis

6a. Please comment on the clarity with which the ES communicates the key information from the Draft Pb ISA.

6b. Please provide recommendations on any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

Overall, the ES communicates key information in a clear and organized manner through the text and figures. However, the ES is missing a summary of contextual factors that are key in understanding increased risk for at-risk groups. I suggest that this section briefly summarize the context related to at-risk groups including the delineation of intrinsic and extrinsic susceptibility as well as intersectionality between at-risk groups.

7a. Please comment on the usefulness and effectiveness of the summary presentation in the IS and provide any recommendations or alternate text that may improve the synthesis of available information across subject areas and the communication of key findings.

7b. The IS includes a summary of evidence related to concentration-response relationships for human health effects and the timing of Pb exposure contributing to nervous system effects. To what extent do these sections appropriately synthesize the available evidence? To what extent do the conclusions in these sections adequately reflect the strengths and limitations of the evidence?

The organization of the IS according to causality between Pb exposure and associated effects by health outcome is an effective and clear description of the scientific evidence. The tables comparing causality conclusions from the 2013 review and the current 2023 review are especially helpful in summarizing prior and recent scientific conclusions.

7c. To what extent does the IS appropriately synthesize the evidence for populations at increased risk of experiencing effects due to Pb exposures, including consideration of children's health? To what extent do at-risk conclusions adequately reflect the strengths and limitations of the evidence?

The draft ISA has improved in its consideration of at-risk groups as well as some of the underlying factors linked to disproportionate exposure and susceptibility. The inclusion of sections on at-risk groups in multiple appendices, where data is available, is an improvement and broadens the inclusion of these groups into assessment of overall risk for the most vulnerable and subsequent policy considerations. The IS of the ISA could be further improved by including greater context as to the identification of each at-risk group. The ISA includes some discussion of intrinsic and extrinsic factors associated with each at-risk group, but does not specifically explain whether intrinsic, extrinsic, or both, factors are linked to a particular at-risk group. For example, there is evidence that genetic factors are determined by intrinsic factors.

The intersectionality of identified at-risk groups is currently complex to quantify. The ISA states, "While a combination of factors (e.g., residential location and ES) may increase the risk of Pb-related health

effects in portions of the population, information on the interaction among factors remains limited." As a result, "this ISA characterizes the individual factors that potentially result in increased risk for Pb-related health effects." However, the relationships could be more thoroughly described qualitatively using the reviewed literature. This qualitative assessment could be meaningful in interpretation of findings for these groups as well as in the subsequent policy assessment.

Suggested edit for IS-60, lines 21-24: Delete "the presence of" in the following sentence, "Lastly, the recent evidence further supports and adds to the collective evidence presented in the 2013 Pb ISA that the presence of absence of certain nutrients may increase Pb-related health effects, while other nutrient deficiencies or surpluses may decrease the risk of a Pb-related health effect among certain populations." This change would increase the clarity of the sentence.

Dr. Philip. E. Goodrum

Appendix 2: Exposure, Toxicokinetics, and Biomarkers

Appendix 2 describes the multimedia nature of Pb exposure, toxicokinetics of Pb in humans, biomarkers of Pb exposure and body burden, as well as models of the relationship between Pb biomarkers and environmental Pb measurements.

2a. Please comment on the choice and emphasis of topics for providing useful context for the evaluation of human health effects of Pb in the ISA. Is the current organization of the appendix clear and logical?

Yes, the organization of the appendix is clear and logical. The general structure of – methods, example applications, and implications of findings – works very well. The level of detail on each element is appropriate. The Summary section is very well written.

Consider including call-out boxes to summarize key points for major sections.

There is overlap and connections between topics, both within this Appendix 2 and across the ISA document. Additional cross-references could be added. For example:

- Potential enrichment of Pb in smaller size fraction (e.g., <150 μm) can be an important consideration when studies involve measurements of soil and dust Pb (e.g., occurrence studies, media-to-media relationships, statistical analysis of associations between media and health effects). More consistent summaries of when studies do/do not account for enrichment would be helpful. Consider including this detail in any summaries of the literature (e.g., Table 2-3; Table 2-13; Section 2.5.3)

- The document describes the connection between inhalation and ingestion pathways for larger particle size fractions, highlighting 2.5 μm as potential tipping point. Page 2-11 (lines 1-4) briefly mentions the relevance of hand-to-mouth activity, but not the additional mechanism of inhalation followed by ingestion via mucociliary transport – this is described well in Section 2.2.1.1 (which also refers to Chapter 4 of 2019 PM ISA (U.S. EPA, 2019c)). A forward cross-reference to that section could be added, or use the same language that appears in the Summary on this topic.

A few observations relevant to overall clarity and completeness:

- Fig. 2-1. CSM diagram – it is not immediately obvious how to interpret: a) newly emitted Pb; b) historically emitted Pb; c) non-air Pb releases, for the following reasons:

    - If (a) and (b) are emissions directly to air, shouldn't they be in the blue shading in the Air circle, rather than in the region of overlap with outdoor soil & natural waters/sediments?

    - Non-air Pb releases may also be newly emitted or historically emitted;

    - Non-air Pb releases can contribute to Pb in air via resuspension.

Also, there are many additional potential sources of Pb in dust (e.g., paint, consumer products)

Consider switching the format of the CSM diagram to show sources > environmental media > pathways > receptors; using arrows to convey the situations whereby there are multiple connections.

- Pages 2-7, 2-8: there appears to be conflicting evidence regarding the relative magnitude of air concentrations measured from personal devices vs monitoring stations (indoors or outdoors). For example, the study summarized in Table 2-2 suggests there is no high-bias associated with personal air monitoring. While this becomes clear as the entire section is reviewed, the introductory sentence (Line 1 on p. 2-7) might be revised to establish this point right up front.

- Particle size distribution is a recurring theme as far as source contributions, exposure pathways (inhalation vs ingestion), and strength of correlation with blood Pb or health outcomes. Consider creating a diagram that gives different particle size distribution ranges on one side, and relevant categories/concepts on the other.

- Page 2-14. Summary statistics of soil ingestion rates are presented with reference to USEPA EFH Chapter 5. Add a sentence that explains that these are the recommended central tendency exposure (CTE) values for the general population, and perhaps why the CTE is more relevant that reasonable maximum exposure (RME) for mechanistic models like IEUBK and AALM. It would be helpful to generate a summary table that lists the age-specific summary statistics (means/medians) and corresponding age-specific parameter values currently used/recommended from USEPA guidance on lead risk assessment (e.g., see Table 16 of USEPA. 2021. Advancing Pb Exposure and Biokinetic Modeling for U.S. EPA Regulatory Decisions and Site Assessments using Bunker Hill Mining and Metallurgical Complex Superfund Site Data. Office of Research and Development. EPA 600/R-21/017F. April; also Table 1 from USEPA (no date). Estimation of Age-specific Soil and Dust Ingestion Rates for U.S. Children: Update to the Default Values for the IEUBK Model for Lead in Children).

- Section 2.1.3.3 Dietary – a key source of uncertainty in the national surveys is the method used to extrapolate from 24-hour recall to long-term average daily intakes. Its unclear if the citations referenced address this through the use of the NCI method, or similar techniques that account for the probability of food group selection on any given day (e.g., Kipnis et al., 2009; Tooze et al., 2006; Institute for Medicine, 2005)[1].

- Page 2-25, lines 16-17. While 2 ppm is correctly cited as the action limit for Pb in spices at the time that Hore et al. (2019) was published, New York has since updated its Class I and Class II

[1] Kipnis, V.; Midthune, D.; Buckman, D.W.; Dodd, K.W.; Guenther, P.M.; Krebs-Smith, S.M.; Subar, A.F.; Tooze, J.A.; Carroll, R. J.; Freedman, L.S. 2009. Modeling Data With Excess Zeros and Measurement Error: Application to Evaluating Relationships Between Episodically Consumed Foods and Health Outcomes. Biometrics.
Tooze, J.A.; Midthune, D.; Dodd, K.W.; Freedman, L.S.; Krebs-Smith, S.M.; Subar, A.F.; Guenther, P.M.; Carroll, R. J.; Kipnis, V. 2006. A New Statistical Method for Estimating the Usual Intake of Episodically Consumed Foods with Application to their Distribution. J. Amer. Diet. Assoc. 106(10): 1575-87.
Institute of Medicine (IOM). 2005. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids / Panel on Macronutrients, Panel on the Definition of Dietary Fiber, Subcommittee on Upper Reference Levels of Nutrients, Subcommittee on Interpretation and Uses of Dietary Reference Intakes, and the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes, Food and Nutrition Board, Institute of Medicine of the National Academies. Washington, D.C., National Academies Press. Available online at: http://www.nal.usda.gov/fnic/DRI/DRI_Energy/energy_full_report.pdf

action levels[2]. The current Class II action level is 0.21 ppm (see Ishida, M., Greene, V., King, T., Sheridan, R., Luker, J., Oglesby, D.V., Trodden, J., Greenberg, J., 2022. Regulatory policies for heavy metals in spices – a New York approach. https://doi.org/10.21423/JRS-V10I1ISHIDA).

- The Appendix refers to IEUBK and AALM, which is appropriate. ALM is introduced much later (p. 2-110). It might be helpful to at least introduce ALM and California's Leadspread earlier, as additional mechanistic modeling tools used to model (predict) relationships between environmental Pb levels and age-specific blood lead (or changes in blood lead in the case of Leadspread).

Please provide any recommendations to integrate exposure and toxicokinetics information more clearly throughout the appendix.

Section 2.1.5.2 Age – nicely integrates concepts of age-specific differences in blood Pb kinetics, and age-specific differences in exposure and physiology.

Section 2.2.1.2.1 Physiologic Factors – nicely references studies that examined variability in GI absorption associated with fed/fasted status.

Does the appendix adequately describe air-related and non-air related pathways of Pb exposure?

Yes. A few minor suggestions:

- Page 2-11. Define atmospheric soil Pb. Is it a subset of atmospheric Pb?
- See comment on CSM figure above.

2b. Please comment on how well Section 2.3 reflects the current state of knowledge of Pb biomarkers and their interpretation as it relates to exposure and dose. Is the focus on blood Pb and bone Pb appropriate given the epidemiologic literature largely assesses exposure through these two biomarkers?

Yes, the focus on blood Pb and bone Pb is appropriate. Other biomarkers are also discussed, and the document explains their limitations well. The difference between whole blood and plasma is covered well, including in the Summary.

Is there sufficient and accurate discussion of the relationship between blood Pb and bone Pb?

Yes. The expected shifts in blood Pb and bone Pb following changes in exposure is well described (e.g., Section 2.3.5.1). Discussions of half-life are accompanied by concepts of quasi-steady state, and continued exchanges between bone and blood (i.e., "prolonged terminal elimination phase"). Variability in rates of bone remodeling with age and pregnancy/lactation periods, as well as differences between trabecular and cortical bone kinetics are well described.

Different kinetics for blood and bone compartments over time, and implications for potentially confounding the interpretation of epidemiological study results, is well described (p. 2-75).

[2] The NYSAGM Division of Food Safety and Inspection (FSI) employs a recall classification system with two numerical designations: Class I recall is for products where reasonable probability exists that the use of, or exposure to, a violative product will cause serious adverse health consequences or death; Class II recall is applied when the use of, or exposure to, a violative product may cause temporary or medically reversible adverse health consequences or where the probability of serious adverse health consequences is remote.

Are relationships between blood Pb and Pb in soft tissues and urine Pb adequately described?

Yes. Urine Pb is marginally useful as a biomarker, and the Appendix makes this point well.

2c. Sections 2.5.1 and 2.5.2 discuss empirical models of the relationship between air Pb and blood Pb from recent and older studies. Please comment on the effectiveness of this section to accurately reflect what is known about air Pb-blood Pb relationships. Please provide recommendations on any studies that should receive less or greater emphasis.

The introduction provides a good overview of numerous sources of potential bias (both positive and negative) in the air Pb – blood Pb relationship. Consider adding a table that lists the various examples with simple checks in columns to show potential direction of bias.

The Appendix should spend additional time on alternative metrics – absolute blood Pb vs delta blood Pb. Empirical models are most easily applied to support a delta blood Pb risk metric, whereas mechanistic models are generally more applicable to absolute blood Pb. Both metrics have relevance to past CASAC assessments of air quality criteria, and both have had utility in the context of risk management of contaminated sites. Both require assumptions and reliance on multimedia lead sources, and source contributions to blood Pb. This discussion might precede the more detailed overviews of empirical models (2.5) and biokinetic models (2.6), to provide additional context.

Mr. Perry Gottesfeld

The draft Lead Integrated Science Assessment is very comprehensive, well prepared and clearly written. Below are some specific suggestions and comments for possible improvement:

Executive Summary and Integrated Synthesis

1) The health effects portion of Integrated Synthesis for Lead is too focused on the causal relationships that are already summarized in this section and fails to provide a simple summary of key exposures and most significant human health effects. It is not necessary to use the synthesis to explain the subtle differences between this ISA and the one from 2013 which is covered elsewhere. This synthesis could benefit by putting airborne lead exposures into context and using more simple language. The following examples of revised or added text are intended to improve clarity for a broad audience and highlight key findings:

- Pb exposures originating from airborne lead are a direct source of exposure and contribute directly and indirectly to the body burden through contributions from Pb contamination of soil, dust, agriculture, food, and water.
- We conclude from a review of the available evidence that Pb exposure in children causes cognitive function decrements and behavioral disorders including attention-related deficits, impulsivity and hyperactivity.
- We conclude from a review of the available evidence that Pb exposure in adults is causal for cognitive deficits and likely to be causal for psychopathological effects including depression, anxiety and panic disorders.
- This review of the available evidence demonstrates a causal relationship between Pb exposure and cardiovascular effects including increases in blood pressure, hypertension, and coronary heart disease.
- There is growing evidence supporting a causal relationship between Pb exposure and cardiovascular mortality and all-cause mortality.
- The available evidence provides clear evidence that minority and low-income communities suffer from greater exposure to Pb and there is suggestive evidence that some subpopulations are more susceptible to the health effects of Pb exposure.

2) The Welfare Effects section of the Integrated Synthesis for Lead does not mention key sources of exposure. This summary should include a statement such as: “Pb ammunition continues to be a prevalent source of Pb contamination in both mammals and birds” (section 11.2.2.4) and should note that contaminated soil is a key source of bioaccumulation through the food web and a direct source for other species.

3) The Executive Summary would benefit by rearranging the order of the topics discussed. Section ES.5 discusses “primary” contribution to blood lead levels without describing which population or subpopulation that this refers to. I would strongly avoid reaching any conclusions on “primary” vs. secondary exposure sources as you correctly indicate in the same paragraph that this is “situation specific.” Overall, this executive summary should start with a discussion to provide context to the multiple sources of lead exposure described in the document and include this key sentence about exposure pathway contributions being “situation specific” and this should not be buried on page 54. The

summary should provide a discussion of the changing nature of lead exposure over time as we regulated and controlled various emission sources and reduced the use of lead in some products (e.g., gasoline, residential paint) while noting the growing use of others (e.g., ammunition, lead batteries) over time.

This discussion would also benefit from pointing out that as the U.S. reduced exposures and median BLLs, that the overall contribution from each remaining source becomes more significant. Similarly in section ES.2 the discussion about the proportion of lead from various sources as listed in the NEI could benefit from the context of how the relative contribution from each shifts over time as some emission sources are reduced. For example, the proportionate contribution of airborne lead from aviation gasoline (AV gas) increases as industrial emissions have decreased since the 2008 Pb NAAQS. The discussion of AV gas should also point out that its relative contribution to airborne lead has increased while annual AV gas consumption has decreased (see data from US Energy Information Administration). This discussion can be combined with the section on Trends (ES.4) and should be moved up to beginning of the Executive Summary.

4) Table IS-1 can be improved as can the clarity of the health appendices by separating the discussion of Mortality from the section on "other organ systems".

5) Section IS 7.3.9 (and later text in health appendices) should provide a separate causal determination for cardiovascular mortality and total (nonaccidental) mortality. Separating the discussion of these two mortality categories would lessen the confusion and be more consistent with research findings as one is a subset of the other.

6) Section IS 7.4.2 discusses "proximity to lead sources" which is confusing. It may be clearer to say proximity to sources of airborne lead emissions. Line 19-20 adds to this confusion as describing an "increased risk of negative health outcomes" as it suggests that the epidemiological studies you rely on for this conclusion found an association between proximity and health effects that was independent of blood lead levels. (This is also stated in Table IS-13 which seems to suggest that the conclusions on proximity have changed since the ISA in 2013.5)

7) Table IS-13 uses the category of "residential factors" but describes only findings on lead levels in dust and ignores residential soil contamination. Instead of "residential factors" for clarity consider using housing age or residential lead contamination along with a description to include both dust and soil lead levels.

8) The discussion on blood pressure under Section 7.4.2.1 is confusing as you state "These racial differences may also reflect a history of greater exposure to Pb among non-Hispanic black populations." This is confusing and seems to contradict the conclusions presented of consistent findings of greater effect levels in African-Americans.

Appendix 1

1) The title is confusing as it is not clear what the "concentration" refers to. Is this specific to airborne concentration or total environmental exposure sources?

2) This appendix fails to provide context to the discussion of lead sources to include:

a) a clear statement that all exposures to lead are significant as lead is bioaccumulative;

b) the relative contribution from sources (exposure pathways) is situation specific and depends on age, geography, nutritional status, seasonal fluctuations and other factors;
c) that the relative contribution of environmental sources are changing as we eliminate or reduce some exposures (e.g., remove lead from gasoline or reduce NAAQS for lead);
d) airborne lead deposited in soil or dust becomes an ongoing source of exposure through resuspension, construction activities and in the case of children from direct contact/ hand to mouth activity.

3) Section 1.2 on sources of atmospheric Pb provides the breakdown of contributions as noted in the 2017 NEI but does not indicate how this data compares with air emissions reported in the TRI database. A discussion and summary of the TRI data for air lead emissions would be useful.

4) Section 1.2 is clear in discussing contributions from resuspension. However, it would be helpful to provide an aggregate estimate or range (if available) for soil resuspension that can help contextualize the contribution from legacy contamination and the importance of controlling air emission sources.

5) Section 1.2.2 fails to define the type of industries that are discussed here. In particular, it fails to provide the breakdown for the largest contributor of industrial lead emissions currently from battery manufacturing plants and battery recycling plants as per TRI data or ambient air monitoring data. These industrial sources are only mentioned in the section on legacy sources and in the context of international studies. The discussion of soil contamination from industrial sources can benefit from examples of studies of lead battery recycling and/or manufacturing sites in the U.S. See:

- An-Min, W., & Johnston, J. (2019). Assessing Spatial Characteristics of Soil Lead Contamination in the Residential Neighborhoods Near the Exide Battery Smelter. *Case Studies in the Environment*, *3*(1), 1-9.
- Small, M. J., & Rose, S. M. (2015). Statistical Analysis of Soil Lead Concentrations in Vernon, CA. *Advanced geoservices: Pennsylvania*.
- Greipsson, S., Tay, C., Whatley, A., & Deocampo, D. M. (2013). Sharp decline in lead contamination in topsoil away from a smelter and lead migration in ultisol. *World Environment*, *3*(3), 102-107.

6) Section 1.2.8 incorrectly states that "Informal industry is defined as industry characterized by a lack of adherence to regulation, including zoning and pollution controls." when in fact, "informal sector" is generally defined as employment in businesses that do not obtain legal permissions and do not pay taxes and are generally not recorded in GDP. Note that the study discussed in this context of soil contamination around 15 lead battery recycling plants in Africa includes only formal sector facilities.

7) Section 1.3.2.2 does not mention soil transport with organic forms of lead (e.g., TEL) from leaks or spills from aviation fuel which can be significant based on consumption figures and the figures for aviation gas discarded on the ground provided in Section 2.1.3.5.

8) Section 1.3.3.4 ends with a statement that the "primary source of Pb was determined to be from mining and manufacturing" but this needs to be corrected to say that these are sources of Pb contamination in surface water.

9) Section 1.6 calls aviation gas the "dominant contemporary source" but perhaps this should be more carefully stated to indicate that aviation gas is the largest contributor to lead air emissions in the U.S. and not the most significant exposure source. It should also provide context that aviation fuel is not the

largest source of exposure for children in the U.S. except for those living in proximity to certain airport facilities.

10) Appendix 1 should discuss the contribution of ammunition production and use to airborne lead. The appendix should also provide some discussion of industrial emissions from waste incineration.

11) Appendix 1 does not mention airborne emissions from short-term construction activities that does not get fully appreciated within ambient air monitoring data. Construction activities that may result in higher emissions for several days or weeks may include removal of road marking paints, maintenance of steel structures, demolition and exposures during abrasive blasting of lead paint. These may have similar short-term impacts to those reported from NASCAR events (see Hollingsworth 2021). Past and ongoing use of lead chromate and other lead pigments in "industrial" paints on roadways, bridges, and other steel structures likely contribute to airborne lead during construction/ demolition activities and through resuspension of soil.

This section could benefit from citing these sources:

- Lee, P. K., Yu, S., Chang, H. J., Cho, H. Y., Kang, M. J., & Chae, B. G. (2016). Lead chromate detected as a source of atmospheric Pb and Cr (VI) pollution. Scientific reports, 6(1), 1-10.
- LeGalley, E., Widom, E., Krekeler, M. P., & Kuentz, D. C. (2013). Chemical and lead isotope constraints on sources of metal pollution in street sediment and lichens in southwest Ohio. Applied Geochemistry, 32, 195-203.

Appendix 2

1) Section 2.1.2. states that Pb in soil can be measured with XRF and ICP but not mention Atomic Absorption (AAS) that is more commonly used for soil.

2) Section 2.1.2. discusses identifying a "primary Pb source" for an elevated BLL but fails to provide context that there often is no single "primary" source on the individual level as all exposures contribute to body burden of lead. Isotopic analysis is a research tool and generally not used to identify sources for an individual but are more often employed in studies designed to investigate sources in a given population or allocate sources of soil/ dust contamination. This should be noted here.

3) Appendix 2 discusses ammunition as a contributor to dietary intake and as a source of occupational exposure but given widespread gun ownership and the growth in ammunition sales in recent years, this exposure source warrants a separate discussion on its contribution to airborne lead exposures. This should provide context to recreational exposures in both indoor and outdoor firing ranges. For example, Section 2.1.3.1 fails to mention airborne Pb exposure from indoor and outdoor firing ranges, range and the frequency and duration of these exposures to customers of these facilities. See for example:

- Laidlaw, M. A., Filippelli, G., Mielke, H., Gulson, B., & Ball, A. S. (2017). Lead exposure at firing ranges—a review. *Environmental Health*, *16*(1), 1-15.
- Beaucham, C., Page, E., Alarcon, W. A., Calvert, G. M., Methner, M., & Schoonover, T. M. (2014). Indoor firing ranges and elevated blood lead levels—United States, 2002–2013. *Morbidity and Mortality Weekly Report*, *63*(16), 347.

4) Section 2.1.3.2.1- The discussion of soil and dust fails to mention that soils are a significant contributor of lead contamination in household dust. This section should also discuss the contribution of airborne lead to agricultural soils and crops.

5) Section 2.1.3.3.2 There is a typo in the last sentence.

6) Section 2.1.3.5 provides a good discussion of occupational exposure but should mention the relatively new military lead standard DOD dated 10/15/17 "Provisional Blood Lead Guidelines for occupational Monitoring of Lead Exposure in the DOD" Available at: https://apps.dtic.mil/sti/citations/AD1169209 Or https://www.esd.whs.mil/Portals/54/Documents/DD/issuances/dodm/605505m.PDF?ver=3u-UoR7v7wydz-RIKQkI_Q%3D%3D

7) Section 2.1.5.1 uses the title "proximity to lead sources" should be revised to proximity to sources of airborne lead emissions. This section may benefit from citing the study conducted by the California Department of Public Health around the Exide lead battery recycling facility that concluded: "children with higher blood lead levels lived slightly nearer to the Exide site." See: CDPH Childhood Lead Poisoning Prevention Branch. "An analysis of children's blood lead levels in the area around the Exide site." (2016) https://dtsc.ca.gov/wp-content/uploads/sites/31/2018/03/An-Analysis-of-Children-s-Blood-Lead-Levels-in-the-Area-Around-the-Exide-Site.pdf

8) Section 2.1.5.4 does not indicate if the reduction in the race/ ethnicity gap is proportionate to the reduction in BLLs in previous decades or if the gap has been reduced as a relative percentage of the subpopulation BLLs.

9) Appendix 2 fails to adequately discuss lead air emissions from waste to energy incinerator plants. There are currently 73 waste incinerators in the U.S. Some of these facilities are reporting air emissions of hundreds of pounds of lead per year. See for example:

- Baptista, A. I., & Perovich, A. (2019). US municipal solid waste incinerators: An industry in decline. *Tishman Environment and Design Center at The New School*.
- Li, Y., Zhang, H., Shao, L., Zhou, X., & He, P. (2019). Impact of municipal solid waste incineration on heavy metals in the surrounding soils by multivariate analysis and lead isotope analysis. *Journal of Environmental Sciences*, *82*, 47-56.

Appendix 3

1) This ISA evaluates teenage and adult crime along with conduct disorders among children in summarizing the literature. However, given that crime covers a wide range of offenses, is associated with multiple risk factors independent of lead exposure, and generally peaks among young adults, the consideration of studies linking lead and criminal behavior should be evaluated in a separate section.

In addition, the discussion of criminal behaviors/ crime is missing reference to an important study on fireman violence. Emer et al. (2020) using a retrospective cohort study design estimated association between childhood BLL (< 6 years) and firearm violence perpetration and victimization among 89,000 Milwaukee residents. They found a statistically significant elevated relative risk for perpetration and victimization for increased BLL with a dose response relationship. The attributable risk of an elevated blood lead level to firearm violence /perpetration was found to be greater than 50%. Mean and peak BLL showed similar results. See:

- Emer, L. R., Kalkbrenner, A. E., O'Brien, M., Yan, A., Cisler, R. A., & Weinhardt, L. (2020). Association of childhood blood lead levels with firearm violence perpetration and victimization in Milwaukee. *Environmental research*, *180*, 108822.

2) The discussion on cardiovascular mortality and all-cause mortality should be combined as suggested above. As noted, one is a subset of the other.

3) Section 4.12 could be strengthened by pointing out the large body of epidemiological studies linking high blood pressure to mortality (without consideration of blood lead levels).

Appendix 10

1) The summary causality determination in Section 10 refers to "cancer incidence and mortality" whereas the causality tables summarized elsewhere in the text refer to "cancer." The discussion in sections 10.1 and 10.6 also refer to cancer incidence and mortality but the evidence largely rests on toxicological studies as noted. Given the evidence presented, conclusions and causality determinations should be limited to cancer and not include cancer incidence or mortality.

Dr. Daven Henze

General: While I appreciate the intent of placing up-to-date figures and tables regarding current emission estimates and concentration trends in a separate document, a general comment regarding Appendix 1 is that this did make it a bit more tedious to evaluate. Will this content be merged into the final ISA?

1a. To what extent are Pb sources and emissions appropriately captured and described in Section 1.2? Are the relative importance and uncertainties in comparing contemporary vs. legacy Pb sources adequately explained?

- Section 1.2.1 discusses the avgas Pb emissions from takeoff / landing / taxing, but does not mention or compare these to in flight emissions. What are avgas emission outside of airports, and how do these contribute to long-range Pb transport or background concentrations?

- 1-22: Fig 1 is a bit lackluster. Is it possible to present a clearer diagram with known budgets / budget ranges? Also, isn't the source from aircraft in the form of PbBr, as explained in the supporting text?

- It would be interesting to connect discussion of Pb sources from fire with the discussion of Pb from paint, in the context of urban fires.

- I didn't see any mention of studies showing how COVID-19 lockdowns effected ambient Pb concentrations (and thus serve as validation of emissions estimates) – for example, Li et al. (2021)… further comments on this later.

1b. To what extent are Pb fate and transport adequately covered in Section 1.3 with appropriate detail and balance across media?

- 1-23/1-24: This is a bit vague in terms of quantitative descriptions of atmospheric lifetime and transport distances. What is meant by "long distances" (mentioned twice)? What is a short lifetime vs a long lifetime in this context? This type of text is repeated I-25 lines 30, 31… "transported over long distances and being deposited in remote environments." I think this is especially interesting with regards to the observation that Pb deposition increases at elevation (despite being remote) given increased wet deposition. However, the discussion of background concentrations (I-66) makes it sound like concentrations would not be elevated far from sources.

- I-49, line 15-17: I found the statement here that neighborhood-scale variability in ambient Pb was not directly related to soil concentrations to be a bit odd, given the explanation of neighborhood variability in ambient air Pb as related to historical impacts on soil explained earlier (eg I-20, section 1.2.7, lines 23-26) as well as statements on the next page (I-50, 6-8) regarding the role of historical sources.

1c. To what extent are recent advances in the development of measurement methods and method performance adequately covered in Section 1.4?

- It is stated that there are no studies of Pb diurnal variability, though I suspect (although I have not searched myself) there are studies showing the diurnal variability of PM factors, which include a strong Pb signal.

1d. To what extent do presentation of recent concentration trends and size distribution observations in Section 1.5 adequately reflect the recent literature?

- The summary of trends in the ISA (section 1.5.1) begins with a discussion of the period 2019-2021. However, there's no mention of declining ambient concentrations in response to COVID-19 lockdowns (e.g., Li et al., 2021).

1e. The topics covered in Appendix 1 were selected to provide useful context ranging from sources to concentrations of Pb relevant to understanding health and ecological effects covered in the ISA. Please identify any missing or incomplete areas of research and provide specific additional studies which would fill any gaps in this appendix. Are any research areas overemphasized and is there specific text that could be reduced in length to remedy this?

- As mentioned in a few places above, it seems the potential for using the COVID-19 lockdown to test of estimates of anthropogenic Pb emissions is a missing area of research. For comparison, for other atmospheric constituents such as NOx, CO2, NH3, there is a tremendous volume of research (e.g., Le Quéré et al., 2020, Liu et al., 2021, Cao et al., 2022) on this topic. In a very brief search, I did notice a study showing significant (~59%, "deweathered") declines in ambient Pb in the North China Plain (Li et al., 2021). I suspect there are other studies which, while perhaps not focusing on Pb alone, may have included this in analysis of PM composition changes during the lockdown. How do these studies challenge or support our estimates of Pb sources? For example, the work of Xu et al. (2019) implied an underestimate of US anthropogenic Pb sources of a factor of five – are changes in response to lockdowns elucidating here? Are there corresponding changes to exposure and or concentrations of Pb in other media that have been noted following 2020?

References:

Cao, H., D. K. Henze, K. Cady-Pereira, B. C. McDonald, C. Harkins, K. Sun, K. W. Bowman, T.-M. Fu, O. Nawaz (2022), COVID-19 lockdowns afford the first satellite-based confirmation that vehicles are an under-recognized source of urban NH3 pollution in Los Angeles, *Environ. Sci. Technol. Lett.*, 9, 1-3, https://doi.org/10.1021/acs.estlett.1c00730.

Le Quéré *et al.*, Temporary reduction in daily global CO_2 emissions during the COVID-19 forced confinement. *Nat. Clim. Chang.* **10**, 647–653 (2020).

Li, R., Zhao, Y., Fu, H., Chen, J., Peng, M., and Wang, C.: Substantial changes in gaseous pollutants and chemical compositions in fine particles in the North China Plain during the COVID-19 lockdown period: anthropogenic vs. meteorological influences, Atmos. Chem. Phys., 21, 8677–8692, https://doi.org/10.5194/acp-21-8677-2021, 2021.

Liu, Q. et al. Spatiotemporal changes in global nitrogen dioxide emission due to COVID-19 mitigation policies. *Sci. Total Environ.* **776**, 146027 (2021).

- It is recognized that the ISA incorporated some systematic review methods, but not wholesale. A statement that discusses the extent to which systematic review methods were used would be useful.
- "Aspects to Aid in Judging Causality" are understood to be related to the Bradford Hill criteria. It would be helpful to have this explicitly stated.
- The document should distinguish susceptibility from increased exposure from susceptibility from increased sensitivity to lead dose. This is an important distinction, given that the former relates to disparities in exposure (which are often related to societal inequities) whereas the latter may be related to nutrition, life stage, genetic traits and other modifying factors.
- Consider changing the blood lead metric from ug/dL to ug/L, to deflect the perception that 1 microgram/dL is necessarily a vanishingly small exposure. It is not, since it is also clear that pre-industrial blood lead levels were well under 0.1 microgram/dL.
- Lead exposure and conduct disorders (e.g., aggression and even violence) are the subject of an on-going systematic review being conducted by Ellen Kirane, Rachel Shaffer, and others in the EPA IRIS in collaboration with outside academics (such as myself). This could be acknowledged. The initial paper on methodology has been published (Shaffer et al., 2022).
- In utero and postnatal effects seem to be the same? (p. 40) Not sure this is true! The evidence for impacts (and lasting impacts) from in utero exposures is increasing (Hu et al., 2006; Liu et al. 2014; Jedrychowski et al., 2009; Merced-Nieves et al., 2022).
- The general issue of pregnancy-associated mobilization of bone lead as risk factor for fetal lead exposure (Committee on Obstetric Practice, 2012) arguably should be acknowledged.
- Are the animal models of aggression well-established?

References

Committee on Obstetric Practice. Committee opinion No. 533: lead screening during pregnancy and lactation. Obstet Gynecol. 2012 Aug;120(2 Pt 1):416-20. doi: 10.1097/AOG.0b013e31826804e8. PMID: 22825110.

Hu H, Téllez-Rojo MM, Bellinger D, Smith D, Ettinger AS, Lamadrid-Figueroa H, Schwartz J, Schnaas L, Mercado-García A, Hernández-Avila M. Fetal lead exposure at each stage of pregnancy as a predictor of infant mental development. Environ Health Perspect. 2006 Nov;114(11):1730-5. doi: 10.1289/ehp.9067. PMID: 17107860; PMCID: PMC1665421.

Jedrychowski W, Perera F, Jankowski J, Mrozek-Budzyn D, Mroz E, Flak E, Edwards S, Skarupa A, Lisowska-Miszczyk I. Gender specific differences in neurodevelopmental effects of prenatal exposure to very low-lead levels: the prospective cohort study in three-year olds. Early Hum Dev. 2009 Aug;85(8):503-10. doi: 10.1016/j.earlhumdev.2009.04.006. Epub 2009 May 17. PMID: 19450938; PMCID: PMC3725459.

Liu J, Gao D, Chen Y, Jing J, Hu Q, Chen Y. Lead exposure at each stage of pregnancy and neurobehavioral development of neonates. Neurotoxicology. 2014 Sep;44:1-7. doi: 10.1016/j.neuro.2014.03.003. Epub 2014 Apr 2. PMID: 24704588.

Merced-Nieves FM, Chelonis J, Pantic I, Schnass L, Téllez-Rojo MM, Braun JM, Paule MG, Wright RJ, Wright RO, Curtin P. Sexually dimorphic associations between prenatal blood lead exposure and performance on a behavioral testing battery in children. Neurotoxicol Teratol. 2022 Mar-Apr;90:107075. doi: 10.1016/j.ntt.2022.107075. Epub 2022 Jan 31. PMID: 35108597; PMCID: PMC8957713.

Shaffer RM, Forsyth JE, Ferraro G, Till C, Carlson LM, Hester K, Haddock A, Strawbridge J, Lanfear CC, Hu H, Kirrane E. Lead exposure and antisocial behavior: A systematic review protocol. Environ Int. 2022 Aug 4;168:107438. doi: 10.1016/j.envint.2022.107438. Epub ahead of print. PMID: 35994796.

Below are references related to my comment on how methods have been developed to impute/predict bone lead levels using blood lead levels and other parameters commonly measured in cohort studies; and how this imputed measure can then be used in existing cohort studies to determine how imputed/predicted bone lead may predict health outcomes that are not clear from simply using blood lead levels.

Park SK, Mukherjee B, Xia X, Sparrow D, Weisskopf MG, Nie H, Hu H. Bone lead level prediction models and their application to examine the relationship of lead exposure and hypertension in the Third National Health and Nutrition Examination Survey. J Occup Environ Med. 2009 Dec;51(12):1422-36. doi: 10.1097/JOM.0b013e3181bf6c8d. PMID: 19952788; PMCID: PMC2939477.

Wang X, Bakulski KM, Mukherjee B, Hu H, Park SK. Predicting cumulative lead (Pb) exposure using the Super Learner algorithm. Chemosphere. 2022 Nov 5;311(Pt 2):137125. doi: 10.1016/j.chemosphere.2022.137125. Epub ahead of print. PMID: 36347347.

Dr. Chris E. Johnson

Appendix 11 – Welfare Effects of Pb Exposure

Question 4a: Introduction (Section 11.1) The welfare effects appendix has an introductory section that includes concepts and tools for evaluating Pb effects on organisms and ecosystems. To what extent do the choice and emphasis of topics in the introduction provide adequate context for the evaluation of ecological effects of Pb in the ISA?

Section 11.1 includes a good, balance discussion of concepts, models and approaches. The discussion of biotic ligand models (BLMs) and multiple linear regression (MLR) models is especially useful.

The LECES approach to culling the literature is fine, but the exposure cutoffs for soil (230 mg/kg) and especially for water (10 mg/L) seem quite low.

Equation (2) on page 11-24 cannot be used to compute the final chronic value (FCV). This section should be clarified/corrected.

Question 4b: Terrestrial (Section 11.2) Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

The discussion of terrestrial ecosystems is comprehensive and highlights a number of important advances since the last ISA. The text is clear and very well-written.

The significance of Pb "aging" on the interpretation of study results is interesting and a useful addition. One must remember, though, that an atom of Pb is an atom of Pb. It doesn't "age" in any meaningful way. What ages are the components with which the Pb is associated – natural organic matter, soil/sediment minerals, etc. This is largely semantics, but it's unfortunate that this jargon has crept into the Pb literature.

The importance of certain types of mycorrhizal fungi in mediating uptake of Pb by plants is a welcome and important observation in this section of the ISA.

It is also encouraging to see a growing literature on the influence, or lack thereof, of Pb on microbial community structure and function.

All in all, this section of the ISA supports the causality determinations in Table 11-2. The literature review and analysis were remarkably consistent with the causality determinations from 2013, supporting the decision to leave them unchanged.

Question 4c: Freshwater (Section 11.3) Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or

removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

This section of the ISA does a good job of summarizing new findings in the context of previous causality determinations. New research refines some key observations and nearly always supports the 2013 determinations.

Based on the new studies cited, it would seem that the causality determination for neurobehavioral effects on aquatic invertebrates could be upgraded to "causal." What was the basis for not doing this?

It is encouraging to see more studies that are trying to discern the relative contributions of dietary uptake and water column uptake to the total Pb burden in aquatic organisms.

The ISA highlights some very interesting new findings from studies of zebrafish at realistic Pb concentrations. This is particularly important because zebrafish are a suitable analog for several human functional traits.

It would seem that the revised CMC and CCC values for low-hardness, low-pH, low-DOC freshwaters are themselves low enough to possibly warrant consideration in the development of a secondary standard.

That said: For the most part, significant negative effects in freshwater systems were found at Pb concentrations well above typical ambient concentrations. The confounding effects of DOC, pH and hardness make it very difficult to envision a workable model relating air-Pb to aqueous-Pb.

Question 4d: Saltwater (Section 11.4) Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

In past assessments, a paucity of studies and data from saltwater systems has made it difficult to assess the significance of Pb in individual organisms and communities. This seems to be changing somewhat, resulting in a few new causality determinations. However, our knowledge of Pb cycling and toxicity in these ecosystems still lags behind terrestrial and freshwater systems. The generally high pH and salinity of these systems, and in some cases high DOC as well, create conditions in which the percentage of free Pb^{2+} tends to be very, very low. Due diligence requires us to continue to look for new insights about Pb in saltwater environments, but it seems much more likely that terrestrial and/or freshwater systems will guide policy regarding environmental Pb toxicity.

Executive Summary

Question 6a: Please comment on the clarity with which the ES communicates the key information from the Draft Pb ISA.

The Executive Summary of the ISA document is generally well written. The treatment of Pb sources, characterization of emitted Pb, and fate and transport were informative, relevant, and sufficient. The health and welfare sections give a concise summary of the key findings in the ISA.

Question 6b: Please provide recommendations on any information that should be added to the ES or information that should be removed and left for discussion in other parts of the document.

The "Venn diagram" (Figure ES-1 – it's not really a Venn Diagram) is not particularly helpful or well-constructed. To the extent that it's important to get across the overlapping nature of Pb sources and fates, perhaps a more professionally designed image is called for.

The causality tables that are included in the Executive Summary are excellent, clear and very helpful.

I am not entirely comfortable with the sentence that occurs at page 17, lines 7-10, near the end of the summary:

> *"In other words, in older studies in which leaded gasoline or local sources were a major contributor to air Pb, there may be a greater likelihood of discerning the true effect of air Pb on blood Pb due to relatively less contribution from non-air exposure pathways."*

While this sentence makes some sense, the fact is that the EPA is charged with setting a standard that pertains in the circumstances that exist now.

On page 17 of the Executive Summary, while discussing both the air-Pb to blood-Pb relationship and the blood Pb to cognitive response relationships the point is made that the slopes of these relationships tend to be greater at lower values of the independent variable. Does it make sense in this case to move away from linear models and adopt some sort of non-linear approach that can more adequately capture the forms of these relationships?

Dr. Susan Korrick

Charge questions for health appendices (3-10):

1. *Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitation of various types of health studies, including epidemiologic and toxicologic studies.*
2. *What are the Panel's [sic: your] views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?*
3. *To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?*

Appendix 3: NERVOUS SYSTEM EFFECTS

General comments:

In Appendix 3, I think there are some areas where clarification would be useful. These include the following:

1. With rare exceptions, tooth lead levels are characterized based on the timing of their collection (e.g., 6-8 years) seeming to imply that this biomarker reflects childhood exposures at these ages. In general, tooth lead estimates prenatal and early postnatal exposures and often is analyzed to characterize specific exposure time windows across pregnancy and early childhood. Being more explicit about the exposure window reflected in these biomarkers in a given study would be useful.

2. Comments such as "Uncertainty regarding past exposures among women in studies reporting associations with maternal or cord BLL" seem misplaced. Cord blood is fetal blood so cord BLL reflects fetal exposure (in later pregnancy) -- with this direct biomarker of exposure during a relatively short time window, maternal past exposures should not be relevant. That is, in a pregnancy without large fluctuations in lead exposure (e.g., absent maternal intermittent occupational Pb exposure or other acute exposure events), a cord BLL of 2 µg/dL in an infant born in 1995 should reflect exposures that are comparable to a cord BLL of 2 µg/dL in an infant born in 2020. A similar argument applies to maternal pregnancy BLL which is highly correlated with cord BLL. Fetal exposure assessed in a short time window is the key measure and this should be relatively independent of maternal exposure history. In short, this exposure assessment scenario should not be considered a limitation.

3. I applaud the emphasis on precision of effect estimates rather than p-values but, in a number of places, conclusions are made without acknowledging substantial uncertainty. A representative example (see section 3-149, page 563): " Choi and Park (2017) measured speech- and high-frequency hearing loss in adolescents (12–19 years) and adults (20–87 years) in the Korea National Health and Nutrition Examination Surveys (KNHANES)…. For each doubling of blood Pb, there was a positive association with speech-frequency hearing loss (OR = 1.2 [95% CI: 0.48, 3.05]) and high-frequency hearing loss (>25 dB) (OR = 1.26 [95% CI: 0.73, 2.16]) among adolescents."

4. I appreciate that Pb's observed impact on FSIQ in children is one of the most salient associations from public health and causation perspectives. FSIQ has been widely studied and has value and utility, not just as an indicator of neurodevelopmental impacts, but also as a predictor of long-term function as reviewed in the ISA (see 3-69, page 481): "FSIQ has strong psychometric properties (i.e., reliability, consistency, validity), is among the most rigorously standardized cognitive function measures, is relatively stable in school-age children, and has been predictive of educational achievement and life success." I am currently a collaborator on the NIH-funded ECHO (Environmental Influences on Child Health Outcomes) consortium and, as such, participate in ECHO's Neurodevelopment Working Group and its recently formed Neurodevelopment DEI Task Force. The latter group is working on recommendations for future research that would employ other metrics for assessing innate intelligence because IQ, as an omnibus measure, integrates tests of innate cognitive skills as well as cognitive skills that are correlated with access to educational opportunities and related resources. In my comments during our in-person discussions, I noted the potential for IQ to be biased. My comment was in reference to observations that different populations may score differently on IQ tests because of differential access to resources, ***not*** because prior studies utilizing IQ are potentially biased (Marks DF. "IQ variations across time, race, and nationality: an artifact of differences in literacy skills". *Psychol Rep* 2010 Jun;106(3):643064; Shuttleworh-Edwards AB. "Generally representative is representative of none: commentary on the pitfalls of IQ test standardization in multicultural settings." *Clin Neuropsychol* 2016 Oct; 30(7):975-98.). In addition, historically in the U.S., IQ has sometimes been used as a mechanism to "justify" discriminatory practices and, as a result, can be perceived as problematic. This additional issue does not invalidate the utility, validity or robustness of Pb-IQ associations demonstrated in previous studies but supports consideration of alternative metrics in future research. Admittedly, the above issues are not a direct commentary on the current ISA, but future research reviews may benefit from consideration of these emerging issues in cognitive assessment.

5. Although the literature summaries in each section touch on a variety of strengths and limitations, they are sometimes limited in scope. For example, study summary text is inconsistent in provision of study n's (some sections do this, others don't) which can be highly relevant, especially for studies reporting wide confidence limits. Summaries of cross-sectional and longitudinal study results are sometimes intermixed without noting the difference in inference that is appropriate for each design. The potential limitations associated with unexpected findings (e.g, biologically improbable "beneficial" associations with Pb) are not acknowledged. Perhaps most importantly, inconsistencies across the epidemiologic literature on a topic may reflect variability in the robustness of study design. When well-designed studies demonstrate consistent findings, that may be a better indicator of likely associations (null or positive) than consideration of all available literature. This distinction was sometimes hard to glean from the ISA analyses and I would speculate is more likely to contribute to under- rather than over-estimation of Pb effects.

6. Potential minor errors to correct:

    a. Taylor 2017 ALSPAC study is reportedly excluded from a summary figure because it used change in the raw IQ scores (see 3-32, page 444). My understanding of this study is that they used standardized (not raw) IQ scores but reported results using "unstandardized beta coefficients". The latter is common and is unrelated to the IQ measure, perhaps that

is a source of confusion? Or perhaps I'm misunderstanding? If not, it's good to include this study.

b. 3-140/page 552 typo: [95% CI: 30.12, -0.79] should be [95% CI: -30.12, -0.79]
c. This is not in Appendix 3 but is related. There are inconsistencies in description of causal assessments for various outcomes in IS-1: (a) summary list neurodegenerative disease as "likely causal" vs. summary table (figure ES-2) as "suggestive"; (b) summary lists total mortality as "likely causal" vs. summary table (figure ES-2) as "causal".

Comments on how accurately and appropriately the strengths and limitations of studies are described:

1. In discussing the literature in which Pb exposure is assessed in the context of exposure mixtures (to other metals or chemicals), a common criticism in this section is that models failed to account for potential interactions of Pb with other components of the mixture. In many cases, such studies have used BKMR to assess mixtures, including the potential for interactions among chemicals and, finding none, have used standard regression models to adjust for mixture components as confounders rather then effect modifiers. This is appropriate and should not be considered a limitation. For example, Oppenheimer used BKMR and found no evidence of chemical interactions (see 3-80, page 492) "Some of the available studies consider co-exposure to other chemicals and metals as confounders (Oppenheimer et al., 2022; Tatsuta et al., 2014) despite evidence that such co-exposures may interact with or modify the association between Pb and the outcomes (Yorifuji et al., 2011)."

2. There are a number of instances where the role of confounder adjustment in determining the validity of an analysis is unclear.

    a. E.g., in analyses of childhood cognition, adjustment for birth outcomes (such as birth weight or gestational age) is described as appropriate confounder adjustment. However, it is also possible that birth outcomes are causal intermediates in the relation of Pb with child cognition such that adjustment may or may not be appropriate, depending on the study question.
    b. In studies assessing bone Pb in relation to adult psychopathology, residual confounding by age is identified as a potential limitation. However, in contrast to cognitive function, the relation of age to psychopathology is not straight forward so the potential for residual confounding by age with these outcomes seems much less problematic than with cognition.
    c. Failure to account for psychotropic medication use is also cited as a potential source of confounding bias (and limitation) in studies of Pb and psychopathology (e.g., see 3-196, page 608 and 3-198, page 610). Although medication use may impact psychological symptoms, it seems unlikely that medication use would also impact Pb exposure and therefore it is unlikely to be a confounder. Instead, it is more likely a source of outcome misclassification which can decrease precision but generally does not bias effect estimates.

3. The interpretation of unexpected findings (e.g. "beneficial" impacts of Pb) and their role as potential study limitations is not always clear. E.g. in describing findings in Kim et al, 2013b/c (see 3-39, page 451): "Further, an increase in MDI was observed in association with early pregnancy maternal BLL among those with Cd levels below the median (β = 2.44 [95% CI: 0.04, 4.83]), indicating a pattern of interaction between Pb and Cd exposure that may be dependent on

the stage of pregnancy." Pb should not be beneficial regardless of exposure timing; study limitations are likely relevant here such as differential confounding, selection bias, or random error, etc.

4. Many epidemiologic studies published since the 2013 ISA assess lower Pb exposures. In terms of childhood FSIQ, there has also been more emphasis on prenatal exposures in more recent studies as compared to studies reviewed in 2013 where early childhood/childhood exposures were the focus. I think this literature supports the conclusion that low-level prenatal Pb exposures may not be as strongly associated with childhood FSIQ as early childhood/childhood Pb exposures. However, I don't think this message is clear.

5. There is potentially greater diversity of populations studied in the more recent literature on child FSIQ with a large number of international cohorts. Some discussion of whether culturally and linguistically appropriate IQ assessment instruments were used would be of use vis-à-vis study strengths and limitations.

6. ISA assumes behavioral check lists (e.g. completed by parent, teacher or self) have more measurement error (or lack validation) as compared to psychometric tests (3-86, page 499 and 3-100, page 512). This is not accurate – the behavioral check lists used in the studies cited are standardized and validated instruments. There are internal checks that can be used to identify less reliable respondents and studies may be inconsistent in whether or not they use these checks, but these are generally reliable and valid instruments and their use should not be considered a limitation. Use of SEMs or other analytic approaches is of interest as described in the text but this does not address their inherent properties. In fact, behavioral check lists are often more predictive of functionally important outcomes than formal psychometric tests.

7. Many epidemiologic studies published since the 2013 ISA applied current state-of-the-art approaches to analyses including assessment of metal/chemical mixtures and potential effect modification by various susceptibility factors. It sounds as if these designs are not as readily incorporated into the ISA evidence as are the single exposure biomarker, main effects analyses that were common in the past. I think this is a lost opportunity. The C-R relationship of Pb exposure in vulnerable population subsets (defined by those with co-exposure to Mn or maternal stress, e.g.) should still be relevant to the ISA. Perhaps I am misunderstanding how such study designs were considered in the ISA C-R/causality assessments? Also, assessing effect modification was identified as contributing to less consistent findings in the most recent literature, consistent with the potential for interaction analyses to be underpowered. I appreciate this may limit the utility, or at least the precision, of these study results. However, if there is true effect modification, assessing Pb associations in more vulnerable subgroups is critically important to estimating Pb's health impacts; failure to do so may underestimate such impacts.

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

See above comments re. some areas to address regarding the strengths and limitations of the evidence. Otherwise, causality determinations seem appropriate. A few general comments re. causality rationale:

1. It seems as if recent epidemiologic evidence (at BLL < 7 gμ/dL) is consistent in findings associations with internalizing behaviors. Better clarity regarding why Pb's relationship with this outcome is considered 'likely causal' would be helpful.

2. Especially for adult neurologic outcomes, it wasn't always clear if the causality determination was informed by there being few toxicology studies on a certain topic versus inconsistencies between toxicology studies and human epidemiology.

3. New animal data at lower exposures and assessing new measures of memory/learning and executive function enhance biological plausibility for these outcomes and related causality determinations.

4. Relationship of childhood BLL with adult cognitive function in epidemiologic literature and parallel toxicological study findings are important new observations re. causality (and public health) implications of Pb exposure.

Appendix 4: CARDIOVASCULAR EFFECTS

Comments on how accurately and appropriately the strengths and limitations of studies are described:

1. Except for cerebrovascular disease, literature on other CVD outcomes has expanded since 2013 at which time there was already a robust and extensive body of research supporting Pb's association with several CVD-related outcomes.

2. The relation of Pb with HTN appears to be strongest when cumulative (e.g., bone Pb) exposure rather than shorter term (e.g. blood Pb) exposures are modelled. Explicit acknowledgement of the role of exposure biomarker choices in assessing study strengths and limitations is not done here but would be useful.

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

1. Causality determinations are concordant with findings from 2013 ISA the strength of which has been enhanced by new, supporting literature especially regarding Pb's relation with BP (and, to a lesser extent, HTN) as well as CVD mortality. Appropriately so, the causality determination discusses emerging evidence of Pb's impact on cardiac electrophysiology and atherosclerosis.

2. Cerebrovascular disease evidence was inadequate in 2013 and remains so. Given the inter-related nature of these cardiovascular outcomes (and thus the justification for combining them), it seems

surprising that there is divergence of Pb associations (and resultant causality determinations). One issue for cerebrovascular disease is that there are fewer published studies (which could potentially reflect publication bias for null findings) but it would be interesting to consider what factor(s) might cause divergence of results for such closely inter-related disorders.

Appendix 5: RENAL EFFECTS

Comments on how accurately and appropriately the strengths and limitations of studies are described:

Several longitudinal studies of renal function outcomes have been done since the 2013 ISA supporting Pb associations and addressing concerns from the 2013 ISA about potential reverse causation. However, there are a number of new cross-sectional (and case-control) studies in the current ISA for which reverse causation is still a concern. In addition, longitudinal studies among populations with chronic kidney disease (CKD) are presented and these are also limited by the potential for reverse causation. This important potential limitation is not explicitly addressed in the review of cross sectional or CKD studies. In Table 5-1 summarizing evidence supporting a causal relationship (see page 1163/5-41), it would be good to indicate which studies in this table are prospective vs. cross sectional.

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done and compelling.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

As above, clarifying the relative importance of cross-sectional studies or even short-term longitudinal studies of renal function (e.g., follow up for 2 menstrual cycles) in determining causality would be helpful as such studies are still subject to reverse causality limitations that undermine certainty re. causality.

Appendix 6: IMMUNE SYSTEM EFFECTS

General comments;

There is an error in describing study results that is also in the summary sections of the ISA (ES and/or IS) – (see page 1243/6-32): "[Pb associations with]…reduced antibiotic resistance in children, as measured by nasal *Staphylococcus aureus* colonization (Eggers et al. 2018)" "Reduced antibiotic resistance" doesn't make sense – the study found Pb exposures associated with *increased nasal colonization with antibiotic resistant Staph aureus* (i.e., methicillin resistant *Staphylococcus aureus* or MRSA) and decreased colonization with methicillin susceptible *Staphylococccus aureus* or MSSA).

Comments on how accurately and appropriately the strengths and limitations of studies are described:

1. Key limitations of the epidemiology literature, including small numbers of cases of asthma or other atopic disorders (i.e., some studies were underpowered) were well described. However, the

diversity of outcome measures included in this section is notable but the relative strengths and limitations of each was not always clear. The ISA acknowledges that functional measures of immune function are preferrable but still reviews a substantial body of research that assesses non-functional outcomes (e.g., serum immunoglobulin levels, WBC populations). Similarly, in discussing Pesce et al (2021) (see page 1232/6-21), e.g.: "Specifically, in a small prospective birth cohort in France, Pesce 29 et al. (2021) reported that neither BLL measured during pregnancy nor cord BLL at birth were associated with incident parental-reported asthma attacks through 5 years of age. Notably, there was a low rate of asthma in the study population, limiting the statistical power to detect an association. However, because asthma can be difficult to diagnose in children under 5, asthma attacks may be the most reliable measure." It's unclear what parental reported "asthma attacks" means in children too young to have an asthma diagnosis but the study included children followed from ages 3-5 years with an additional follow up on most of the children at age 8 years. Diagnosis of asthma is probably reasonable in this age range. This same study observed associations between cord BLL and eczema which may be a cleaner parental reported outcome. The importance of this latter finding was overshadowed by concerns re. multiple testing. I would argue that differential associations may also be a consequence of variation in outcome sensitivity and validity (depending on how robust parental reported asthma attacks are in this cohort – by comparison, study guidelines for parental eczema reporting were very detailed).

2. The above section of text goes on with: " In a cross-sectional NHANES analysis including slightly older children (2–12 years old), Wells et al. (2014) also observed a null association between BLL and prevalent asthma." There is no discussion of the relative value of cross-sectional as compared to prospective studies of this topic. Timing of exposure relative to outcome is likely important in the development of asthma and other atopic disorders but is not ascertainable in a cross-sectional analysis.

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

Initially, I was unclear on why the causality assessment for "Sensitization and Allergic Response" was "likely causal" in 2013 but "suggestive" in the current ISA. For example, the toxicology literature supporting Pb's immunotoxicity is reportedly unchanged but more recent findings from the epidemiologic literature findings are inconsistent. This is an important reflection of research prioritization wherein relatively strong toxicologic evidence is not sufficient in the absence of supportive epidemiologic evidence. See Page 1235/6024): "Though limited in number, recent PECOS-relevant animal toxicological studies continue to support the findings from the last review. Specifically, these studies consistently report effects of Pb on sensitization and allergic responses including two studies of the effects of Pb exposure on production of cytokines relevant to immediate-type hypersensitivity. In contrast, recent epidemiologic evidence is not consistent with studies evaluated in the 2013 ISA. Specifically, recent studies provide little evidence of an association between exposure to Pb and atopic disease, and inconsistent evidence for immunological biomarkers involved in hypersensitivity and

allergic response. Similar to cohort studies evaluated in the 2013 ISA, recent longitudinal analyses are limited in number and have limited statistical power because of small case numbers."

Appendix 7: HEMATOLOGIC EFFECTS

Comments on how accurately and appropriately the strengths and limitations of studies are described:

Accurate discussion of limitations related to study design (most of this literature is cross-sectional) and validity (newer studies generally have better/more comprehensive approach to confounding adjustment than previously).

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done – mechanistic plausibility for Pb's hematologic effects is clear given experimental and epidemiologic evidence of the fundamental role high level Pb exposure has been shown to play in altering heme synthesis and downstream events such as RBC survival and function as well as clinical disease (e.g., anemia).

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

The causality determinations seem appropriate for the evidence but the approach for this outcome category seems different than for most others in the ISA. Specifically, causal assessment seems largely motivated by very strong toxicology data and very strong historical epidemiologic evidence (in occupationally exposed cohorts and heavily exposed populations whose Pb biomarker levels were much higher than today's U.S. population). There seems to be less focus on assessing causality at Pb exposure levels characteristic of today's population than is the case for other ISA outcomes. E.g., the epidemiologic evidence in 2013 and in the current ISA at lower Pb exposure levels is still limited (e.g., not a large number of studies, largely cross-sectional designs but reasonable confounder adjustment) with few studies in adults and most associations seen in children.

Appendix 8: REPRODUCTIVE AND DEVELOPMENTAL EFFECTS

Comments on how accurately and appropriately the strengths and limitations of studies are described:

1. In this section (and others), studies in which bone Pb and health outcomes are assessed at the same time are included among those limited by cross sectional design. When bone Pb is used as an exposure marker it represents years to decades of cumulative exposure, as a result, many of the limitations of cross-sectional design (e.g., uncertainty re. the directionality of the association) may not be applicable.

2. For pregnancy outcomes related to fetal loss, one of the more informative study designs assesses women seeking ART treatment as early pregnancy losses are ascertainable and this is otherwise a major limitation of most other designs. Results included Pb associated increased risk of of SAB before 12 weeks and lower risk of maintaining pregnancy. Studies in this category were noted to

be limited because of concerns about generalizability of women seeking fertility treatment but the important design strengths that result from ART cohorts, and therefore the value of findings, were not clearly acknowledged.

3. In text considering Pb's relation with age at menarche (see page 1381/8-34): "While these cross-sectional studies reported imprecise associations, the pattern of association is important to note. As BLLs decline, the association between blood Pb and age of menarche may be attenuated by potential confounders such as body weight and/adiposity." I am not sure what this means as some of the described study results are adjusted for BMI or %body fat. Also, it is unclear why this would be more problematic as BLLs decline. Is this comment confusing measures that may be on the causal pathway (↑Pb → Δadiposity → Δage at menarche) with those that are confounders? Adjustment for causal intermediates can attenuate associations but, again, this should apply to analyses of both high and low Pb exposures.

4. In the summary of the literature re. male puberty, height, weight and BMI are also considered potential confounders and adjustment for these factors is considered a study strength. However, as above, these measures are potentially on the causal pathway linking Pb to puberty and, as such, adjustment may not be appropriate (and may attenuate associations), depending on the study question.

5. Identified limitations in studies of female fertility (and male reproductive health) include use of women (and men) seeking help for fertility problems as this population is not generalizable to other women (men) and may reflect substantial selection bias. Is the assumption that, e.g., women (men) with fertility problems who seek help are more (or less) likely to be Pb exposed than women (men) with fertility problems who don't seek help? The text's explanation is hard to follow (see page 1400/8-53): "Additionally, by recruiting men who were seeking help at fertility clinics, there could be selection bias, as their fertility status is already known and those seeking help at fertility clinics may be different from men who have fertility issues who may be unaware of their condition and not seeking help at a fertility clinic."

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

1. For some reproductive/developmental outcomes (e.g., birth defects, postnatal growth, implantation loss) recent toxicology studies at lower Pb exposures are largely null whereas studies reviewed in 2013 ISA found Pb associations but sometimes in conjunction with maternal toxicity because of high Pb exposures. It's not clear how this context is applied to causal analysis – it seems as if lower Pb exposure in animal models (particularly rodents), absent maternal toxicity, may not be sensitive to the health outcomes observed in human studies. It is unclear how these null toxicology study findings are considered.

2. It seems as if the causality determination for male reproductive function (=causal) is largely driven by the association of BLLs with semen quality as this was the most consistent finding

among multiple measures of male reproductive function. One of the most consistent findings for female reproductive function was earlier menopause but this did not appear to be used as the basis for causality determination (=suggestive) despite its importance vis-à-vis women's health. It's unclear how the many different measures included as part of female and male reproductive function were prioritized for purposes of causality determination as the strength of evidence varied widely across measures for *both* men and women.

Appendix 9: EFFECTS ON OTHER SYSTEMS AND MORTALITY

Comments on how accurately and appropriately the strengths and limitations of studies are described:

1. See Page 1619 (9-6) regarding Hepatic Effects: "Similar to the Pollack et al. (2015) study, the median GGT levels in this study were within the normal range, making it difficult to interpret the clinical relevance of the results." I'm not sure if this was intended as a limitation but it should not be. When assessing a continuous outcome biomarker with an expected distribution in the general population, changes within the "normal" range don't have to be clinically significant to have public health importance. E.g., an exposure that shifts the distribution of a continuous outcome in a population can result in a disproportionate proportion of individuals in the population being at risk for abnormal outcomes.

2. See page 1651 (9-38) regarding Musculoskeletal Effects, specifically bone mineral density (BMD): "Importantly, the cross-sectional nature of the studies does not rule out the possibility that the association is driven by increased BLLs due to higher bone turnover in individuals with osteoporosis." This is *the* critical limitation of cross-sectional studies of BLL and BMD. In fact, the possibility of reverse causation is very likely and undermines the utility of such studies although the summary of the literature still considers these cross-sectional studies of BMD (or related measures such as osteoporosis) and BLL as contributory to causality determinations, etc.

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was acceptable.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

1. For some outcomes in this Appendix (e.g., Gastrointestinal Effects, Ocular Health, Respiratory Health, etc.) it seems as if the limited number of studies is potentially more important than the strength or even consistency of findings in determining causality. Is there some guideline used to determine an adequate number of studies on a topic?

2. For Total Mortality, if one excludes CVD mortality, the number of studies reporting associations of Pb with all-cause (non-accidental) mortality is relatively small and several utilize the same NHANES cohort. It's not clear to me how much of the causality determination is informed by CVD mortality and therefore how independent all-cause mortality is as an outcome. E.g., (see page 1684, 9-71): "The evidence for Pb-associated all-cause and cardiovascular mortality and strong supporting evidence for Pb-associated cardiovascular effects indicates **there is sufficient**

evidence to conclude that there is a *causal relationship* between Pb exposure and total (nonaccidental) mortality."

Appendix 10: CANCER

Comments on how accurately and appropriately the strengths and limitations of studies are described:

Strengths and limitations of individual studies were described reasonably and challenges of varying risk, depending on Pb's form (e.g., Pb chromate encompasses two potential carcinogens). Perhaps more importantly, the broader limitations to available literature on this topic was also reviewed. That is, since 2013, there have been essentially no new PECOS relevant toxicology studies. In 2013, epidemiologic studies were largely cross-sectional or ecologic designs often without direct exposure biomarkers or confounding adjustment. Limits to newer literature include that most studies are in occupationally exposed cohorts and often do not have measures of exposure biomarkers or confounders (e.g. smoking status, other occupational exposures).

Comment on integration of evidence from mechanistic studies to inform conclusions on biologically plausibility.

This was well done with a detailed review of evidence of Pb's role in genotoxicity/mutagenicity, disruption of DNA repair, and as a potential source of cancer promoting epigenetic alterations.

Comment on whether causality determinations appropriately reflect strengths and limitations of the evidence.

The causality determinations appear to be based on toxicologic data as the human data findings are very inconsistent, even among those with occupational Pb exposure.

HEALTH EFFECT APPENDICES: OVERALL

Overall comments:

In assessing the strengths and limitations of reviewed studies, some sections focus on the consistency of effect estimates across studies (patterns of findings) rather than strict requirements regarding precision/statistical significance whereas others dismiss study findings that fail to meet standards of statistical significance despite effect estimates of potential interest. As a result, the approach to summarizing the literature seems to vary by section/topic.

It also seems like there is variation in the framework for causality determination, depending on the outcome(s). In some cases, epidemiologic literature reflecting current low level Pb exposures seems to be key (e.g., for nervous system effects), in other cases (e.g., hematological effects or male reproductive function) more heavily exposed populations seem to play a more central role in the causality determination. In addition, the relative importance of toxicology vs. epidemiology evidence and their concordance (or lack of concordance) also seems to vary, depending on the outcome with cancer being one extreme wherein toxicology findings dominate the causality determination. Lastly, where a health category has multiple different outcome measures (e.g, renal effects, sensitization and allergic response,

or female reproductive function, etc.), it is unclear how discordance of likely causation across outcome measures is considered in causality determinations. More fundamentally, some outcome measures within a category may be more important (e.g., in terms of health consequences or likely mechanistic relationship with Pb) and therefore merit greater weight in causality determinations but, generally, the relative importance of outcome measures within a health category are not discussed.

Dr. Bruce Lanphear

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

Nervous Systems Effects

In general, the appendix of the draft ISA *For Nervous Systems Effects* was excellent.

It was unclear, however, how the EPA reached a causality determination when it primarily focused on reviewing new studies published since the 2013 ISA. It was useful to have an overview of the newer studies to determine whether they are consistent with previous conclusions, but many of them are too small, didn't examine the shape of the dose-response relationship, or were not critical for causality determination. This makes it difficult for the reader to focus on the critical studies for their causality determination. It would be valuable if the EPA screened new studies for quality and relevance prior to describing them in the ISA. An appendix of new studies that were not of sufficient quality to include in the ISA review could be included in future ISAs.

I recommend the EPA focus the ISA on high-quality studies (e.g., longitudinal studies that adjusted for key confounders, studies with high quality biomarkers, serial biomarkers) it relied on for their causality determination. For example, the description of the pooled analysis (Lanphear, 2005, Lanphear, 2019; Crump, 2010) was essential, but describing the other ten studies that examined the shape of the dose-response relationship and consistently found it exhibited a decelerating curve whether they used plasma, whole blood, or bone lead as biomarker would have been valuable in reaching a conclusion that no threshold exists (see, for example, Hu, 2006; Wasserman, 2003; Evens 2015).

It is also important to describe the studies of lead exposure and academic abilities in the section on cognitive abilities (Evens, 2015, Miranda, 2009). Those studies – which included large populations at the lowest measurable levels – also found the steepest decrements in academic abilities at lowest blood lead concentrations. Academic abilities are highly correlated with cognitive abilities and many of the children in these two studies would have had blood lead concentrations below 20 µg/L. Thus, these studies confirm that no threshold exists for lead-associated cognitive deficits.

Cardiovascular Effects

Page 1003, 4.6.2 Endothelial dysfunction

Using electron microscopy, van Strijp and others examined the effect of various concentrations of lead (0.14 µg/L, 4.1 µg/L and 8.2 µg/L) on EA.Hy926 endothelial cell line membrane integrity. They found an increase in perforations and blebs in the lead-exposed endothelial cells compared with controls. This new study is intriguing and raises important questions about how endothelial damage from lead and other toxic metals may initiate atherosclerosis (van Strijp, 2023).

Page 4-62, lines 9-11: "Notably, adult BLLs may be representative of contributions from both recent Pb exposures and mobilization of legacy Pb from bone, therefore it remains unclear as to what extent either

recent, past, or 10 cumulative Pb exposures contribute to the observed associations with cardiovascular mortality."

I agree that this is a limitation, but several lines of evidence indicate that we can draw conclusions about cardiovascular disease and deaths from existing cohort studies that rely on blood lead concentrations. First, fifteen prospective cohort studies conducted in Europe (4) and the United States (11) examined blood lead concentrations and cardiovascular mortality; all found that lead was a risk factor for CVD mortality (Navas-Acien, 2021). Second, many of the associations observed in prospective cohort studies that relied on blood lead concentrations were also observed in studies of bone lead concentrations. It is likely that studies that relied on blood lead underestimated the contribution of lead to cardiovascular disease, but they also indicate that blood lead levels are a valid and reliable biomarker. Finally, natural experiment studies and randomized controlled trials found that decrements in blood lead levels were associated with reductions in hypertension and fewer cardiovascular events within months or years, not decades (Adrian Ruiz-Hernandez, 2017; Lamas, Pirkle, 1985). In a RCT of chelation, decrements in blood lead concentrations were followed by reductions in cardiac events in about 18 months after the intervention began (Escolar, 2014).

The ISA reported that no effect modification was observed by Menke et al, but they did not report effect modification results reported in Lanphear et al (Lanphear, 2018) who relied on the same cohort with longer follow-up. Lanphear et al reported: "Examination of effect modification of the relation between concentration of lead in blood and key characteristics showed that HRs for participants younger than 50 years were significantly larger than were those for participants aged 50 years or older, for all-cause mortality (HR 2·24, 95% CI 1·50–3·34 *vs* 1·53, 1·18–1·98; p=0·003 for interaction), cardiovascular disease mortality (2·93, 1·60–5·36 *vs* 2·08, 1·35–3·19; p=0·01), and ischaemic heart disease mortality (4·68, 2·42–9·05 *vs* 2·46, 1·51–4·01; p=0·02). The HR for cardiovascular disease mortality was significantly larger for non-smokers than smokers (HR 2·19, 95% CI 1·47–3·26 *vs* 1·32, 0·86–2·05; p=0·03 for interaction)." (Lanphear, 2018).

3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

No comments.

3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

Renal Effects

I concur with the EPA's determination that lead is a causal risk factor for renal disease. Please include a description of a critical randomized controlled study by Lin et al. (Lin, 2003). This single-blinded RCT provides evidence that supports the EPA's causality determination. Please also include a description of the representative study by the Cadmibel Study group (Staessen, 1992).

Reproduction and Developmental Effects

The one area I have questions about is on the causality determination of gestational lead exposure and preterm births.

The 2013 ISA review concluded that lead exposure was "suggestive of a causal relationship" with pregnancy outcomes. In my opinion, new studies published since the last review support a causal relationship for lead exposure with preeclampsia and preterm birth. I encourage EPA staff to review the evidence, focusing especially on high quality studies with lower gestational blood lead concentrations. Inclusion of novel studies, like Linda Bui's study of deleading of NASCAR, would also be critical to describe.

Lead is a causal risk factor for preeclampsia, a disorder of severe hypertension in pregnant women. In a meta-analysis – a study of several high-quality studies – Arthur Poropat found that higher concentrations of lead in the blood of pregnant women was a risk factor for pre-eclampsia (Poropat, 2017). For every 1 µg/dL (10 µg/dL or 10 ppb) increase in blood lead in pregnant women, the risk of pre-eclampsia rose by 1.6% (Poropat, 2017). A review of the literature will show other studies published since the last ISA that support the results of this meta-analysis.

Lead is a causal risk factor for preterm birth (Taylor, 2014; Li, 2017; Vigeh, 2011). In a pregnancy and birth cohort study in Bristol, England, pregnant women with a blood lead > 50 ppb) were 1.9-fold more likely to give birth preterm (Taylor, 2014). In the China-Anhui Birth Cohort Study with a mean blood lead of 1.5 µg/dL (15 µg/L), Jun Li found that the risk of PTB was elevated in those with moderate (11.8-17.9 µg/L; OR=2.33, 95% CI:1.49, 3.65) and high (≥16.1 µg/L; OR=3.09, 95% CI: 2.01, 4.76) serum lead concentrations compared with women who had lower exposure (<11.8 µg/L) (Li, 2017). In an Iranian cohort of 348 pregnant women with a geometric mean blood lead of 35 µg/L, Mohsen Vigeh found using logistic regression that higher blood lead concentrations, measured between 8 to 12 weeks gestation, were associated with an elevated odds of preterm birth (OR=1.41, 95% CI: 1.08, 1.84) (Vigeh, 2011).

In a natural experiment study, Linda Bui examined the impact of short-term lead exposure on birth outcomes in 147,673 births following NASCAR's decision to eliminate leaded gasoline (Bui, 2021). After leaded fuel was no longer used, newborns of mothers residing within 4,000 meters of the racetrack gained an average of 104 grams and 0.36 weeks longer duration of pregnancy. The probability of low birth weight declined by 4.1%, preterm births by 2.7%, and small for gestational age by 4.1% (Bui, 2021). The authors concluded that the EPA's National Ambient Air Quality Lead Standard, which is based on a 3-month moving average, failed to protect against risks from short-term exposures.

In the pan-Canadian MIREC birth cohort, Mandy Fisher and her colleagues found that the association of low-level lead exposure and preterm birth was stronger in women with insufficient 25OHD (<50 nmol/L). Of 1,851 live births, 6.1% of births in MIREC were PTBs; 4.9% were spontaneous PTBs. Geometric mean blood lead concentrations were 6.2 µg/L or 6.2 ppb. Fifteen percent of women had insufficient 25OHD concentrations. For all women, a 10 µg/L increase in blood lead was associated with an elevated risk for PTB (OR=1.48, 95% CI:1.0, 2.2) and spontaneous PTB (OR=1.7, 95% CI: 1.1, 2.6). For women with insufficient serum 25OHD (< 50 nmol/L), a 10 µg/L increase in blood lead was associated with a larger risk for PTB (OR=2.4, 95% CI: 1.0, 5.8) and spontaneous PTB (OR=3.0, 95% CI: 1.2, 8.0)

In another study from the MIREC Canadian pregnancy and birth cohort (n=1,560), Mike Borghese and his colleagues found that third trimester blood lead concentration was associated with the development of preeclampsia. The geometric mean blood lead concentrations were 5.8 µg/L. They found that each doubling of third trimester blood lead concentration was associated with a significantly higher risk of developing preeclampsia (RR=1.54; 95%CI:1.06,2.22).

3d. Please comment on the current organization of this appendix and the decision to develop a single causality determination for cardiovascular effects.

I agree with the decision to incorporate a single causality determination for lead exposure and cardiovascular disease. Still, the EPA needs to be thoughtful about how to calculate separate population attributions of lead as a causal risk factor for blood pressure (or hypertension), atherosclerosis, and coronary heart disease.

3e. Please comment on the current organization of this appendix and the decision to incorporate a separate causality determination for exposure to Pb and total (nonaccidental) mortality.

I agree with the decision to incorporate a separate causality determination for exposure to lead and total mortality.

Executive Summary and Integrated Synthesis

ES-17, lines 5-10: "Although slope factors increase with decreasing air Pb concentration, it is possible that the contribution from non-air exposure pathways may lead to the higher slope factors at lower air concentrations. In other words, in older studies in which leaded gasoline or local sources were a major contributor to air Pb, there may be a greater likelihood of discerning the true effect of air Pb on blood Pb due to relatively less contribution from non-air exposure pathways. However, overarching distinctions between old and new studies should be made with caution given that Pb in all media, not just air, has decreased over time."
I don't agree with this statement. If sources of lead are often correlated (e.g., paint, soil, and airborne lead in densely populated cities), then I would expect confounding would be greater at higher air lead concentrations and higher blood lead levels. In other words, it is possible that the flattening of the relationship of air-lead and blood lead is due to non-air exposure pathways attenuating the relationship at higher levels, not at lower levels.

I encourage the EPA to convert the blood lead measure from µg/dL to µg/L to be more compatible with other heavy metals. Moreover, blood lead levels of 1 µg/dL to 10 µg/dL are 10- to 100-times higher than exposures of pre-industrial humans. A consensus statement of scientists has been submitted to EHP calling for a shift to µg/L (Caravanos, et al. EHP resubmitted).

IS-9.1, lines : The study by Richmond-Bryant and others at the US EPA is one of the most important studies published since the 2013 ISA (Richmond-Bryant, 2014). The steeper slopes at lower air lead concentrations are especially important because most of the population is exposed at lower concentrations. A small increase in risk across large populations result in a large population attributable fraction.

The dose-response relationship found by Richmond-Bryant, et al, study was supported by Sammy Zahran's recent analysis of air lead-blood lead levels at the Reid-Hillview airport (Zahran, 2023). This study and others indicate that the EPA should anticipate air lead concentrations well below the existing standard are hazardous. As such, many more air monitors (i.e., over 1000 monitors) should be deployed to measure airborne lead across the United States. Research to examine the particle size most relevant to lead exposure and toxicity is urgently needed.

IS-11, lines 5-9: It could be true that size distribution of samples collected near roads, industrial sources, in rural locations, and in urban locations within the U.S. have become larger. But how do you know if you only have 100 air monitors measuring airborne lead across the United States and you aren't even analyzing their results? I encourage the EPA to be more careful about extrapolating from the few air monitors in their national lead monitoring network. Indeed, I am not even comfortable calling it a network.

IS-19, lines 23-24: The ISA recognizes that a single measurement of lead in blood cannot be used as a good reflection of exposure. Yet some of the new studies published since the last ISA relied on a single blood lead test. While it is useful to provide a comprehensive review of new studies, it would be more helpful to screen out the low-quality studies and focus the review all the high-quality studies that were critical for the causality determination, even if they were published before the last ISA.

IS-26, lines 37-39: I encourage the EPA to indicate whether a threshold exists in the final summation when it is relevant (e.g., for IQ decrements and cardiovascular disease mortality).

References

Borghese MM, Fisher M, Jillian Ashley-Martin, Fraser WD, Trottier H, Lanphear B, Johnson M, Helewa M, Foster W, Walker M, Arbuckle TE. Individual, independent, and joint associations of toxic metals and manganese on hypertensive disorders of pregnancy. Environ Health Perspect 2023; https://doi.org/10.1289/EHP10825.

Escolar E, Lamas GA, Mark DB, et al. The effect of an EDTA-based chelation regimen on patients with diabetes mellitus and prior myocardial infarction in the trial to assess chelation therapy (TACT). Circ Cardiovasc Qual Outcomes 2014;7:15-24.

Evens A, Hryhorczuk D, Lanphear BP, et al. The impact of low-level lead toxicity on school performance among children in the Chicago Public Schools: a population-based retrospective cohort study. Environmental Health 2015;14(1):21.

Fisher M, Marro L, Arbuckle TE, et al. Associations between toxic metals, vitamin D, and preterm birth in the Maternal-Infant research on chemicals study. Paediatr Perinat Epidemiol 2023 doi: 10.1111/ppe.12962. PMID: 36864001

Lanphear BP, Rauch S, Auinger P, Allen RW, Hornung RW. Low-level lead exposure and mortality in US adults: a population-based cohort study. Lancet Public Health 2018; 3(4): E177-E84.

Hu H, Tellez-Rojo MM, Bellinger D, et al. Fetal lead exposure at each stage of pregnancy as a predictor of infant mental development. EHP 2006;114:1730-1735.

Li, J. et al. Maternal serum lead level during pregnancy is positively correlated with risk of preterm birth in a Chinese population. Environ Pollut 2017;227:484–489.

Lin J-L, Lin-Tan D-T, Hsu K-H, et al. Environmental lead exposure and progression of chronic renal diseases in patients without diabetes. NEJM 2003;348:277-286.

Liu J, Liu X, Wang W, et al. Blood lead concentrations and children's behavioral and emotional problems: a cohort study. JAMA Pediatrics. 2014;168(8):737-745.

Navas-Acien A. Lead and Cardiovascular Mortality: Evidence Supports Lead as an Independent Cardiovascular Risk Factor. NCEE Working Paper.

Poropat AE, Laidlaw MAS, Lanphear B, Ball A, Mielke HW. Blood lead and preeclampsia: A meta-analysis and review of implications. Environ Res 2017;160:12-19.

Richmond-Bryant J, Meng Q, Davis A, et al. The influence of declining air lead levels on blood lead-air lead slope factors in children. Environ Health Perspect 2014;122:754-760.

Ruiz-Hernandez A, Navas-Acien A, Pastor-Barriuso, et al. Declining exposures to lead and cadmium contribute to explain the reduction of cardiovascular mortality in rteh US population, 1988-2004. Int J Epidemiol 2017;46:1903-1912.

Staessen JA, Lauwerys RR, Buchet JP, Bulpitt CJ, Rondia D, Vanrenterghem Y, Amery A, for the Cadmibel Study Group. Impairment of renal function with increasing blood lead concentrations in the general population. N Engl J Med. 1992;327:151–156.

Taylor CM, Golding J, Emond AM. Adverse effects of maternal lead levels on birth outcomes in the ALSPAC study: A prospective birth cohort study. BJOG 2014;10.1111/1471-0528.12756.

Tsuji, M. et al. The association between whole blood concentrations of heavy metals in pregnant women and premature births: The Japan Environment and Children's Study. Environ Res 2018;166:562–569.

van Strijp L, Rooy MV, Serem J, Basson C, Oberholzer Hm. Investigating the effect of heavy metals cadmium, chromium and lead, alone and in combination on an endothelial cell line. Ultrastructural Pathology 2023; https://doi.org/10.1080/01913123.2023.2189986
Vigeh, M, et al. Blood lead at currently acceptable levels may cause preterm labour. Occup Environ Med 2011;68:231–234.

Wasserman GA, Litvak-Factor P, Liu X, et al. The relationship between blood lead, bone lead and child intelligence. Child Neuropsyhchol 2003;9:22-34.

Zahran S, Keyes C, Lanphear BP. Leaded aviation gasoline exposure risk and child blood lead levels. PNAS Nexus 2023 doi:10.1093/pnasnexus/pgac285. PMID: 36712926

Dr. Joel G. Pounds

Appendix 2 - Exposure, Toxicokinetics, and Biomarkers

2a. Please comment on the choice and emphasis of topics for providing useful context for the evaluation of human health effects of Pb in the ISA. Is the current organization of the appendix clear and logical?

The EPA CASAC Team is commended for synthesis of Pb toxicokinetics in Appendix 2! This document is remarkably comprehensive, clearly written, literate, and a cogent. The EPA should feel very good about this Appendix!

Please provide any recommendations to integrate exposure and toxicokinetics information more clearly throughout the appendix?

The exposure-toxicokinetics information is nicely described. I'm not aware of overlooked key publications that would illuminate or modify the exposure-kinetics information.

Several Figures are presented in Sections 2.2 and 2.3 which illustrate simulated Pb in selected compartments (blood, bone, body, etc). These simulations include children to the age of ten years age. The Figure legends state" Simulation based on ICRP Pb Biokinetics Model and cites Leggett 1993. However, Leggett's 1993 model was not adequately scaled for children. Moreover, Leggett provided a fairly detailed assessment of the limitations of his model for children (a) increased fractional deposition from plasma to bone, decreased fractional transfer of Pb from plasma to soft tissue, and (c) an elevated rate of transfer from from nonexchangeable bone to fusible plasma. These parameters were incomplete in the 1993 model because these data were not available. I'm confident that the EPA modelers are aware of the limitations of the Leggett model when applied to children and thus used an updated Leggett code. My question is, which Leggett code was used for these simulations? The cited paper suggests the original code was used. EPA, and others, have implemented the Leggett-1993 code to include incorporation into the AALM, and other modifications for a variety of purposes including, to be more appropriate for children. Appendix 2 should clarify the specific model source code implemented for these simulations.

Also, these simulations define Pb intake as 20 or 120 µg/d. Is this Pb intake respiratory, intestinal, or direct to blood?

Figure 2.8 illustrates the blood, bone, and body Pb with a high (120 µg/d) or very high (4,020 µg/d) Pb intake. This extremely high, daily Pb intake for thirty years, seems unlikely for the 21st century, not very relevant and thus detracts from the flow of the Appendix 2.

Figure 2.8 also illustrates, in part, the slow decline of blood Pb due to high bone Pb stores. This point is not well illustrated by the current figure. This point could be made more evident by (a) zooming in the plot to focus on the ages ~45-60 years, and/or (b) quantifying the rate of decline in blood Pb between

ages 50-70 (the halftime can be easily quantified by fitting the blood, bone, total body curves to a two or three-term exponential equation).

Does the appendix adequately describe air-related and non-air related pathways of Pb exposure?

Yes. The Air Pb-Blood Pb relationship is adequately described.

2b. Please comment on how well Section 2.3 reflects the current state of knowledge of Pb biomarkers and their interpretation as it relates to exposure and dose? Is the focus on blood Pb and bone Pb appropriate given the epidemiological literature largely assesses exposure through these two biomarkers?

Yes, of course, due to the long history of blood and bone Pb as biomarkers of exposure, and the central role of blood and bone to Pb kinetics in the body.

It would be useful to briefly summarize (or refer to elsewhere in the document) the analytical advantages, disadvantages, and challenges to measure blood and bone Pb at ever decreasing levels using the typical analytical methods.

Is there sufficient and accurate discussion of the relationship between blood Pb and bone Pb?

Yes

Are relationships between blood Pb and Pb in soft tissues and urine Pb adequately described?

Yes, the simulations help illustrate the relationships.

2c. Sections 2.5.1 and 2.5.2 discuss empirical models of the relationship between air Pb and blood Pb from recent and older studies. Please comment on the effectiveness of this section to accurately reflect what is known about air Pb-blood Pb relationships. Please provide recommendations on any studies that should receive less or greater emphasis.

Minor editorial suggestions:

- The text relevant to the figures is sometimes a page or two remote from the figure. It would be helpful to italicize or bold "Figure x.xx" in the text to facilitate the reader's comprehension of the text.
- Figure legends 3.1 and 3.2 cite Shadbegian *et al.* 2019 as the source for the two Figures. Shadbegian *et al.* is a very nice paper, but it doesn't include the figure, or even discussion from which the figure could be derived. Is there a better citation for the figure?

Dr. Brisa Sánchez

Congratulations to the writing team for creating the ISA. It is clear that a great deal of work went into it.

Integrated Synthesis

- Overall, the Integrated synthesis is well organized and appropriately represents/summarizes the data included in the appendices
- Page IS-1, Summary box: The last bullet point under human health effects may give the appearance that race/ethnicities and genetic backgrounds are synonymous. As the remaining text in the ISA correctly separates these into a social construct and biological risk, respectively, the summary box should not list these factors within the same parenthesis. The bullet point can be streamlined to "Many population subgroups and different life stages have been shown …".
- Page IS-1, Summary box: The last bullet point under human health effects conflates disproportionate exposure and susceptibility into a combined higher risk label. It may be useful to explicitly separate the two, since the summary box will likely get the highest attention. Separating differential exposure from differential susceptibility is important b/c policies can be put in place to reduce differential exposure.
- Section **2.1.5.4** conflates individual race/ethnicity with place level factors (e.g., urbanicity and segregation). This section should be separated into two, with one specifically focused on place-level factors including segregation.
- Along with revisions on Section 2.1.5.4 on race/ethnicity disparities in exposure (above), Table IS-13 in the ISA should be amended as follows: (a) the row of Race/ethnicity should include that new evidence suggests racial/ethnic differences in BLLs are related to racialized segregation and urbanicity; (b) the row on residential factors should include that new evidence suggests racialized residential segregation as a residential factor associated with higher exposure.

Appendix 4: Cardiovascular

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

- Strengths and limitations of some of the longitudinal studies are not consistently stated in the text. For example, the Malmo study (pg 4-32, lines 4-11), it is not clear what was the rate of lost-to-follow up, making it unclear if the lack of association at follow-up is due to selection bias (healthy survivors). Similarly, one of the other longitudinal studies is among very highly exposed workers, at an exposure range where the non-linear response curve may have already be flattening out. Thus, the very small effect size (even though it was found significant), may be due to the exposure range. Making these clarifications is important, since typically longitudinal studies are conceptualized as having stronger designs, but is not always the case.
- There is lack of clarity on whether some of the effect measure modification results may be due to differential exposures or due to unmeasured confounders/stressors among the sub-populations. Given that the association between lead exposure and outcomes tends to be non-linear for many outcomes, differences in associations may be purely because sub-group exposure is at different

levels for different sub-groups. Hence, if the association is modeled linearly (which is almost always the case) the results appear as though effect modification is present (i.e., different subgroups tend to be at different locations in the non-linear dose-response curve, which results in an approximated linear slope of different magnitude).
    - Related to the point immediately above, the results of Appendix 2 on differential exposure could be brought to the discussion of effect measure modification for health effects.
- The study examining effect modification by allostatic load (figure 4-9) may have some circularity built into it– AL measures often have cardiovascular outcomes as part of the AL calculation. It may be warranted to include a comment along these lines.
- The presentation of results based on NHANES data is very important, given the importance of NHANES in the assessment of US population health. At times, however, the evidence from NHANES appears contradictory. Given the importance of NHANES, it may be useful to explicitly acknowledge or identify potential reasons for differences in results (e.g., different covariates were adjusted for, different NHANES waves).
- Exposure biomarkers have been decreasing, often with decreases in the between-individual variability in exposure. Because reduced exposure variability limits power to detect associations, it may be useful to explicitly acknowledge this reduction in variability over time, and how that may affect study results of more recent research.
- The epidemiological literature has for some time now reserved the use of the words "increases/decreases" for longitudinal and experimental studies where the researcher can observe within-subject change in exposure and used the words "higher/lower" for crossectional human studies where the researcher can only observe between-person differences in exposure. Since the Appendix includes experimental studies and both longitudinal and crossectional human studies, it would be more consistent with the literature if the terminology to describe association/effect estimates became consistent with the study designs of the reported studies.
- Some of the above comments also apply to other health effects Appendices.

Typos or other issues

- Sections 4.1.1 and 4.1.2 make reference to section 4.1.2, but should probably be 4.12
- X-axis in figures 4-7 and 4-8 include negative values, which seems strange given the label of the x-axis
- Figure 4-14 appears to be missing association estimates from the 4th study listed in the figure (stratified by race/ethnicity?)
- Line 2 of pg. 4-16 states "outcomes", should it be "exposures", "covariates" or "factors"

3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?

- The mechanistic studies strongly support biological plausibility. These sections are a strong contribution and essential to the causality determination. On the whole, these studies along with the consistent evidence from epidemiological studies among humans support the causality determination.

3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

- It is not always clearly stated how or why the findings of effect measure modification play a role in causality determinations. It would be useful to clarify this in the section where the effect modification results are presented, and in the causality determination section.

3d. Please comment on the current organization of this appendix and the decision to develop a single causality determination for cardiovascular effects.

- Appendix 4 is organized well.
- A single causality determination for cardiovascular effects is consistent with newer ISAs, and seems appropriate for cardiovascular outcomes given the interrelationship among the CVD outcomes studied.

Other human health appendices

- The comment in 3a. above regarding word choices used to describe longitudinal vs. crossectional studies (e.g., "increasing" vs. "higher"), also applies to other health appendices.

Dr. Brian Schwartz

I reviewed relevant sections of Integrated Science Assessment for Lead, External Review Draft, March 2023. The filename was "Draft Pb ISA_Combined_HEROnet.pdf." Appendices were as follows: 3 = *Nervous System Effects* (file pages 402-941); 4 = *Cardiovascular Effects* (file pages 943-1115); 5 = *Renal Effects* (file pages 1116-1204); 6 = *Immune System Effects* (file pages 1205-1294); 7 = *Hematological Effects* (file pages 1295-1338); 8 = *Reproductive and Developmental Effects* (file pages 1339-1604); 9 = *Effects on Other Organ Systems and Mortality* (file pages 1605-1770); and 10 = *Cancer* (file pages 1771-1823).

I first reviewed **Appendix 12** (file pages 2081-2120), *The Process for Developing the Pb Integrated Science Assessment*. The process for Health Studies was in Section 12.4.3 (page 12-6 to 12-9); Section 12.5 on literature search and study selection (primarily 12.5.1.3 on Health); Section 12.6 on Study Selection and Study Quality (primarily 12.6.1.1 on Health, notable phrase *"As described in the Preamble, causality determinations are informed by integrating evidence across scientific disciplines (e.g., exposure, animal toxicology, epidemiology) and related outcomes, and by judgments of the strength of inference in individual studies. For health outcomes, study quality is evaluated using a uniform approach that considers study strengths and limitations, including the possible roles of chance, confounding, and other biases that may influence results."*); and Table 12-5 on scientific considerations for evaluating strength of inference from studies on health.

Some general **comments** about the approach in Appendix 12 and comments relevant to the health appendices overall:

- In **Table 12-3**, it is unclear how different study designs were considered to give evidence of a range of quality levels. The study designs are listed and seem to be equally weighted. However, evidence provided by cross-sectional and other designs vs. longitudinal studies should not be equally weighted. Longitudinal studies allow inferences about reversibility, persistence, and progression of effects and overcome several weaknesses regarding temporality, confounding (particularly important), and other threats to the validity of non-longitudinal studies. I would recommend that these issues are more directly and explicitly addressed.
- I think evidence should be hierarchically used: human > animal > exposure alone; longitudinal > other designs; biologically sound associations > more difficult to explain (e.g., related to comments below, we might expect a study to report that blood lead was associated with systolic blood pressure at cross-section but not with cognitive decline over time, and if the study reported the latter, there should be a strong explanation for it). The studies of kidney function that were longitudinal and showed effects of chelation should be more highly weighted. Studies that discover effect modification in hypothesized directions should be more highly weighted, because while confounding is a nuisance that should be removed, effect modification can often inform possible mechanisms.
- I would like to see more direct consideration of how associations with lead biomarkers are interpreted differently depending on the biomarker. How should associations with blood lead, patella lead, tibia lead, cumulative blood lead index, and other biomarkers be interpreted in terms of acute effects, chronic effects, recent dose, and cumulative dose. In our work, we have written that blood lead allows inferences about the acute effects of recent dose that are more likely to be reversible; and that tibia lead allow inferences about the chronic effects of cumulative dose that are more likely to be irreversible. How should patella lead associations be interpreted, especially

in studies that did not find other lead biomarker associations? The documents should explain how tooth lead is interpreted, generally as a measure of prenatal and early natal lead exposure. More consideration should be given to how past exposures should be integrated into much later health studies, depending on the biomarkers that were evaluated.

- I would like to see more direct consideration of exposures during the life course, and how earlier life effects could have significant implications for health over time. An earlier life exposure could cause a health impact that progresses over time, or that itself is a risk factor for other health impacts (e.g., lead → hypertension → cerebrovascular disease). There are few studies that formally addressed these questions, but this should be written to indicate we understand that this trajectory can occur so our conclusions about lead's life course health impacts are probably under-estimated.
- In the prose, rather than the tables, about specific studies, sample sizes are not provided, so large confidence intervals cannot be easily interpreted.
- Emissions and environmental redistribution from legacy lead sources and uses seem particularly problematic. These create very localized risks that are not, in my view, well accounted for in the EPA consideration. These kinds of emissions, the resulting human lead doses, and the connected health effects are obscured by the "average" relations between lead in soil and lead in air, lead in air and lead in blood, and lead-related health effects that have been generally relied upon. This is acknowledged in some locations in the ISA but perhaps should be separately emphasized. Recent papers have discussed these issues (e.g., Caballero-Gómez H, et al., *GeoHealth* 2022; O'Shea MJ, et al., *IJERH* 2021). Urban redevelopment, demolitions of buildings, legacy soil contamination, and related activities and sources may need focused attention and regulation.
- Throughout, I am a bit uncertain about why some pre-2006 and pre-2013 studies are considered, discussed, and weighed and others are not. Is this explained somewhere that I missed? Old studies are frequently referred to as "recent" (for studies from 2013 or earlier). Similarly, although it is stated that occupational evidence is not of primary interest and would not generally be considered, it is covered many times throughout the health outcome appendices.
- The bodies of scientific literatures are so extensive and compelling that consideration should be given, based on science alone, to the elimination of new, controllable environmental discharges of lead. Given that most biases and measurement errors are likely to lead to underestimation of effects, it would seem that the quantitative exposure and risk analyses for policy making may also be likely to underestimate overall impacts. This should be clearly stated.
- As discussed further below, I believe "interaction" is often sub-optimally written about when such analyses were in papers considered across the appendices. In addition, I generally believe that "effect modification" or "moderation" is a better term to use.

Appendix 3, Nervous System Effects. This appendix had seven causality determinations for nervous system effects ascertained during childhood, adolescent, and young adult lifestages; and four additional causality determinations ascertained during adult life stages, tabulated on file page 413, document page 3-1.

- Nervous System Effects Ascertained during Childhood, Adolescent, and Young Adult Life Stages
    - Cognitive Function: causal
    - Attention, Impulsivity, and Hyperactivity: causal
    - Conduct Disorders: likely to be causal
    - Motor Function: likely to be causal

    - Anxiety and Depression: likely to be causal
    - Sensory Function: suggestive of, but not sufficient to infer, a causal relationship
    - Social Cognition and Behavior: suggestive of, but not sufficient to infer, a causal relationship
- Nervous System Effects Ascertained during Adult Life Stages (Section 3.6 and following, document page 3-171, PDF page 583)
    - Cognitive Effects: likely to be causal
    - Psychopathological Effects: likely to be causal
    - Sensory Function: suggestive of, but not sufficient to infer, a causal relationship
    - Neurodegenerative Disease: suggestive of, but not sufficient to infer, a causal relationship

I had the following general comments about the determinations in adults:

- I think the causality determination for adults should be causal. There are several, large longitudinal studies of lead and cognitive function and decline in adults that examined associations with both blood lead and bone lead, and found, as would be hypothesized, more consistent associations with cumulative dose than with recent dose. In my view, this is strong, and difficult to come by, evidence.
- The lack of strong consistency in associations of lead biomarkers with cognitive domains and cognitive domain scores should not be considered evidence in favor of a lower causal determination. There are many reasons that there could be instability in associations with domains across studies, including many individual characteristics that are likely to bias associations towards the null.
- The 2013 Pb ISA concluded that there was likely a causal relationship for cognitive function in adults, and the current Pb ISA states "Recent prospective cohort studies with longer follow-up periods, multiple and repeatedly measured cognitive outcomes, and adjustment for multiple risk factors and confounders reduce uncertainties and strengthen the overall evidence related to the association of Pb exposure with cognitive function in adulthood." It seems the lack of stronger toxicological evidence is reducing the overall confidence for a causal determination (and other features reviewed on pages 3-178 and 3-181), but the human evidence should be more strongly weighted compared to animal evidence. Longitudinal studies also reduce concerns about confounding.
- It is also stated "This suggests that early life Pb exposure contributes to cognitive dysfunction that persists into adulthood, which is new evidence in this review." This is important evidence for a causal conclusion in childhood, as persistent (and progressive) effects may be less likely to be due to confounding but seems to be giving pause about separate causal conclusions in adults. I find this distinction artificial; lead exposure occurs across the lifespan and determining critical exposure periods, latency and duration of exposure and effects, the shape of the C-R function, and other features seems to be a technical exercise that is demanding too much of existing studies, and not needed for causal inference. Studies in adults by my research group, albeit occupational exposures, that distinguished cross-sectional, historical, and longitudinal effects, I believe provide evidence that adult cognitive effects are due to adult exposures.
- Similar arguments could be made about the causal determination for psychopathological effects, but I feel less strongly about this one.
- Small point: page 3-209, line 3, amyotrophic lateral sclerosis is not a proper noun.
- I do not find it a compelling argument that MMSE is useful to identify Alzheimer's disease.

- The statement that tremor is the most common indicator of neurological disease is not referenced and seems inaccurate to me (page 3-212, line 26). All sorts of tremors are benign and indicative of nothing of concern.
- Throughout, I am concerned about conflation regarding the shape of the C-R function and the issue of causal inference. Does the document imply that the former is necessary for the latter?

Regarding the specific questions for Appendix 3:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

For childhood, adolescent, and young adults:

3a: I think this is appropriate.

3b: I think this is well done.

3c: I agree with the causality determinations.

For adults:

3a: I think this is not unreasonable. I understand why occupational studies were excluded (not all actually are), but I believe very strong longitudinal evidence is provided by the papers of the Korea Lead Study (co-authors Brian S. Schwartz and Byung-Kook Lee on all these papers, which generally were published before 2010 [especially *Epidemiology* 2005]). I think this evidence, when combined with other more recent evidence, supports a causal determination in adults.

3b: I think the committee should discuss how important it is to have evidence from mechanistic studies if the human evidence is compelling.

3c: With the one caveat above regarding the causality determination for cognitive effects in adults, I think this is reasonable.

Appendix 4, Cardiovascular Effects. A single causality determination was made for Cardiovascular Effects as "causal." The Review Panel was given a specific question about the single causality determination, in view of the observation that cardiovascular effects are inter-related, either through biological pathways or due to common causes. The 2013 Pb ISA made four cardiovascular disease causality determinations, for hypertension, subclinical atherosclerosis, coronary heart disease, and cerebrovascular disease. A causal determination was made for hypertension and coronary heart disease.

The single cardiovascular causal determination in the Pb-ISA is consistent with the 2019 Particulate Matter and 2020 Ozone ISAs.

I had a few general comments:

- The studies of cardiovascular outcomes seem to have much more commonly evaluated effect modification. The appendix seems to describe these analyses and results the way the authors did, and hence there is considerable heterogeneity on how these analyses are described and interpreted. I wonder whether an overview section about all the effect modification studies would be helpful. Define effect modification and how evaluation of it can help; explain all the terms used by authors to describe what are essentially the same kinds of analyses (e.g., interaction, effect modification, moderation, "modified," effect measure modification, subgroup analyses, stratified analyses); and describe all the effect modifiers that have been evaluated in the lead-CVD literature and how several of them could be particularly helpful to causal inference.
- I would recommend that the consideration of effect modification not just be about identifying higher risk subgroups, but also consider how effect modification may inform causality determinations, biological plausibility, and mechanistic pathways. I would recommend that there be more consistency in describing why such analyses were done, how they can be helpful, how they are presented, and how they are interpreted.
- One suggested edit: page 4-3, line 9, I believe the word "consistent" would be better than "coherent."
- Page 4-8, section 4.3.1. I was surprised that the earlier paper by Martin D, et al., *American Journal of Epidemiology* 2006, wasn't discussed. This was a paper of community dwelling older adults that measured both blood lead and tibia lead and modeled systolic and diastolic blood pressure and hypertension.
- Relating to a prior comment, several figure titles and text use the term "interactive effect." In other places, this is written as subgroup analyses or stratified analyses. In other places the idea is not even conveyed (e.g., Figure 4-11 title: "Relationship between blood Pb levels and systolic blood pressure by sex and age, Canadian Health Measures Survey"). I believe these figures and sections would be better written consistently and be formally about evaluation of effect modification. "Effect modification of relation between blood lead and systolic blood pressure by race/ethnicity," for example. Figure 4-21 is also about effect modification by age of the blood lead association with pulse wave velocity. There are <u>many</u> other examples.
- Section 4.3.1.1.1 refers to Effect Measure Modification by "a variety of outcomes." These effect modifiers are not outcomes.
- For Figure 4-9, I would recommend that the measurement of allostatic load be defined in a footnote, legend, or title. It is not even described in the text. This is effect modification of the relation between blood lead and blood pressure by allostatic load. This can give biologic and mechanistic insights, can it not? Why was this an analysis of interest? How has allostatic load been measured and did this study use the best approach?
- I have similar concerns about the presentation of the Peters et al. 2007 study on self-reported perceived stress, how it is described, how it is presented, and how it is interpreted. What is stress? Does lead cause stress? How does evaluation of effect modification by perceived stress of the tibia lead and SBP relation contribute to our thinking about this literature? This paragraph on page 4-24 needs to be rewritten in my view.
- Page 4-35, lines 7 to 9, "further modified by sex" is not optimally written. Why evaluate effect modification if there is no main effect?

- Page 4-36, why is BMI an interesting effect modifier to evaluate? Was BMI pre-specified as an effect modifier or was this exploratory and post-hoc? So much evaluation of effect modification in the epidemiologic literature on lead is exploratory and post-hoc. Is that a problem?
- Like in other sections, there are many references to "recent" studies that are not very recent, throughout Appendix 4.

Regarding the specific questions for Appendix 4:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?
3d. Please comment on the current organization of this appendix and the decision to develop a single causality determination for cardiovascular effects.

3a: Please see the comments above.

3b: I believe this is acceptable.

3c: A single causality determination was made for cardiovascular effects as causal. I agree with this determination.

3d: I agree with the approach that was used. Disentangling and assigning different causal conclusions to different, interconnected cardiovascular outcomes, seems like an impossible, and possibly arbitrary, exercise.

Appendix 5, Renal Effects. A single causality determination was made for renal effects, as "Reduced Renal Function," as causal. This is an upgrade from the 2013 Lead Integrated Science Assessment, which was judged to be "suggestive of a causal relationship."

I had the following general comments:

- I continue to find inconsistencies with how effect modification is written about and interpreted. For example, the text interpreting Figure 5-1 notes that associations were stable after stratification. I assume this stratification was done to evaluate effect modification, not confounding, because p-values for "interaction" were presented. The effect modification evaluation was null, so why are these results important? The figure title is "Modification of association ...," inconsistent with the way other such figures were titled. I would recommend more consistency. Similar issues are present in Figure 5-7, labeled "subgroup analyses" with "interaction" p-values presented, almost all of which were null.
- This chapter highlights the problem of not being clear about the hierarchy in the quality and causal determination relevance of study designs. For example, there are papers by Ja-Liang Lin

from 1999 and the following years that randomized persons with chronic renal insufficiency and "mildly elevated" (authors' term, assessed by provocative chelation with EDTA) body lead burden to two months of chelation and followed renal function over time. I think the evidence from this type of study – randomized, prospective, tried to estimate cumulative dose – should be more heavily weighted.

- Minor issue: type 1 and type 2 diabetes are not proper nouns, so the 't' need not be capitalized. In the document, sometimes it is, and sometimes it is not.
- The consideration of studies in this section illustrates my concern about how human evidence is evaluated, seemingly weighting different designs equally. My concerns are mitigated by the overall causal determination.
- I would refer to "persons with diabetes," and not use the term "diabetics" to refer to people. A global search should be done because I believe this is present in other chapters as well.
- I continue to remain confused about why some pre-2013 and pre-2006 studies are considered, discussed, and weighed and others are not.
- Section 5.8, Reverse Causality, is particularly important and nicely done.
- Page 5-32, line 1, how is eGFR different from creatinine clearance?

Regarding the specific questions for Appendix 5:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: Subject to the comments above, and the fact that the determination is causal, I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I agree that the causality determination should be causal.

Appendix 6, Immune System Effects. Three causality determinations are made: 1) Immunosuppression = "likely to be causal"; 2) Sensitization and Allergic Responses = "suggestive"; and 3) Autoimmunity and Autoimmune Disease = "inadequate."

I had the following general comments:

- Species names are not capitalized (e.g., page 6-6, lines 5 and following). This is recognized later (page 6-10, line 1), where *monocytogenes* is not capitalized, but later "*Coli*" is capitalized, probably due to the errant period.
- Hepatitis B is not a proper noun.

- I believe the consideration of immune system effects is nicely done given the complexity of this literature, the immune system, its many measures, its many functions, and its inter-connections.
- The Integrated Summary sections (Sections 6.3.3, 6.4.3, and 6.5.3) are nicely done, as are the Biological Plausibility sections (Sections 6.6.1 and 6.6.2).

Regarding the specific questions for Appendix 6:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I agree with the three causality determinations that were made.

Appendix 7, Hematological Effects. A single causality determination was made, for Altered Heme Synthesis and Decreased Red Blood Cell Survival and Function as "causal."

I have no general comments on this chapter.

Regarding the specific questions for Appendix 7:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I agree with the single causality determination that was made.

Appendix 8, Reproductive and Developmental Effects. There were four causality determinations for this Appendix: 1) Effects on Pregnancy and Birth Outcomes = "suggestive"; 2) Effects on Development = "causal"; 3) Effects on Female Reproductive Function = "suggestive"; and 4) Effects on Male Reproductive Function = "causal."

I had the following general comments for this appendix:

- There are several studies in the past 2-3 years that provide strong evidence in these categories. Some may have only appeared after the publication cutoff year for paper inclusion.
- Seven separate categories of pregnancy and birth outcomes were considered (page 8.5). There were several specific health outcomes that could be within these seven categories. This seems comprehensive.
- Effects on Development endpoints considered postnatal growth, bodyweight, and stature; puberty among females; puberty among males; and other developmental effects. I think this is sufficiently comprehensive.
- Effects on Female Reproductive Function considered hormone levels, menstrual cycle, female fertility, and morphology and histology of female sex organs. I think this is sufficiently comprehensive. However, there are several recent studies that have very strong study designs that should be weighted more heavily. A study that leveraged the removal of lead from NASCAR fuels as a natural experiment in relation to birth outcomes provides strong causal evidence. This seemed to appear after the date cutoff for inclusion of studies, but earlier versions may be available (Clay K, Hollingsworth A, Severnini E. The Impact of Lead Exposure on Fertility, Infant Mortality, and Infant Birth Outcomes. IZA Institute of Labor Economics 2023). There are numerous other studies from this group using the same design that could be considered for other health outcomes as well. Another study leveraged a natural experiment provided by lead in water in Newark in 2016 and found evidence that should be considered very helpful to causality determination (Dave DM, Yang M. Lead in drinking water and birth outcomes: a tale of two water treatment plants. *J Health Economics* 20220.
- Effects on Male Reproductive Function considered sperm and semen production, quality, and function; hormone levels; male fertility; and morphology and histology of male sex organs. This is sufficiently comprehensive. I believe that effects on female reproduction should be stronger than "suggestive." I am unclear how the existing evidence was used to land on suggestive rather than likely to be causal.
- Overall, this is a complex appendix that is presented well.

Regarding the specific questions for Appendix 8:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I believe that two causality determinations should be reconsidered – birth outcomes and female reproduction – as discussed above.

Appendix 9, Effects on Other Organ Systems and Mortality. A total of eight causality determinations were made in this appendix: 1) Hepatic Effects (direct measures of liver injury, serum biomarkers of liver function, serum lipids) = "suggestive"; 2) Metabolic Effects (diabetes mellitus and insulin resistance, metabolic syndrome, body weight in adults) = "inadequate"; 3) Gastrointestinal Effects = "inadequate"; 4) Endocrine System Effects = "inadequate" (regarding thyroid hormones, cortisol/corticosterone, and vitamin D levels); 5) Musculoskeletal Effects (bone mineral density, osteoarthritis, oral health [periodontal disease in adults, dental caries in children]) = "likely to be causal"; 6) Effects on Ocular Health = "inadequate"; 7) Respiratory Effects = "inadequate" (in populations without asthma); and 8) Mortality (all-cause, cause-specific) = "likely to be causal" (for all-cause, nonaccidental, mortality).

I had the following general comments for this appendix:

- I am uncertain about the mechanism of osteoarthritis mediated through depressed protein synthesis (Figure 9-2).
- It seems that the range of musculoskeletal effects that were considered is too large and heterogeneous to give a single causality determination to musculoskeletal overall. This would seem to depend on how strong the evidence is for the potential biologic pathways linking what seem to be disparate outcomes.
- The causal determination in the table on page 9-1 for mortality is "likely to be causal" but on page 9-71 this is written as a causal relationship.
- I believe the study by Hollingsworth and Rudik (*Am Economic J: Economic Policy* 2021) is compelling evidence for mortality, exploiting regulatory exemptions as a quasi-experimental design. This highlights the strength of certain study designs.

Regarding the specific questions for Appendix 9:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I agree with the eight causality determinations that were made, with the caveats and inconsistency explained above.

Appendix 10, Cancer. This appendix had a single causality determination, Cancer Incidence and Mortality as "likely to be causal."

I had the following general comments:

- The somewhat different approach that this appendix took to consideration and evaluation of evidence seems reasonable. It started with mechanistic pathways and markers of carcinogenesis, first evaluating animal models of carcinogenicity, then evidence for genotoxicity (human and non-human studies), oxidative stress, cell viability, cytotoxicity, apoptosis, DNA damage repair enzymes and gene expression, epigenetic regulation of gene expression, gene expression and extracellular matrix, and inflammation. It then proceeded to cancer incidence and mortality, with overall cancer incidence, overall cancer mortality, studies of lung cancer, studies of brain cancer, studies of breast cancer, and studies of other cancers.
- IARC considers lead to be a class 2A carcinogen, "probably carcinogenic." The EPA determination may appear to be consistent with that one. However, IARC class 2A is assigned when there is limited evidence of carcinogenicity in humans AND sufficient evidence of carcinogenicity in experimental animals OR strong evidence that the agent exhibits key characteristics of carcinogens. It is thus surprising that the causality determination is for cancer incidence and mortality.

Regarding the specific questions for Appendix 10:

For the health appendices:

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.
3b. What are the Panel's views on the integration of evidence from mechanistic studies to inform conclusions on biological plausibility?
3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

3a: I think the consideration of studies is acceptable.

3b: I think the integration of mechanistic studies is acceptable.

3c: I would like to discuss the wording of the causality determination that was made. I think a causality determination of lead as a carcinogen as causal is supportable, but not with cancer incidence and mortality.

Dr. Elizabeth A. (Lianne) Sheppard

Comments on the epidemiologic evidence for cancer mortality discussed in Appendix 10

I disagree with EPA's summary of the cancer mortality epidemiologic evidence as inconsistent, both with respect to the evidence extracted from the 2013 ISA as well as the newer evidence reported in the 2023 Pb ISA. This is documented in Table IS-11. I conclude that the epidemiologic evidence is far from inconsistent.

EPA claims that the evidence cited in the 2013 ISA is inconsistent. This is based on two NHANES III studies: Menke et al 2006 and Shober et al 2006. The cited results are a hazard ratio (HR) of 1.10 (95% CI .82 – 1.47) for BLLs ≥ 3.63 μg/dL vs <1.93 μg/dL (Menke et al 2006) and 1.44 (95% CI: 1.12, 1.86) for BLLs of 5–9 μg/dL compared to < 5 μg/dL (Shober et al 2006). I note that while their categories are not the same, both studies report a range of elevated HRs and report overlapping elevated HRs in the range of 1.12 to 1.47. While the Shober study estimates suggest elevated BLLs convey excess risk, the Menke study estimates are consistent with a wide range of risks, including that the higher blood Pb category (with lower bound below the upper bound for lowest category in Shober) offers nearly 20% protection up to that it confers nearly 1.5 times the risk compared to individuals in the lowest BLL category. These studies have considerable overlap in their confidence intervals and should not be characterized as being inconsistent merely because they differ with respect to statistical significance.

EPA also claims that the new epidemiologic study evidence available since the 2013 ISA is inconsistent. This is based on three studies: two US studies from NHANES data (van Bemmel et al 2011 (also NHANES III) and Duan et al 2020 from a more recent time period) and one Korean NHANES study (KNHANES, Byun et al 2020). The van Bemmel study among 3,223 individuals greater than 40 years of age was described as "null" based on their reported HR of 1.083 [95% CI: 0.983, 1.194] for BLL ≥ 5 μg/dL vs <5 μg/dL. EPA fails to note that given its lower CI bound of 0.98, that this study is consistent with a toxic effect of Pb where a protective effect is unlikely. The more recent study with approximately the same duration of follow-up, larger sample size (n = 26,056), and larger age range (ages 20+) reported a HR of 1.47 [95% CI: 1.22, 1.78] per 1 unit increase of BLL for a distribution of BLLs with a median of 1.49 μg/dL and a 75th percentile of 2.31 μg/dL. The KNHANES study reports larger HRs for Pb the second and third tertiles (relative to the first) with low exposure levels (i.e., the top tertile has a minimum BLL of 2.71 μg/dL). Based on the distinct characterization of BLL increments in these studies, some additional work is needed to better align the HR reporting and to address the apparently large range of BLLs in van Bemmel in contrast to Duan. Nonetheless, all three studies report HRs that are consistent with a toxic effect of Pb. Thus, they are all consistent with reporting an adverse association of higher BLL Pb with cancer mortality.

My interpretation is that this evidence is consistent. While two of the five studies report lower CI HR estimates below 1, the lack of statistically significant estimates from these two studies should not dominate the evidence synthesis. All of the studies report evidence that suggests that elevated Pb is consistent with a toxic effect and three of the studies report that elevated Pb is associated with increased mortality even at the lower bound of its 95% CI. Whether this consistent evidence should lead to the conclusion that there is elevated total cancer mortality associated with Pb exposure, including for exposures well below 5 μg/dL, is a different question. To reach that conclusion requires consideration of additional information, such as whether blood Pb is an appropriate biomarker for cancer outcomes, and whether reverse causality may be a factor in the epidemiologic study findings.

Overall, I believe that EPA is misusing the concept of "inconsistent", basing their judgement almost exclusively on statistical significance, without consideration of the increment of BLL in the HR, or the width or overlap of the confidence intervals. This is particularly evident in their characterization of the van Bemmel study as a null finding. I recommend that in all future distillations and summaries of epidemiologic evidence, EPA integrate the guidance provided by Cummings et al (2004)[1] into its evidence syntheses.

Reference

1. Cummings P, Rivara FP, Koepsell TD. Writing Informative Abstracts for Journal Articles. *Arch Pediatr Adolesc Med*. 2004;158(11):1086-1088. doi:10.1001/archpedi.158.11.1086

Dr. William Stubblefield

Welfare Effects of Pb Exposure

Appendix 11 evaluates the available welfare effects evidence and presents causality determinations for Pb effects in terrestrial, freshwater and saltwater systems. Within the scope and context of the Draft Pb ISA to synthesize the most policy-relevant science to inform the review of the NAAQS for Pb, please evaluate if the welfare effects sections adequately address the following considerations for terrestrial, freshwater, and saltwater biota and ecosystems.

Introduction (Section 11.1)

The welfare effects appendix has an introductory section that includes concepts and tools for evaluating Pb effects on organisms and ecosystems. To what extent do the choice and emphasis of topics in the introduction provide adequate context for the evaluation of ecological effects of Pb in the ISA?

Appendix 11 Introduction provides an excellent synopsis of the available toxicity data and the approach used in the lead ISA assessment. The document is well-written and well-organized and does an adequate job of summarizing the "new" published ecotoxicolgical data (post-2013). The organization of the document based on endpoints and biological complexity is logical and appropriate and addresses the levels of uncertainty with the various organizational levels. Additional discussion/justification for the application of the literature cut-off values may be useful and warranted. In some cases, the cut off values were very close to the current criteria/standards. For example, the aqueous freshwater cutoff was 10 µg/L while the current AWQC is 2.5 µg/L at 100 hardness. Section 11.1.4 provides a good description of how sub-organismal, organismal, and ecosystem level endpoints relate to each other and how they are addressed in the assessment. This section also makes a very important point, i.e., the difficulty of attributing observed effects in the field to Pb (or even more difficult…atmospheric Pb). This point must be stressed in the later sections. Section 11.1.6 provides a good overview summary of the importance of Toxicity Modifying Factors (TMF) in the various matrices and this reflects the current "state-of-the-science" for evaluating bioavailability and predicting toxicity. This is a major improvement over past evaluations and reflects the recognition of all the research that has been developed in the past decade and the value of EPA/Industry cooperative research agreements, like the metals CRADA. I really appreciated Section 11.1.7, this provides an excellent synopsis of criteria/standards for each matrix from different locations around the world and reflects the state of the science for the jurisdictions. It is important to note that much of the recent research and toxicity model development have come from Europe, Australia, and Canada rather than the US; perhaps there could be benefit to international research cooperative arrangements.

One concern was noted, that is that there does not seem to be a clear discussion of environmental fate and source apportionment for Pb. This would seem to be important and necessary if the ultimate goal is to establish air quality criteria. Much of the lead in the environment today no doubt comes as a consequence of terrestrial runoff, direct discharge from industry or publicly-owned treatment facilities, and/or transport from non-point source into freshwater systems. Atmospheric input is no doubt important and will influence sources like terrestrial runoff, but some discussion about the environmental fate processes and the contribution of atmospheric sources is important and should be addressed somewhere in the document.

Terrestrial (Section 11.2)

Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.2 provides a good summary of the available data and new data (post-2013). Available new data seem to support previous observations and provides no basis for heightened concerns, thus the justification for the conclusion that no changes to the previous causality determinations is adequate and appropriate.

Regarding the question about the relationship between Pb exposure and effects on individual organisms and ecosystems, the statement in section 11.2.1 summarizes the situation quite well:
"Although evidence for effects on growth, reproduction, and survival at the individual organism level and in simple trophic interactions makes the existence of effects at higher levels of organization likely, direct evidence is relatively sparse and difficult to quantify. The presence of multiple stressors, especially including other metals, continues to introduce uncertainties in attributing causality to Pb at higher levels of organization." The conduct and interpretation of community level toxicological studies is extremely difficult due to the number of experimental variables. This is especially difficult in conducting tests in a real-world system (non-laboratory) or an uncontrolled field-based assessment. The characterization of "exposures" is very difficult especially if the goal is to ascribe observed environmental changes to a single contaminant (Pb; there are almost always multiple metals in real-world exposures) and even more difficult, if you want to apportion those exposures to atmospheric Pb sources verses waterborne or other fate pathways.

The complex relationship between atmospheric Pb and the deposition and subsequent reactions that can affect Pb speciation, bioavailability, and toxicity is an area that needs more investigation. As was noted in the text, investigations using Pb-spiked soils are difficult to interpret and may ultimately contribute very little to real-world exposures because the methods may not duplicate natural processes. Similar complexities have been observed in aquatic sediment investigations with other metals (i.e., Ni) and test procedures have been developed to artificially reproduce those processes that occur under natural conditions making it possible to conduct meaningful laboratory metal-spiking investigations (Besser et al. 2011).

I am not currently aware of terrestrial studies that have been conducted that should be considered at this time.

Freshwater (Section 11.3)

Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.3 provides an excellent description of the aquatic research that has been conducted since the 2013 ISA effort. It is well-organized and comprehensive. Overall, I like the presentation of the data followed by the summarization of the salient findings. Substantial new data were identified and included. The authors should be praised for the inclusion of the toxicity modifying factors and bioavailability section (11.3.2). A large amount of new data has been developed since the 2013 ISA effort and it was appropriately captured in the section.

The authors are to be commended for the level of effort that must have been required to produce these sections. That said, the report is in many places so detailed and inclusive that the reader can easily get lost and forget the reason for the discussion. In many ways this document is too detailed for the purpose, i.e., to address salient new data that affects the conclusion of causality. Sections like 11.2.2.1.10 "Parasites" or 11.2.2.1.11 "Bioturbation/association with sediment" are interesting, but have little to do with the issue at hand. Many of the sections like most of 11.3.4 are very useful and might well be better served as a contribution to an update to the 1984 Ambient Water Quality Criteria (AWQC) document rather than in this ISA. Some of the sections discuss large amounts of data that ultimately may be of little importance, for example, while Section 11.3.4.1 "Effects on freshwater microbes" (an area not currently considered in developing AWQC or state standards) provides 3 pages of text, it is ultimately summarized by concluding that the data are negative or uninterpretable. Relating the importance of the endpoints to the conclusion of causality, as discussed in Section 12.6.1 would be helpful.

Although, as I said, I like the general organization based on organism type (plants/ invertebrates/vertebrates) followed by biological complexity (suborganism/organism level) it makes the document very difficult to read and follow. Perhaps presenting the data in tabular form, like Table 11-5, without the detailed discussion would help. I also notice that in several places reference is made to studies that are based on "nominal" or unmeasured concentrations; I suggest that these studies be rejected and not further considered and these criteria should be equally applied across all of the data considered in the consideration of causality. In fact, I wonder what the standards for acceptability for any of the reported studies were. Is there a review procedure available that states what the minimum criteria for acceptability of a study is? These should be included in this document and referenced as the standard by which the review was conducted. Some minimum standards for relevance and reliability should be documented and applied. I believe that guidelines for study acceptability are used in both OPPT and the Office of Water.

Section 11.3.6 Community and Ecosystem effects does a good job of identifying new studies that were not previously considered (post 2013). No clear findings were reached from the studies, as stated they "reported either negative, positive, or null associations between sediment or porewater Pb concentration and community and ecosystem effects." No mention was made as to how the study results compared to single-species tests or to extant AWQC or state standards. These types of tests are theoretically ideal for "validating" regulatory values.

I am not currently aware of additional freshwater aquatic studies that should be considered at this time.

Saltwater (Section 11.4)

Please comment on the synthesis of the available information regarding the relationship between Pb exposure and effects on individual organisms and ecosystems. Please provide recommendations on any subject area that should be added, expanded, shortened, or removed. Is the panel aware of any important missing studies for characterizing Pb effects on biota and ecosystems within the scope and context of the ISA? Please comment on the application of available scientific evidence to inform the causality determinations in this section.

Section 11.4 provides an excellent synopsis of the available ecotoxicity information for Pb in the marine environment. Although there is not as much information as is available for freshwater organisms, there is sufficient new information to modify some of the 2013 causality determinations that were judged "inadequate" at the time, to "suggestive" or "causal" based on current data. In fact, there is now sufficient data to permit the calculation of regulatory marine values (HC5) using species sensitivity distribution approaches, e.g., Church et al 2017. This suggests that a great deal of research has been developed to characterize the response of marine organisms to lead since the 2013 ISA and the information would be useful in updating the current 1984 Ambient Water Quality Criteria. All of these studies characterize the effects of lead in the marine environment and are no doubt an important contribution to our understanding of the potential effects of Pb in the marine environment. The ISA does a good job of describing the environmental "effects" side of the risk characterization paradigm and the importance of considering bioavailablility to address site-specific toxicity; however, no discussion of the relationship between atmospheric concentration and waterborne (or soil) concentration is provided. No discussion of source apportionment is included, so the relationship between concentrations in air and concentrations in water or soils is missing.

The environmental fate processes affecting Pb in the marine environment are decidedly different that in the freshwater environment, but this is not addressed. For example, $PbCl^{+}$ is the dominant species over the normal pH range in saline waters, while in freshwater, Pb^{+2} is the predominant species. This is important in that PbCl is poorly soluble and will tend to precipitate and be found in bottom sediments. Thus, some discussion of Pb speciation and environmental fate in marine waters would be beneficial in assessing the effects of atmospheric lead in the marine environment. I must ask, has EPA considered whether the contribution of direct atmospheric Pb to marine waters significant or is it primarily from contributions to freshwater sources? I could not find it addressed in Appendix 1. I would have guessed that Pb in marine waters comes about primarily as a consequence of terrestrial runoff, direct discharge from industry or publicly-owned treatment facilities, and/or transport from non-point source into freshwater systems. So the question is, are atmospheric contributions of Pb (either direct or via runoff from freshwaters) likely to be an environmental concern in marine environments and how high would loads have to be (on a local level) to result in marine concerns?

I am not currently aware of additional marine studies that should be considered at this time.

Section 11.4.6 addresses new information addressing saltwater communities and ecosystem effects. This is a nice synopsis of the available data and it highlights that adverse effects were noted in some studies, but not in others. It fails to address the important question, were adverse effects noted at Pb concentrations that were not expected to cause impact? How do reported effect concentrations in

community level studies compare to single-species tests or to derived criteria/standards…in other words, do they validate the methods used for developing regulatory protective values?

Reference

Besser, J.M., Brumbaugh, W.G., Kemble, N.E., Ivey, C.D., Kunz, J.L., Ingersoll, C.G., and Rudel, David. 2011. Toxicity of nickel-spiked freshwater sediments to benthic invertebrates—Spiking methodology, species sensitivity, and nickel bioavailability: U.S. Geological Survey Scientific Investigations Report

Dr. Kathleen Vork

Appendix 2 - Exposure, Toxicokinetics, and Biomarkers

Appendix 2 describes the multimedia nature of Pb exposure, toxicokinetics of Pb in humans, biomarkers of Pb exposure and body burden, as well as models of the relationship between Pb biomarkers and environmental Pb measurements.

Overall, Appendix 2 is well written and organized. There are a few areas in the appendix where I have suggested rewording or reorganizing content for clarity and accuracy. Finally, I have suggested adding or removing text pertaining to issues raised in charge questions as follows.

2a. Please comment on the choice and emphasis of topics for providing useful context for the evaluation of human health effects of Pb in the ISA. Is the current organization of the appendix clear and logical? Please provide any recommendations to integrate exposure and toxicokinetic information more clearly throughout the appendix? Does the appendix adequately describe air-related and non-air related pathways of Pb exposure?

Some rewording and reorganization among sections is recommended for clarity and accuracy. Specific recommendations are given for parts of Section 2.2.1.1

In the beginning of section 2.2.1.1 Inhalation, the first sentence is confusing. I suggest rewording to say to something like: Particle size influences the pattern of regional deposition within the respiratory tract and, along with other factors, systemic absorption.

Also reword the text on lines 21 – 27 to say something like:

Section 4.2.4 of the 2019 PM ISA (U.S. EPA, 2019c) specifically addresses biological factors affecting particle deposition, such as.... on children. Chapter 4 of that document provides a detailed discussion of factors affecting particle deposition…

Particle size deposition and retention is mentioned as the focus of Chapter 4. Since absorption appears to be the main topic of this section, does that chapter also discuss absorption?

The following sections summarize empirical findings on absorption of inhaled particles from human experimental, urban, and industrial area exposures. However, Section 2.2.1.1.2 – Occupational is confusing and in some areas inaccurate.

Two paragraphs of this section summarize empirical findings from personal air samples of lead and experimental finding on absorption (bio accessibility) from the Dartey et al (2014) study fit with the material in the previous sections. The other material in section 2.2.1.1.2 does not seem to fit the rest of the section.

ICRP's classification system for estimating absorption of inhaled materials from particles deposited in the respiratory tract seems to fit in the first paragraph of section 2.2.1.1 as an example of chemical factors affecting absorption of inhaled particles. Recommendations for classifying various forms of inhaled lead from occupational exposures seems to fit better in section 2.6.

The description of CalEPA's simplified method for estimating a default inhalation transfer coefficient as a component of a physiologically based model is misrepresented. A more appropriate description is provided in Vork et al (2023) and would also fit better in section 2.6.

2b. Please comment on how well Section 2.3 reflects the current state of knowledge of Pb biomarkers and their interpretation as it relates to exposure and dose?

Is the focus on blood Pb and bone Pb appropriate given the epidemiologic literature largely assesses exposure through these two biomarkers?

Yes, this seems appropriate.

Is there sufficient and accurate discussion of the relationship between blood Pb and bone Pb?

Add a brief discussion of the evidence in the occupational studies by Fleming and colleagues of a supra linear vs linear curve in the relationship between bone and blood Pb. Also recommend describing job tenure cohort slope differences in that relationship and other related findings in the Fleming et al papers.

Are relationships between blood Pb and Pb in soft tissues and urine Pb adequately described?

Add a summary of the potential impact of disposition and elimination of lead during initial days following exposure through the inhalation such as that described in Leggett 1993.

2c. Sections 2.5.1 and 2.5.2 discuss empirical models of the relationship between air Pb and blood Pb from recent and older studies.

Please comment on the effectiveness of this section to accurately reflect what is known about air Pb-blood Pb relationships.

From the Meng et al supplemental Table – point out PM2.5/PM10, TSP differences in sample size as a potential factor in statistical significance. Also, it looks like there is a statistically significant finding from exposure to Pb in PM2.5 in the age 6-11 cohort. Please add a summary of this finding.

Add empirical models from occupational studies that provide some indicator or measure of particle size as a modifier of the relationship between air and blood Pb (see studies described in Vork et al 2023)

As with nonoccupationally exposed populations in different geographic locations (i.e., near vs remote from sources of lead contaminated air), it is also difficult to compare occupationally exposed cohorts that experience considerably different exposure conditions as demonstrated in the following occupational studies.

Rodrigues et al (2010) - measures the change in blood lead over a two-week exposure and air concentrations are respirator corrected.

Pierre et al (2002) - workers with long job tenure of 10 or more years and no mention of workers wearing masks.

Lai et al (1997) - workers with between six months to 3 or 4 years of job tenure and a mix of workers wearing masks vs not wearing a mask.

In addition to the potential temporal and current sources of exposure misclassification, different units such as mg/m3 vs ug/m3, ug/L vs ug/dL, log10 vs natural log regression and are all factors that need to be identified so that the reader can reproduce, verify, and interpret the reported results. A table with footnotes providing these differences would be helpful here.

Please provide recommendations on any studies that should receive less or greater emphasis.

I recommend including key insights from occupational studies that examined particle size-related air-blood lead relationships.

Various findings are presented in different units. In the discussion of absorption fraction in relation to concentration, mass, and surface area, common metrics are needed to compare findings. For clarity and comparison purposes, make those adjustments in this document. Also, briefly explain the relationship between bio accessibility and surface area of small particles vs large particles.

Comments on additional sections within Appendix 2

Section 2.6 Biokinetic models of exposure-blood Pb relationships

I recommend adding to this section, a summary of the influence of particle size of inhaled lead on respiratory tract deposition, clearance, transfer into the GI tract and uptake into blood and/or cite summaries found elsewhere as a component of the Leggett model adapted for assessment of occupational exposure scenarios (e.g., Vork et al 2023).

New reference

Kathleen L. Vork, Joseph P. Brown & James C. Carlisle (2023) Evaluation and updates to the Leggett model for pharmacokinetic modeling of exposure to lead in the workplace – Part II adjustments to the adult exposure model, confirmation of Leggett+, and modeling of workplace exposure, Journal of Occupational and Environmental Hygiene, 20:2, 55-83, DOI: 10.1080/15459624.2022.2150767

Dr. Marc Weisskopf

3a. Please comment on the degree to which the appendix accurately describes and appropriately interprets the strengths and limitations of various types of health studies, including epidemiologic and animal toxicological studies.

In general, the appendices do a good job describing the health studies and appropriately comments on the strengths and limitations of the epidemiological studies. The studies are placed in context well.

3c. To what extent do the causality determinations appropriately reflect the strengths and limitations of the evidence?

Overall, I think the causality determinations seem appropriate given the strengths and limitations described.

Some specific comments:

1) Appendix 3, 3.6.1.1, p.3-174, ~10 lines from bottom: Please indicate here when first mentioned that the case-control study is of Parkinson's disease.

2) Appendix 3, 3.6.1.3.1, p.3-180, 1st paragraph: The results from Wang et al and Weisskopf et al seem to go in the opposite direction, so saying "Similarly,…" is not correct. In the former there was a larger decline at higher tibia Pb levels, and in the latter, the opposite (unless the studies are not described correctly).

3) Appendix 3, 3.6.1.3.1, p.3-181, towards the end of the 1st paragraph on the page: Something seems wrong in the description of the adulthood results in Reuben et al, 2017 study. The overall effect is not between those at higher and lower Pb levels. And the childhood to adulthood results also seem odd—the overall result is very close to that for those over the level of concern. Are most of the people over the level of concern?

4) Appendix 3, p.3-182, Pre-existing conditions, 3rd line: Again indicate at the first mention that the case-control study is of Parkinson's disease.

5) Appendix 3, 3.6.2.3.3, p.3-196, l. 25 (p. 3-198, l.17 too): anti-depressant use seems unlikely to be a confounder—it does not affect Pb levels, but rather might be affected by them. It is clearly an important variable to consider as it could be on a causal path and so affect results, but confounding is not the right term.

6) Appendix 3, 3.6.3.3.1: Need to include Grashow et al., Neurotoxicology. 2015 Jul;49:158-64.

7) Appendix 3, 3.6.4.1.1, 3-210, l.5: Also exposure metric not capturing relevant exposure window is an important limitation. (add to summary too.)

8) Appendix 3, 3.6.4.1.2: Need to add Peters S, et al. Ann Neurol. 2021 Jan;89(1):125-133. Also consider Dickerson AS, et al., Occup Environ Med. 2019 Apr;76(4):208-214 and Wang TW, et

al. Amyotroph Lateral Scler Frontotemporal Degener. 2023 Feb;24(1-2):100-107. But these two are occupational exposure, should they be excluded? The use of job-exposure matrices in these latter two get around the concerns of reverse causation when using BLL.

9) Appendix 3, 3.6.4: I wonder whether at some point different neurodegenerative disorders should be considered separately for causality determination.

10) Appendix 4, 4.10.1, p. 4-63, l.8: Need to cite Weisskopf et al. Environ. Health Perspect. 123(11), 2015, which looked at bone lead and cardiovascular mortality (at least in addition to Weisskopf et al., Circulation, 2009, if not instead of). See results of Table 2, model 4.

3e. Please comment on the current organization of this appendix (#9) and the decision to incorporate a separate causality determination for exposure to Pb and total (nonaccidental) mortality.

Overall, the organization of these appendices seems good. I also think it is appropriate to have a separate determination for overall mortality.

Some specific comments on Appendix 9:

1) In many places (e.g. 9.5.6, p. 9-45, ll.29-30) it is stated that cross-sectional studies have several limitations, in particular, related to difficulties being able to determine temporality. This is correct and appropriate, but it is important to keep in mind that this is more problematic for a biomarker like blood lead that has a relatively short half-residence time. Lead in bone has a much longer half-residence time and as such reflects much longer-term cumulative exposure more than does blood lead. As such the concerns about temporality are a little different when this exposure is measured at the same time as an outcome. This is probably worth stating explicitly.

2) 9.1.3, ll. 16-20. (other places too, though. Like 9.5.3, p. 9-37, para. starting l. 15): It is true that "…with BLL, it is difficult to characterize the specific timing, duration, frequency, and level of Pb exposure that contributed to associations…", but it should be noted that this is also an issue with bone lead since bone lead only provides an estimate of cumulative exposure and so "specific timing, duration, frequency, and level of Pb exposure" is also not something that can be teased apart.

3) 9.2.2.3, Elevated fasting glucose (and some other places, e.g. end of 9.2.3.4, and 9.7.3.1, p. 9-58, l.16): at the end of the paragraph it is stated that "…the small sample size (n = 150) in this study reduces statistical power, as well as the likelihood that an observed result reflects a true effect." The first part I find correct, the second part I feel mischaracterizes things a bit. With a continuous outcome, that the observed result might not reflect a true effect is captured by the confidence limits, which is really already captured by the first part of the sentence. There is really no reason beyond that to doubt a true effect here any more than a larger study. The situation is a little different for dichotomous outcomes, because when the number of those is very small, then getting just one or two of those wrong can dramatically affect the point estimate, but this I think is not the same with a continuous outcome.

4) 9.4.3, p. 9-30, l.6: Better adjustment for thyroid medication use is touted as a strength here, but I don't see how that is a confounder (why related to lead?). Instead it would seem to me, if anything, it might be an effect modifier.

5) 9.4.3, p. 9-30, paragraph starting at l.24 (but this issue comes up throughout the text and in other appendices too): In discussing findings from a cross-sectional study, I would avoid using language of "increase" and "decrease" of exposure and outcome. These imply longitudinal changes within a person, which these studies do not get at. Instead, language of "higher" and "lower" for exposure and outcome is more appropriate.

6) 9.5.3.1 & 2: The studies of these 2 outcomes seem particularly subject to reverse causation. This gets stated at the end of 9.5.3.2, but not 9.5.3.1. In both sections, though, the concern is so great, I feel they each need more text about this. This is discussed a little in the summary, 9.5.6, but it is not clear to me how much the causality determination considers this. The "likely causal", to me, would have to be driven very heavily by the tox data.

7) 9.5.3.2, p. 9-39, l.5: As in point 4 above, here I think hormone therapy use is likely an issue of effect modification rather than confounding.

8) The all-cause mortality section does not discuss Weisskopf et al. Environ. Health Perspect. 123(11), 2015, which looked at bone lead and all-cause mortality (better to cite than Weisskopf et al., Circulation, 2009). Table 2, model 4 are most appropriate results to report.

Appendix 8

8.3.2.1, p.8-14, 1st para: In describing project VIVA, it is stated that in the metal mixtures model there were no associations between lead and the birth outcomes. But actual point estimates need to be given—the power will undoubtedly be less because of correlations among metals, so confidence intervals will certainly get wider. Thus, it is important to know what the point estimates are, as similar point estimates with wider confidence intervals that include the null may still suggest an effect of lead that lost significance because of correlation between metals.

8.3.2.1: In general in this section it is important to pay close attention to timing of exposure and what the specific lead measure reflects. It is possible some effects are predominantly during a specific exposure window (e.g. 1st trimester of pregnancy) and so inconsistencies in findings could relate to different lead measures capturing that window specifically (like 1st trimester maternal blood sample, vs. tooth lead [teeth only start developing in the 2nd trimester] or cord blood lead) more or less well.

8.3.2.1, p. 8-10 (other pages too): In describing studies using maternal blood (or serum) Pb, it is necessary to say when the maternal blood was taken. This is critical to understanding what exposure window the measure reflects.
8.3.2.1, p. 8-15 (also 8.3.3.1): I view the Bui et al. 2022 study as very strong evidence of causal effects given its quasi-experimental design. It is immune to much confounding and reverse causation.

8.8.1, p. 8-60, l. 26: There is something missing in the phrase "…Pb exposure and maternal in the 2013 Pb ISA, recent…" -- "maternal health outcomes"?