CLARK HILL

Clark Hill PLC
601 Pennsylvania Avenue NW
North Building, Suite 1000
Washington, DC 20004
T 202.772.0909
F 202.772.0919

clarkhill.com

Jane C. Luxton
T 202-572-8674
F 202-772-0923
Email: JLuxton@ClarkHill.com

September 22, 2014

Via Federal eRulemaking Portal
Document Control Office (7407M)
Office of Pollution Prevention and Toxics (OPPT)
Environmental Protection Agency
1200 Pennsylvania Ave., N.W.
Washington, DC 20460-0001

Re: Approach for Estimating Exposures and Incremental Health Effects From Lead Due to Renovation, Repair, and Painting Activities in Public and Commercial Buildings; Notice of Availability and Request for Comment; EPA-HQ-OPPT-2010-0173

Dear Sir or Madam:

Please find attached for filing the Comments of the Commercial Properties Coalition in response to EPA's Request for Comment on its Approach for Estimating Exposures and Incremental Health Effects From Lead Due to Renovation, Repair, and Painting Activities in Public and Commercial Buildings, 79 Fed. Reg. 45796 (Aug. 6, 2014). The Coalition's Comments are in the form of the Exponent Technical Review of Documents Related to EPA's *Framework for Identifying and Evaluating Lead-Based Paint Hazards from Renovation, Repair, and Painting Activities in Public and Commercial Buildings* (September 2014).

Please contact me if you have any questions.

Respectfully submitted,

Jane C Luxton

Jane C. Luxton

Enc.

Exponent®

Technical Review of Documents Related to EPA's

Framework for Identifying and Evaluating Lead-Based Paint Hazards from Renovation, Repair, and Painting Activities in Public and Commercial Buildings

Technical Review of Documents Related to EPA's *Framework for Identifying and Evaluating Lead-Based Paint Hazards from Renovation, Repair, and Painting Activities in Public and Commercial Buildings*

At the request of the Commercial Properties Coalition, scientists within Exponent performed a technical review of the *Framework for Identifying and Evaluating Lead-Based Paint Hazards From Renovation, Repair, and Painting Activities in Public and Commercial Buildings* (Framework). The Exponent team assembled for the review comprises individuals with specialized expertise in air modeling, Monte Carlo analyses, health effects of lead exposures, and industrial hygiene practices.

The May 30, 2014 Federal Register contained an advance notice of proposed rulemaking (2014 ANPR [U.S. EPA 2014]) and requested comments on the Framework. The Framework contains some discussion of the general procedures proposed for use in evaluating potential hazards associated with renovation, repair, and painting (RRP) activities in public and commercial buildings. These procedures include air dispersion modeling to estimate impacts of RRP activities downwind of renovated buildings, Monte Carlo procedures to characterize exposure and risk, health effects/risks at low levels of lead exposure, and risk modeling procedures. Each of these topics was reviewed by one or more members of the Exponent team. Our review of these procedures is provided below. Most of the comments provided below are based on review of the Framework initially released in May 2014. However, more recently, additional technical reports that provide details of EPA's proposed approach have become available. For two of these, additional focused comments are provided. Specifically, these more recent documents include EPA's June 2014 document, *Developing a Concentration-Response Function for Pb Exposure and Cardiovascular Disease-Related Mortality* (Abt Associates, June 2014) and information contained in *Appendices to the Approach for Estimating Exposures and Incremental Health Effects from Lead due to Renovation, Repair, and Painting Activities in Public and Commercial Buildings.*

We begin with a more general comment related to the need for a technical Framework to deal with EPA's concerns about lead exposures associated with renovation.

The Need for a Framework is not Established

The Framework is premised on the need for a relatively sophisticated analytical approach to assess exposures that may occur as a result of building RRP activities. The Framework acknowledges that other aspects of lead exposures are already addressed, but that there is or may be a need to address the specific topic of lead exposure associated with public and commercial buildings. However, EPA has not established the case for the evaluation of these buildings. In fact, the Framework specifically states that exposures from the sources being considered are likely to be very small. In light of the lack of data on hazards and on EPA's preliminary judgment that risks are likely to be small, it would be prudent for EPA to employ

the commonly phased approach for assessing hazard and/or risk. This would logically begin with evaluating or confirming whether renovation of buildings poses a substantive risk of contributing to lead exposures. If confirmed, then subsequent exposure analyses might be employed, including, if appropriate, the detailed and complex analysis that is proposed in the Framework. The process of determining whether there is a need for a special approach to the evaluation of commercial buildings should be based, in part, on a reliable technical assessment of potential risk and/or hazard. This would be a ***Strategic Plausibility Analysis*** or ***Informed Bounding Evaluation*** that might provide EPA with insight into possible risks/hazards and indicate whether more sophisticated and complex probabilistic evaluations are even warranted. This type of bounding analysis could entail commonly-used "backward" risk calculations to determine whether intermittent exposures could even get close to resulting in exposures of health concern. EPA states in the Framework that they expect exposures to be very small, so it would make sense to confirm that before undertaking a large-scale evaluation. This assessment would still require that EPA identify blood lead levels against which incremental exposures could be judged. At present, there are two conceptual views on this, and it would be worthwhile to consider both.

As discussed above, we question whether a need exists for the highly detailed process proposed in the Framework, and we believe that simpler, more direct methods are available that would fulfill EPA's obligation to address public and commercial buildings. Nevertheless, it appears that EPA is constructing a complicated analytical approach. We examined technical elements of this approach and point out substantial limitations that would affect the reliability and appropriateness of applying the Proposed Framework.

Reliability of Framework in Light of Data Limitations

Our principal concern with the use of the Framework is that it will generate unreliable results. The approach requires a substantial number of input parameters. While data may be available for some variables, it is likely lacking for others. There are, for example, significant data limitations with respect to source terms and parameters needed for air modeling. No amount of probabilistic analysis can overcome these data limitations. The results of analyses based on weak or assumed input data can be very misleading. In the absence of data for input parameters, estimates would likely be used based on professional judgments or sparse information. Combining estimated parameters and weak data into a probabilistic approach can result in an enormous spread of results that would likely overestimate exposures. This is because estimated input parameters would necessarily be conservative to ensure that the tails of the estimated distributions are not missed. Using a probabilistic approach can be deceptive in such cases. The appearance is created that a reliable and useful analysis has been completed in which uncertainties and variability have been taken into account. However, the lack of knowledge embedded in the uncertainties we identify in our comments will lead to an inflation (i.e., spreading) of the Monte Carlo output well beyond the bounds of reality.

Understanding Incremental Exposures and Baseline Conditions

Background conditions and associated levels of lead in blood can vary considerably and can confound and confuse analyses and interpretations. Blood lead levels in children reflect the aggregate of exposures. However, where elevated baseline levels occur, there are usually one or a few predominant sources. Current community health practices involve identifying and managing those sources. To our knowledge, there is no evidence to suggest that RRP activities in public and commercial buildings are a significant source of lead to children, and EPA states its view that this potential source is likely a small contributor to exposure. Further, there are geographic locations and/or urban conditions (e.g., areas with high natural background and older urban centers) where baseline blood lead levels are elevated relative to other areas. Assuming that RRP activities are very small incremental sources, the ***relative contribution*** of RRP to blood lead levels in areas with more elevated blood lead levels is actually a smaller fraction of the total burden than for areas where baseline blood lead levels are already very low. This underscores the importance of: (1) understanding the predominant sources of lead in areas of concern; (2) employing an incremental and relative risk approach for judging specific sources, such as RRP activities for public and commercial buildings; and (3) directing lead management programs in areas where they are needed and to sources that are predominant contributors to exposure.

Estimating lead dust loading from RRP work tasks in public and commercial buildings requires information from real-world studies of these tasks. While we are aware of studies of worker exposure in designated work spaces, we are unaware of any comprehensive study of dust loadings from work spaces to building areas outside these work areas.

Limitations of blood lead levels for assessing lead exposure associated with disease in older adults

EPA appears to be selecting a dose-response approach that has substantial uncertainty and will almost certainly overestimate risks associated with exposures in children. In EPA's June 2014 document, *Developing a Concentration-Response Function for Pb Exposure and Cardiovascular Disease-Related Mortality* (Abt Associates, Inc. 2014), the authors choose blood lead as the metric for lead exposure and consider four epidemiologic studies (Menke et al. 2006; Schober et al. 2006; Khalil et al. 2009; Weisskopf et al. 2009) in which a single blood lead measurement was taken for each subject. However, the measured blood lead in adults is not only a function of recent exposures but also reflects past exposures. Bone lead stores associated with past exposures may reenter the bloodstream with bone remodeling (which is a constant process, but is more active early in life and perhaps later in life with osteoporosis). The potential for misalignment of current blood lead levels with current versus past exposures is further exacerbated by the fact that today's adults experienced lead exposures that were substantially higher before the early 1980s. As children, many of these adults experienced elevated lead associated with lead in gasoline, paint, food cans, and other products. Past lead exposures may also be high in those who grew up in older housing with eroding lead paint. Therefore, adults with blood lead levels measured later in life that show higher levels than their

peers may have had very high past exposures in early childhood, or moderately elevated lead exposure in adulthood (e.g., from occupational exposure), or both. In general, however, with attenuation from storage in bone and excretion over time, blood lead levels measured in older adults will be lower than earlier in life, and do not reflect current exposure levels that might result in the effects of concern.

Thus, it is unclear whether associations observed with adult blood lead levels are due to maximum lifetime exposures (or at a sensitive life stage), recent exposure, or cumulative lifetime exposure (assuming that associations are not explained by bias or confounding). For chronic diseases such as cardiovascular disease (CVD), many causal exposures are thought to exert their effect over a period of years, rather than acting instantaneously. If this is the case for lead and CVD mortality, then blood lead may not indicate the etiologically relevant exposure. However, our understanding of the mechanism by which lead may cause CVD at low levels is incomplete, as acknowledged in the above studies.

Moreover, to the extent that current blood lead levels in adults reflect the release of bone lead accumulated from exposure to much higher levels of lead in the past, as well as recent external exposure, the relative risk of CVD mortality per unit of recent lead exposure is overestimated. Thus, if the relative risk of CVD mortality per 1-µg/dL increase in blood lead is determined not only by recent lead exposure but also by past exposure to substantially higher lead levels, then the preventive impact of interventions to decrease current lead exposure will be exaggerated.

Declining blood lead levels over time complicate exposure assessment

Historically, the national geometric mean blood lead level in U.S. children aged 1–5 years during the first National Health and Nutrition Examination Survey (NHANES 1976–1980) was 15 μg/dL, with about 90% of this population having blood lead levels of 10 μg/dL or higher. The average blood lead level of all ages in 1976 was approximately 16 μg/dL, which declined to about 9.5 μg/dL in 1980 (NCHS 1984). Blood lead levels of children and all ages declined in subsequent surveys with the phasing out of lead in gasoline and other consumer products (Pirkle et al. 1994). Nationwide blood lead data are not available prior to 1976, although Chisolm (1970) noted an upper limit for normal blood lead levels in urban areas of 40 μg/dL.

The four epidemiologic studies considered for derivation of the concentration-response function between lead and CVD mortality are based on blood lead levels measured in 1988–1994 (Menke et al. 2006; Schober et al. 2006), 1986–1988 (Khalil et al. 2009), and 1992–1999 (Weisskopf et al. 2009). Lead exposure and blood lead levels in the United States have decreased substantially over recent decades, including during the period since blood lead was measured in these studies. In these studies, the blood lead levels among study participants were geometric mean = 2.58 µg/dL (Menke et al. 2006; assumed to be similar in Schober et al. 2006, which did not report the average blood lead level, but was based on an older subgroup [ages $\geq$ 40 years instead of $\geq$ 20 years] of the same study population as Menke et al. 2006), mean = 5.3 ± 2.3 µg/dL (Khalil et al. 2009), and geometric mean = 4.8 µg/dL (interquartile range, 3–7) (Weisskopf et al. 2009).

By contrast, as of 2009–2010, the geometric mean blood lead level in U.S. adults was 1.23 µg/dL (95% confidence interval = 1.19–1.28) (CDC 2013). This level is near the lower limit of detection of 1 µg/dL reported by the epidemiologic studies (Menke et al. 2006; Schober et al. 2006; Khalil et al. 2009), meaning that these studies were unable to characterize exposure variability for subjects with blood lead levels close to the average of today's U.S. national population. Therefore, in the absence of robust data on the shape of the concentration-response function at blood lead levels near 1 µg/dL, results from these studies cannot reliably be assumed to be relevant to current average blood lead levels in U.S. adults. Instead, almost nothing is known about the relationship between blood lead levels and CVD mortality in this exposure range.

Observed associations of lead exposure and CVD at high exposures should not be extrapolated to lower exposures

In light of the paucity of data on blood lead levels below 1 µg/dL, which are most relevant to today's U.S. population, the validity of extrapolating concentration-response functions down to lower levels in this range is tenuous. Especially given the evidence of non-linearity in the concentration-response function for blood lead in relation to CVD and all-cause mortality outcomes (i.e., inverse associations in Menke et al. 2006 and null to inverse associations in Schober et al. 2006 at low blood lead levels), it is inappropriate to assume that relationships observed at higher blood lead levels can be extrapolated to lower levels.

Of note, when blood lead levels were analyzed categorically, statistically significant associations with CVD mortality were observed only with the highest exposure category in each study: ≥3.63 µg/dL in Menke et al. (2006), ≥10 µg/dL in Schober et al. (2006), and ≥8 µg/dL in Khalil et al. (2009). Weisskopf et al. (2009) did not observe a statistically significant association between the highest tertile (>6 µg/dL) of blood lead and CVD mortality. Given the open-ended nature of the highest exposure categories, these associations could have been driven largely by blood lead levels well above those typically observed in the United States today. Those with higher blood lead levels in these studies would likely have had even higher past lead exposure and blood lead levels. This observation again calls into question the validity of applying the results of these studies to current blood lead levels.

Abt Associates, Inc. (2014) inappropriately discounted the conclusions by NTP regarding limited evidence of an association with cardiovascular disease mortality at low exposure levels

U.S. EPA concluded in its 2013 Integrated Science Assessment for lead that there is a "causal" relationship with CVD mortality, without specifying whether this relationship exists at low levels of lead exposure. By contrast, the National Toxicology Program (NTP) concluded in its 2012 monograph focused on low-level (<10 µg/dL) lead—which is the exposure level of interest for the EPA in the current Framework document—that there was "limited" evidence of an association with CVD mortality. In particular, Abt Associates, Inc. (2014) quotes NTP as

stating that the "association between increased CVD mortality and increased blood Pb was supported by three prospective studies but not supported by two prospective studies, one of which reported a significant association with bone Pb." Bone lead is more representative of cumulative life-long lead exposure than blood lead, although it also has limitations for assessing the magnitude, timing, and frequency of the exposure.

On pp.2-5 to 2-6, Abt Associates, Inc. (2014) largely discount the NTP conclusions by noting that one of the two negative studies (Møller and Kristensen 1992) combined fatal and non-fatal CVD cases, and that the other negative study (Weisskopf et al. 2009) "suffers from selection bias" because the study cohort "is weighted toward individuals without CVD, given that in order to be entered into the cohort you could not have prior CVD. For older individuals this creates a strong selection bias toward heart-healthy people." In fact, selecting cohort members based on the absence of a history of CVD does *not* create selection bias, which occurs when study participation is associated with both the outcome *and* the exposure. In the Weisskopf et al. (2009) study, eligibility was dependent on relatively low risk of the outcome (CVD), but independent of the exposure (blood and bone lead concentration), and therefore would not have resulted in a distortion of estimated relative risks.

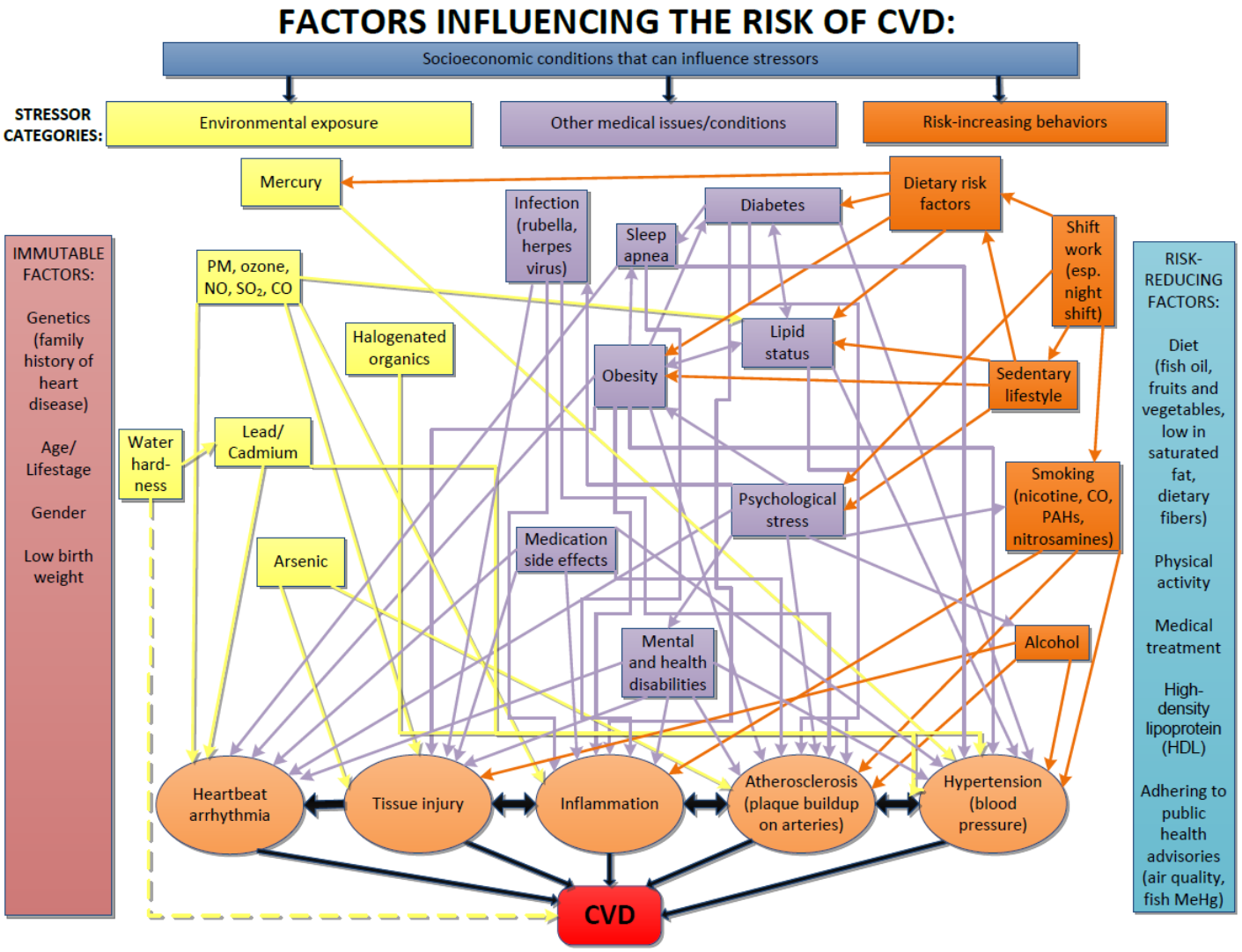
Rather, the apparent selection bias described by Weisskopf (2013) seems to have been due to socioeconomic status, which can influence lead exposure, future CVD risk, and the decision to participate in a cohort study (including study entry and follow-up). This type of selection bias can affect any cohort study, not only that conducted by Weisskopf et al. (2009), as discussed further below.

The inclusion of both fatal and non-fatal CVD cases in the study by Møller and Kristensen (1992) should not have diminished the association with blood lead levels if lead causes the development of CVD, rather than increasing CVD severity (such that it would be associated with fatal but not non-fatal disease). Given that EPA (2013) concluded that a causal association exists between lead and blood pressure increases and hypertension, and NTP (2012) concluded that there is sufficient evidence that low-level blood lead is associated with these outcomes, this scenario seems implausible. Thus, the null association between blood lead level and fatal and non-fatal CVD reported by Møller and Kristensen (1992) should not have been dismissed as irrelevant to the association with CVD mortality.

In summary, Abt Associates, Inc. (2014) inappropriately discounted NTP's conclusion about the "limited" evidence of an association between low-level blood lead and CVD mortality (NTP 2012), and should have given greater credence to the two studies with null results before proceeding to estimation of a concentration-response function for an association that may not be causal. Of note, two of the three prospective studies with positive results (Menke et al. 2006; Schober et al. 2006) were based on the same study population. Thus, among the studies cited by NTP (2012), only two independent prospective studies found a positive association between blood lead and CVD mortality (primarily at higher lead exposure levels), whereas two found no significant association. Further, Hara et al. (in press) conclude that small and inconsistent effect sizes in the associations of blood pressure with blood lead likely exclude current environmental lead exposure as a major hypertension cause in the United States.

Assessing causation for CVD is especially complicated and there exist predominant causal factors that should not be ignored when assessing causation for any one possible factor, namely

exposure to lead. The flow chart below represents a conceptual model the potential cumulative risk of a multitude of factors with regard to CVD as a health endpoint (Menzie and Kashuba, 2013). This conceptual model was prepared as part of the Agency's effort to develop guidance on consideration of cumulative risks. However, the model also displays the myriad factors that need to be taken into account when assessing causation. It should be evident from this figure that correlations that may be present in data sets may not reflect causal relationships. We consider the significance of some of these factors later in our comments.



Use of only a single analysis from Menke et al. (2006) to develop the concentration-response function should be expanded to consider the full evidence from this study and others

The concentration-response function for blood lead and CVD mortality by Abt Associates, Inc. (2014) is ultimately based on the results of only one study (Menke et al. 2006). The selection of this study is reasonably well justified based on the following considerations: (1) Menke et al. (2006), but not the other three studies considered (Schober et al. 2006; Khalil et al. 2009; Weisskopf et al. 2009), reported results using a continuous concentration-response function; (2) Menke et al. (2006) was based on a nationally representative adult population, whereas Khalil et al. (2009) and Weisskopf et al. (2009) were not; and (3) Menke et al. (2006) had the largest study population. Also, Menke et al. (2006) considered both confounding and effect modification by hypertension and estimated kidney function, among other factors, whereas the other studies did not (although Khalil et al. [2009] adjusted for confounding by hypertension and numerous other factors). Finally, in a personal communication, Dr. Weisskopf "revealed that there were errors in the analysis" (Abt Associates, Inc. 2014, p.2-6) and recommended against using the results of Weisskopf et al. (2009) before corrected results were available, thereby ruling out this study.

Although Menke et al. (2006) appears to be the most appropriate study of the four considered for development of the concentration-response function, several issues nevertheless diminish its utility for this purpose. For instance, results of a quadratic spline model were presented graphically for all-cause mortality, myocardial infarction mortality, stroke mortality, and cancer mortality (Figure 1 of Menke et al. 2006), but not for overall CVD mortality. Evidence of an inflection point in the models for both all-cause and myocardial infarction mortality at approximately 2 µg/dL blood lead, with *negative* slopes between 0 and 2 µg/dL (as well as above approximately 7 µg/dL) suggests that a similar inflection point might be detected also for CVD mortality, indicating important non-linearity in the concentration-response function. However, because quadratic spline results were not presented by Menke et al. (2006) for CVD mortality, whether a similar non-linear dose-response trend exists for this endpoint is unclear. Instead, only the hazard ratio (HR) for CVD mortality based on the possibly incorrect assumption of a (log-log) linear concentration-response function was used by Abt Associates, Inc. (2014).

Two other issues not fully addressed by Menke et al. (2006) are modification of the concentration-response function by age and time period. Schober et al. (2006) analyzed the same data set from the National Health and Nutrition Examination Survey (NHANES III) limited to adults aged ≥ 40 years (presumably because CVD mortality among adults under age 40 years may be etiologically different from that in older adults). Instead of setting time since the NHANES III examination as the time scale for the Cox proportional hazards regression model—the approach used by Menke et al. (2006)—Schober et al. (2006) used age as the time scale and stratified the baseline hazard by birth cohort and survey phase, to allow for the decline in cumulative lead exposure over time. Schober et al. (2006) reported that "[f]or each category of deaths, statistical testing did not support the null hypothesis that proportional hazards were constant by age" (pp. 1539–40)—that is, the association varied significantly by age—and the authors therefore reported results stratified by age group. The hazard ratio for CVD mortality

was not statistically significant in any age group, but was stronger for ages 75–84 years than for 40–74 or >85 years. By contrast, Menke et al. (2006) found no statistically significant subgroup heterogeneity in the hazard ratio for CVD mortality, although the results suggested that the association was stronger for <60 than ≥60 years. Thus, whether the association between blood lead level and CVD mortality varies by age group in the NHANES III data set is unresolved. If the relationship is indeed modified by age, then the concentration-response function may not be generalizable to all age groups.

Menke et al. (2006) did not stratify results by birth cohort or survey phase, nor did they allow for different baseline hazards by these factors, as did Schober et al. (2006). Menke et al. stated that they evaluated the proportionality assumption of the Cox model using Schoenfeld residuals, but they did not report whether they detected any violation of this assumption—that is, whether the hazard ratio varied over calendar time (the time scale that they used). Therefore, it is unclear whether the association between blood lead level and CVD mortality varied over calendar period or birth cohort. If such variation was observed, it would also limit the generalizability of the concentration-response function (as well as the validity of the Cox model results).

Abt Associates, Inc. (2014) did not attempt to validate their concentration-response function against the results of the other three studies. If the results of Menke et al. (2006) are valid in terms of showing little subgroup heterogeneity of the hazard ratio between blood lead and CVD mortality, then the results should apply to other populations, including older women (Khalil et al 2009), male veterans (Weisskopf et al. 2009), and an older subset of the same study population (Schober et al. 2006). Validation in other study populations is thus seemingly reasonable. (If, on the contrary, the association does vary by age and sex, then the function is not as broadly generalizable as suggested by the authors.)

The authors of the document suggest that differences in health outcomes evaluated by the four studies hamper comparison of results (Abt Associates, Inc. 2014, p.3-15). However, an evaluation of the International Classification of Disease (ICD) codes used in the four studies reveals that results should be comparable at least among Menke et al. (2006), Schober et al. (2006), and Weisskopf et al. (2009). As their outcome measure, Menke et al. (2006) considered all diseases of the circulatory system except for transient cerebral ischemia (ICD 9th Revision [ICD-9] codes 390–459, except 435, for deaths in 1988–1998 and ICD-10 codes I00–199 for deaths in 1999–2000). This is essentially the same approach taken by Weisskopf et al. (2009), who considered ICD-9 codes 390–459, including 435; and Schober et al. (2006), who considered ICD-10 codes I00–I78, excluding I80–I99 (diseases of veins, lymphatic vessels, and lymph nodes, not elsewhere classified; and other and unspecified disorders of the circulatory system—i.e., diseases not likely to make a major contribution to CVD mortality). Khalil et al. (2009) selected ICD-9 codes 401–404 (primary hypertensive disease), 410–414 (ischemic heart disease), 425 (cardiomyopathy), 428 (heart failure), 429.2 (unspecified cardiovascular disease), 430–438 (cerebrovascular disease), and 440–444 (atherosclerosis, aortic aneurysm and dissection, other aneurysm, other peripheral vascular disease, and arterial embolism and thrombosis), as well as 798 (sudden death, cause unknown). The comparability of results for CVD mortality in this study to those from the other studies depends on the proportion of CVD deaths that are attributed to these codes—probably the majority, in which case, the results would be reasonably comparable.

Thus, it should have been feasible and appropriate to compare the derived concentration-response function to the results of the other studies. Consistency of results would lend credence to the validity and generalizability of the derived function, whereas inconsistency would suggest sources of error or heterogeneity that need to be considered before applying the function to the general population for a benefits analysis.

Overall epidemiologic evidence is sparse

Only four epidemiologic studies were considered in the development of the concentration-response function between blood lead and CVD mortality (Menke et al. 2006; Schober et al. 2006; Khalil et al. 2009; Weisskopf et al. 2009), and two of these were based on overlapping study populations (Menke et al. 2006; Schober et al. 2006), leaving only three independent studies of this association. Especially in light of important methodological limitations of these studies, this is an insubstantial body of literature on which to base public health and policy decision making.

All four studies relied on a single measure of blood lead level as the exposure metric. (Weisskopf et al. [2009] also used bone lead measurements taken at two anatomic sites—the tibia and the patella.) The limitations of using a single blood lead sample to assess lead exposure were discussed earlier.

All four studies used death certificates to determine cause of death, although Khalil et al. (2009) supplemented this information with hospital discharge summaries for 33% of deceased participants (n = 41 deaths from all causes). Misclassification (especially poor sensitivity) of CVD mortality based on death certificates is well known (e.g., Herrett et al. 2013; Wexelman et al. 2013; Harriss et al. 2011). Such misclassification could have been either non-differential or differential by lead exposure status, leading to unpredictable bias in the observed associations.

Confounding is a key concern in studies of the potential adverse health effects of lead exposure. Socioeconomic and many other environmental and behavioral factors are associated with lead exposure, and many of these factors are also associated with a variety of health outcomes, including CVD, raising the possibility of confounding. Although all four studies attempted to adjust for potential confounding by socioeconomic, demographic, and other factors—with more covariates included in the multivariate models used by Menke et al. (2006) and Khalil et al. (2009)—residual confounding cannot be excluded, whether due to failure to control for unmeasured factors or inadequate control for measured factors. Adjustment for confounding is particularly difficult for adults, because many earlier-life covariates that may be important for lead exposure and CVD risk are unknown.

For example, Menke et al. (2006) controlled for cigarette smoking as current, former, or never; alcohol consumption as yes vs. no; education as high school or below; physical exercise as ≥3 times per week or less; and household income as ≥$20,000 per year or below. Blood lead levels were statistically significantly higher among current smokers, alcohol consumers, those without a high school education, those without regular physical exercise, and those with a lower income—all of which are risk factors for CVD mortality. Therefore, the observed associations with elevated blood lead levels, with hazard ratios generally below 2.0, could readily be

explained by residual confounding. (Adjusted relative risks between these factors and blood lead levels were not reported by Menke et al. [2006], but they are typically stronger than 2.0 and could, therefore, be responsible for the observed hazard ratios.)

As noted by Weisskopf (2013), selection bias based on socioeconomic status is also a concern, even in prospective cohort studies: "Because socioeconomic status (SES) is often a strong predictor of cohort study participation, and many environmental toxicants have strong associations with SES, studies of health effects of environmental toxicants may be particularly susceptible to this bias." That is, if study enrollment and/or completion of study follow-up are related to both lead exposure and risk of future CVD mortality, as might occur if participation is associated with socioeconomic status, then estimated associations will be biased. In the cohort studied by Weisskopf (2013), non-participation resulted in bias of the association between bone lead levels and ischemic heart disease mortality toward the null. However, bias away from the null may occur in other settings. None of the four cohort studies under consideration is immune from this potential bias.

Other issues related to the Abt Associates, Inc., report

Other issues that are important for U.S. EPA to address before moving forward are identified below.

- The Abt Associates, Inc., document summarizes the NTP (2012) and EPA (2013) conclusions about other CVD morbidity endpoints, such as blood pressure, hypertension, coronary heart disease, peripheral artery disease, and cerebrovascular disease, but a concentration-response function is developed based only on CVD mortality. The decision not to use CVD morbidity endpoints was not thoroughly discussed and justified, but should have been considered, perhaps even for the purpose of sensitivity analysis. In addition, the decision not to consider other organ systems (e.g., neurological, renal, reproductive/developmental) as endpoints for development of the concentration-response was not thoroughly discussed and justified.

- The document briefly summarizes potential modes of action for lead and CVD, based on discussion in the U.S. EPA Integrated Science Assessment for lead. The summary implies a degree of certainty about the effects of lead on reactive oxygen species, nitrogen dioxide, inflammation, calcium ions, etc., and the downstream impact on CVD. However, the strength of this experimental evidence and its relevance to clinical CVD development and progression should be discussed more fully.

- High lead exposure is toxic to the kidneys, and kidney damage can result in hypertension, which is a risk factor for CVD. Less clear is whether moderate to lower lead exposure causes such effects. Because lead excretion is mainly through the kidneys, those with kidney disease and a reduced glomerular filtration rate (and hypertension) may have a higher blood lead level due to a reverse-causation phenomenon. Such a possibility has not been considered thoroughly by the authors of this document or the authors of the underlying studies.

Dispersion Modeling Procedures

Although the Framework mentions that the EPA AERMOD model will likely be used to model the dispersion of lead-containing dust from renovated buildings as part of the exterior analysis, and mentions some of the parameters that will be incorporated in associated Monte Carlo analyses, the notice and the Framework both acknowledge that the Framework "does not provide significant detail regarding modeling inputs and results" and that "further details … would be provided for review and comment in any future proposal." Consequently, it is not possible at this time to provide detailed comments on the proposed modeling approach, because it is not yet well defined. Indeed, it is not even clear whether EPA has a good idea of how it intends to model the downwind transport of dust emissions that would be generated by RRP activities.

One of EPA's specific criteria for the models to be used is to be consistent with the analysis used for the 2008 RRP rule. However, neither the April 22, 2008, final rulemaking notice nor the January 10, 2006, proposed rulemaking notice for the RRP rule provide any information regarding any air dispersion modeling that may have been conducted to support development of that rule. Indeed, there is no mention of air dispersion modeling in these prior notices, and there is no sign that any was conducted. However, in an earlier ANPR published on May 6, 2010 (2010 ANPR), EPA requested public comment on several issues that may provide some insight into EPA's thinking regarding the dispersion modeling.

Given the lack of detail regarding dispersion modeling inputs and procedures in the Framework, our comments focus on the suitability of the general approach, the appropriateness of the model proposed for use (AERMOD), and issues that should be taken into account in any modeling analysis that might be used to support rule development.

The use of a dispersion model, like AERMOD, to describe the downwind transport of dust emissions from a source is well established for various regulatory purposes. However, if AERMOD is used to predict dust concentrations and/or deposition downwind of RRP activities, and if the results are to be used in an absolute (rather than relative) sense (which appears to be the case here), then the results will be no better than the accuracy with which the modeled emission rates can be specified. Therefore, how EPA intends to estimate dust emission rates is a critical consideration.

Emissions to the outdoor environment due to dust generated by indoor RRP activities will likely be small, particularly if containment measures are used in accordance with standard workplace practices. Emissions to the outside environment from dust generated by exterior RRP activities would likely be larger. The 2010 ANPR references a 2007 report ("Characterization of Dust Lead Levels After Renovation, Repair, and Painting Activities") that includes information on the amount of lead collected per square foot for various types of interior and exterior RRP jobs. EPA could consider using these data to estimate dust generation rates if the underlying data are suitable and reasonably representative of the activities to be analyzed. It's unclear whether this is EPA's intention.

Another possible approach would be to use some sort of mass-balance approach, as implied by Figure 3 and Appendix B of the Framework, based on the listing of factors such as the fraction of paint emitted in bulk and aerosol form and the containment efficiency. However, the

Framework text mentions the use of a mass-balance approach with respect to the interior, not the exterior, analysis.

Another possible approach would be to use existing AP-42 factors for certain types of activities, such as abrasive blasting and some construction activities. However, the associated emission factor ratings are likely not good enough (e.g., the abrasive blasting emission factor is "E" [poor]) to provide meaningful emissions estimates.

The important factor is whether emission rates associated with interior and exterior RRP activities can be estimated accurately enough to provide a means of obtaining meaningful predicted downwind impacts for use in other portions of the analysis. Without this, estimates of exposure and impacts on blood lead become highly uncertain.

Other concerns relate to source characterization. In the 2010 ANPR, EPA asked whether dust drifting from exterior renovations would resemble smelter plumes. Because of differences in the sources, we know that this would not be the case. Plumes from smelters are hot and buoyant and typically emitted from high stacks or from roof monitors. Smelter emissions would be modeled either as point sources (for stacks) or as buoyant line sources (for roof monitors). In contrast, emissions of dust from renovations, particularly from exterior work, would be expected to be largely non-buoyant and would not likely be released from an identifiable chimney, stack, or vent. The emissions would be fugitive in nature and would be best characterized as a volume source, not as a point source.

In the Framework, EPA specifically mentions the ability of AERMOD to incorporate consideration of "obstruction adjustment." We assume that EPA is referring to the building downwash algorithms within AERMOD. However, these algorithms are invoked only for point sources, not for volume sources. Therefore, one of the attributes that EPA cites for selecting AERMOD may not be compatible with or relevant to the likely nature of the sources that would be modeled. The only adjustment that might be applied to volume sources is characterizing the initial horizontal and vertical plume dispersion or size as a function of building size. Although this will account to some extent for enhanced initial dilution due to the source building, it will not account for any subsequent dilution associated with downwind structures. If the downwind target structures where impacts are to be estimated are tall or in clusters, AERMOD will treat the plume as if it moves through (rather than around) those structures, and the associated plume dilution will be underestimated. This could lead to an overestimation of downwind concentrations.

Another potential issue relates to the treatment of wet and/or dry deposition in the exterior analysis. It's not clear whether EPA intends to use AERMOD to explicitly predict deposition of dust generated by RRP activities, or if the consideration of deposition is limited to the use of the separate dust model cited for use in the interior analysis. It would seem that estimates of deposition would be relevant to tracking in dust from the outdoors downwind of public and commercial buildings and residences. However, given the relatively short downwind region of impact that would be expected from most RRP activities, the amount of dust that would be deposited in the near-field area is likely to be small. If AERMOD is used to estimate deposition, then additional parameters (such as the particle size distribution and/or a mean particle size) would be needed. This sort of information may not be readily available.

RRP activities are often of fairly short duration and generally would be expected to occur during daylight hours. If this is the case, then any modeling conducted by EPA to support rulemaking should account for the expected times of occurrence and duration of these activities. If the RRP activities under consideration will not occur at night but are modeled as if they do, then the AERMOD results will likely significantly overestimate actual impacts. Studies have shown, and EPA has acknowledged, that AERMOD, in its regulatory default mode, greatly overestimates actual impacts from low-level sources during light wind and stable hours (i.e., the conditions that tend to occur at night). Therefore, it will be important to use model options in AERMOD that can account for variations in emissions with time of day and season, if applicable. In addition, it would be advisable to use certain "beta" (non-default) options within AERMOD and the associated AERMET meteorological pre-processor to help reduce the degree of model over-prediction that has been observed during light-wind, stable hours. Otherwise, impacts from RRP activities may be overestimated significantly, and the analyses may reach spurious conclusions regarding associated hazards.

EPA does not indicate the averaging times of concern for potential health effects. The National Ambient Air Quality Standard for lead has a rolling 3-month average basis. Is this the averaging time associated with potential health effects from RRP activities? Many RRP activities are of relatively short duration, so any modeling to determine impacts should account for the duration of the activities of interest, as well as the averaging time for any potential associated health effects.

Like most atmospheric dispersion algorithms, AERMOD predicts only average (i.e., first-order expectations of) patterns, and does not accurately account for either short-term temporal fluctuations or short-term spatial heterogeneity (i.e., second-order expectations), of atmospheric dispersion. Potential exposure mischaracterization of relatively large magnitude (e.g., ~10-fold [Bogen and Gouveia 2008]) within modeled atmospheric dispersion "footprints" have been shown to be associated with ignoring expected second-order atmospheric dispersion effects over periods of several hours. To the extent that the durations of RRP activities are substantially shorter than 3 months, it is likely that AERMOD predictions may similarly mischaracterize impacts of deposition due to external RRP sources. Dispersion modeling for the Framework should discuss this issue, and possibly address it using proposed, relatively simple post-processing methods that can be applied to adjust the output of atmospheric dispersion models such as AERMOD that generate only first-order dispersion-modeling predictions (Bogen and Gouveia 2008).

The Framework mentions a variety of factors that may be incorporated in the Monte Carlo analyses. These include climate region, rain frequency, and obstruction adjustment, as well as several "receptor building characteristics" that include distance of receptor from renovated building, receptor use type, area of building, receptor location (urban or rural), location of receptor relative to renovated building, and height of receptor building. Issues associated with obstruction adjustment have already been discussed. Issues associated with climate regime and rain frequency could be addressed by selecting a variety of meteorological data sets that would span a range of geographic locations and climate categories. AERMOD can be run with urban or rural dispersion coefficients, to account, to some extent, for receptor location. AERMOD can also account for local surface characteristics and effects on meteorology through the specification of representative values of albedo, surface roughness, and Bowen ratio.

Interior Modeling

EPA has stated in the framework document that they will likely use a mechanistic mass-balance model, together with Monte Carlo analysis, to evaluate lead hazards from RRP activities in public and commercial buildings. This model assumes that the indoor air is well mixed and contains no concentration gradients in the space.

Mage and Ott (1996) have stated that there is no scientific basis to adjust modeling calculations with a simple mixing factor to account for rooms that are not well mixed. Rooms undergoing renovation, repair, and painting (i.e., the source rooms) will have significant concentration gradients during the lead-dust-generating work tasks, and this will likely invalidate the assumption of a well-mixed room. In addition, other work spaces in buildings will also have concentration gradients created by particle resuspension from foot-traffic areas and HVAC zones, and workstation design or layout. While there are some modeling techniques that can be used to overcome this issue, they will significantly add to the complexity of the models and overall uncertainties of the results. For example, use of multiple zones or compartments or use of computational fluid dynamics (CFD) modeling, as recommended by Mage and Ott, can be attempted. However, the framework document does not discuss the use of multi-zone modeling or CFD modeling as a likely option.

The model must provide an inventory of dust in the air and on the floor through time in order to arrive at lead dose. The large size and time-dependent nature of concentration gradients found in commercial buildings negate the assumption of heterogeneous environments, which assumption is necessary to estimate exposure over time under the static conditions assumed by the model.

Furthermore, the model as it stands now is very complex and requires estimates of central tendencies for some parameters, and distributions for others, that are the most sensitive to estimating exposure. Notwithstanding the large number of parameters that exist in commercial buildings that affect particle transportation and deposition on a daily basis, estimating values for renovations will be extremely difficult, due to the paucity of empirical data that could be used to validate an estimated distribution(s). Also, the model purposely ignores HVAC filtration, which is an important sink for particulates in commercial buildings.

In summary, predicting particle behavior and deposition with mathematical modeling from interior sources, such as from RRP in public and commercial buildings, has not been fully developed to the point where one can reliably use models to estimate exposures for purposes of hazard or risk assessment, as proposed in the Framework.

Monte Carlo Procedures to Characterize Exposure and Risk

The proposed Framework indicates that "EPA would assess elevations in lead exposure resulting from a broad range of scenarios, considering variations in types of renovation activities, building types, sizes and configurations, use and occupancy patterns, cleaning frequencies, etc., which are designed to be reflective of actual P&CB settings … in both

children and adults," and that because the scenarios are not equally likely, "EPA would provide a discussion of the relative frequency of these high incremental IQ changes," and the "hazard finding would be made based on an overall judgment of the frequency and magnitude of incremental health effect changes resulting from P&CB renovations." Two types of RRP activities are considered: interior and exterior. Within each category, the Framework defines a set of scenario variables, together with alternative values for each such variable. An external (or internal) scenario is defined by one among all possible combinations of values of the entire set of external- (or internal-) scenario variables. Conditional on each scenario considered, the Framework also defines a set of constants and variables that will be used in Monte Carlo analysis of exposure and dose-response characteristics associated with that particular scenario. Specifically, in Appendix A of the Framework (U.S. EPA 2014b), Table A-3 (Exterior Analysis Monte Carlo Inputs) lists a total of 65 distributed variables for use in Monte Carlo analysis of exterior scenarios, 34 of which are modeled as continuous random variables, and the remainder as random samples from discrete distributions. In Appendix B of the Framework (U.S. EPA 2014b), Table B-3 (Interior Analysis Monte Carlo Inputs) lists a total of 12 distributed variables for use in Monte Carlo analysis of interior scenarios, seven of which are modeled as continuous random variables, and the remainder as random samples from discrete distributions.

The Framework uses Monte Carlo analysis to estimate exposure and corresponding risk (i.e., "hazard"). To perform this last step, the Framework explains:

> "To analyze potential hazards, EPA would examine the various distributions across percentiles in a single scenario (Figure 1A) and across different scenarios (Figure 1B). The collection of these distributions helps account for the total variability in exposure owing to environmental, lifestyle, and biokinetic differences across the population. EPA would place the less frequent, high incremental IQ changes within the context of expected (mean) incremental IQ changes. Because each scenario is not equally likely, EPA would provide a discussion of the relative frequency of these high incremental IQ changes. The hazard finding would be made based on an overall judgment of the frequency and magnitude of incremental health effect changes resulting from P&CB renovations. … To perform the Monte Carlo analysis, each scenario would be run 20,000 times (where each run is referred to as an "iteration"). Preliminary testing indicates that 20,000 iterations would be appropriate in order to optimize the combination of accuracy and run-time efficiency."

The Framework thus makes clear that EPA intends to evaluate percentiles as well as the mean value of hazard estimated probabilistically using Monte Carlo methods.

Although the Framework makes reference to a total of 20,000 Monte Carlo realizations of each scenario-specific set of Monte Carlo input variables, Issue 1 ("Monte Carlo and Sensitivity Analyses") of the "Draft Peer Review Charge Questions" for the Framework (U.S. EPA 2014c) indicates that "All combinations of scenario variables have been modeled and a distribution of results for each scenario was developed by iterating 10,000 times for exterior scenarios and 3,000 times for interior scenarios from the sampled variables." These simulation sample sizes (3,000 and 10,000) appear to contradict the simulation size of 20,000 indicated in the Framework (as quote above). This contradiction needs to be resolved, and a consistent explanation needs to be provided concerning how Monte Carlo methods were or will be used to develop Framework-based exposure and risk estimates. As explained below, regardless of

which simulation sizes within this range are used, the fact that these values are relatively small may undermine the reliability of Monte Carlo results obtained and used to develop the Framework.

The probabilistic approach to hazard characterization summarized above raises the following questions, which should be resolved before the proposed Framework can be evaluated to assess its scientific merits and technical feasibility.

1. The first question concerns the issue of what is being represented by each of the input distributions to be used for probabilistic hazard analysis. Several authors, including EPA and NRC, have previously recommended that distinguishing between uncertainty and variability is required for this type of probabilistic analysis (Bogen and Spear 1987; Bogen et al. 2009; NRC 1994; EPA 2011). In light of this understanding of appropriate application of probabilistic methods, the Framework should specifically address how input variable distributions representing uncertainty would be distinguished systematically from those representing inter-individual variability. If the Framework intends all input variables to represent inter-individual variability, this should be stated explicitly and justified with an explanation of why none of the variables are considered to involve uncertainty relevant to the analysis. If not, how is joint uncertainty and variability to be evaluated in order to distinguish acceptable from unacceptable scenarios? The proposed Framework should be modified to provide details concerning how it proposes to address this general issue.

2. With so many input variables, statistical test(s) should be applied to ensure that intended multidimensional statistical independence (i.e., lack of significant inter-correlations) is actually realized in each Monte Carlo calculation performed, if such statistical independence is intended. Algorithms (e.g., Iman and Conover 1982) that typically are applied to induce a target correlation (including a target of zero-correlation) among multivariate random samples contain no built-in objective test to assess the magnitude of the extent to which a realized correlation matrix deviates from a user-specified target correlation matrix, or from a default identity (i.e., zero-correlation) matrix. The more that such Monte Carlo calculations are performed, the more likely it is that at least some will have realized correlations that deviate statistically significantly from any target correlation. For example, the Jennrich chi-square test (Jennrich 1970) can be used to assess the statistical significance of any realized deviation from a target correlation matrix.

3. As mentioned above, the proposed Framework indicates that preliminary EPA tests have indicated that a Monte Carlo sample size of $N = 20{,}000$ "would be appropriate in order to optimize the combination of accuracy and run time efficiency." As noted above, Issue 1 ("Monte Carlo and Sensitivity Analyses") of the "Draft Peer Review Charge Questions" for the Framework (U.S. EPA 2014c) indicates that even smaller simulation sizes were used, when it states, "All combinations of scenario variables have been modeled and a distribution of results for each scenario was developed by iterating 10,000 times for exterior scenarios and 3,000 times for interior scenarios from the sampled variables." The reliability of Monte Carlo results using sample sizes this low is dubious and needs to be verified, especially given the very large number of input

variables considered. It is possible that one or more combinations of relatively (but not extremely) unlikely values of some (perhaps even a small) subset of input variables may produce a very large upward shift in the value of modeled hazard. It is consequently impossible to characterize an upper (e.g., 95^{th} or 99^{th}) percentile, or even the mean value, of estimated hazard with any guaranteed degree of reliability using Monte Carlo methods, unless the number (n) of input variables is small, or the sample size (N) used is very large (e.g., $N >> 20{,}000$). This is true regardless of the Monte Carlo sampling technique used (e.g., Latin Hypercube vs. uniform sampling). The relatively large number (n) of variables (and corresponding sampling dimensions) involved guarantees that the parameter space that must be sampled to achieve reliability of Monte Carlo calculations is very large—very likely exceeding sample sizes of 3,000, 10,000, or 20,000 referred to in the Framework and in its Draft Peer Review Charge Questions. This general, well-known problem—commonly referred to as the "curse of dimensionality" in statistical and computer-science literature—arises because the volume of sample hyperspace grows (hence its sample density shrinks) exponentially as a function of the sampling dimension n (Bellman 1961). Mathematically, this dimensionality problem is related to the classic "coupon collector" problem of determining occupancy waiting times (Feller 1971; Bogen 2003). Importantly, this Monte Carlo sampling-efficiency problem pertains not only to continuous random variables, but also to nearly all other types of random variables, whenever a relatively large number of these must be sampled jointly.

To illustrate the dimensionality problem, suppose each of n input variables is known *a priori* only to be related positively and monotonically to predicted (e.g., mean or upper-bound) hazard. Suppose further that relatively high values of a subset of m (i.e., m-tuple subset) of these n variables, which all fall within a commonly defined upper P-percentile tail of the m corresponding probability distributions used to model variation in the subset of m variables, may potentially interact greatly (e.g., highly synergistically) to increase the value of corresponding hazard. Under these assumptions, any reasonably reliable assessment of uncertainty in estimated risk must be based on N risk realizations that jointly reflect ≥1 sample of each m-tuple combination of upper tail values. The likelihood that such a specific "upper-bound combination" (UBC) occurs clearly becomes very small quickly as m increases (e.g., for values of $P \leq 20\%$). For m equal to just 2 or 3, this small likelihood might possibly be balanced by a large upward shift in corresponding predicted hazard. The key point is that, absent *a priori* knowledge about the true distribution of hazard, there is no way of knowing that such a disproportionately increased risk level could be produced by a particular combination of m upper P-percentile tail values for m corresponding inputs, without actually sampling each potentially relevant combination at least once. For example, with $n = 65$ (the number of Monte Carlo input variables used by the Framework for external scenarios), there are a total of 2^{65}, or approximately 3.69×10^{19} possible unique combinations of upper-vs.-lower bound values for all 65 variables. Only $C(n,m)$ potentially relevant m-tuple UBCs need be sampled in order to include each at least once, assuming that the Latin hypercube method of random sampling is used, where $C(n,m)$ is the number of unique combinations of n different items taken m at a time. For example, $C(65,m)$ is approximately 6.8×10^{38}, 8.4×10^{57}, and 7.7×10^{76} for values of m equal to 2, 3, and 4, respectively. Note that

C(65,4) is ~0.1% of Eddington's cosmic number, which estimates the total number of elementary particles in the universe. Consequently, for Monte Carlo analysis of all exterior scenarios, the requirement of investigating such astronomically large values of sample size *N* can be avoided only if it can be established *a priori* that substantial synergism cannot possibly affect the hazard-response model, as would be true, for example, in the case of an entirely linear multivariate model. Without such an *a priori*, analytical investigation of fundamental departures from linearity associated with the entire hazard model, and explicit verification that Monte Carlo analysis is not meaningless due to high dimensionality of the problem being addressed, the application of Monte Carlo analysis provides only a façade of scientific rigor, without any actual substance.

To add power to addressing the dimensionality issue, an approximating (e.g., even two-point) probability mass function can be used to replace each continuous-variable input distribution, and then discrete probability calculus can be used instead of Monte Carlo sampling to obtain a probabilistic characterization of exposure or risk (e.g., Bogen 1995). However, even this approach fails when the number (*n*) of input variables becomes even moderately large (e.g., $n > 10$), unless the sample size *N* is also very large (e.g., $N >> 20{,}000$) (Bogen 2003). A mathematical "stopping rule" was proposed to define a value of *N* that addresses the dimensionality problem with a corresponding, specified degree of statistical confidence (Woo 1991). However, the mathematical proof offered for this rule was shown to be defective (Bogen 2003). Consequently, any generic specification of *N* by the proposed Framework will need to be justified using rigorous, peer-reviewed methods.

4. How will the output be used to distinguish between acceptable and unacceptable conditions? The specific combination(s) of population mean response and upper-bound individual response that would be used to distinguish acceptable from unacceptable scenarios are not specified in the proposed Framework. If the rationale for using any other response measure besides population mean response is an equity concern, how will the issue be addressed that extremely rare exposure scenarios (associated with very little likelihood) may have an unacceptably high level of risk to a 99%ile child at risk, but the number of such children expected to be at risk is so low that it is, for example, 99% certain that zero actual cases will arise?

Monte Carlo Case Study

The Framework provides a case study for which Monte Carlo methods were used to estimate blood lead levels in children. While this was intended as an example, it is unfortunate that EPA included this material, because it depicts unrealistic conditions and an approach that is largely grounded on judgment. Appendix C states that:

> The metrics are incremental IQ change, averaged over ages 1 through 7, for a child who experiences a renovation at age 1. In other words, the hypothetical child experiences the renovation at age 1, which results in a short-term spike in blood lead. The child's blood lead then returns to the pre-renovation level, typically within 1 year. For this preliminary

analysis, the child's AVERAGE blood lead, over ages 1 through 7, is then used to approximate the actual long-term impact on the child's IQ.

This is essentially a "playing with numbers exercise" that does not convey reliable information and may be misleading. Not only is the scenario inappropriate and unrealistic, but it conveys information to the public about exposures that may be incorrect given the assumptions that were made. If this is an example of the approach at a preliminary level, it points to a major problem with implementing the more complex Framework and with generating results that will be useful.

Risk Modeling Procedures

The proposed Framework specifically addresses the "short-term nature of the exposure resulting from renovation activities." For non-cancer health endpoints such as IQ-decrement associated with low-level lead exposure by the risk model(s) to be used, what is the biological basis for predicting that any biologically significant effect may occur from a transient (e.g., less than 2- or 4-week) exposure in the absence of any human or experimental-animal data demonstrating directly that such effects can possibly occur? For example, it is not clear that any such data were described in EPA's recent health assessment for lead (U.S. EPA 2013).

Summary and Implications

Based on the limited information provided by EPA, it is not possible to provide specific feedback regarding the nature or magnitude of likely errors associated with application of the models in the context of assessing impacts on blood lead levels from the RR&P activities. However, the technical comments on specific components of the modeling proposed in the Framework, provided above, illustrate the myriad challenges that EPA will face in implementing the Framework. In evaluating the application of both the proposed air modeling approach and the probabilistic exposure modeling, it is clear that the absence of strong input data and cautious application of the modeling methods could result in spurious or biased results. Due to the tendency to make "conservative assumptions" in the face of uncertainty, the modeling could then result in suggesting significant risk where none exists. Or it could well become clear that the effort is so fraught with uncertainty as to make it unusable. These considerations reinforce the recommendations provided in the early sections of this document that EPA would be well advised to perform a simple "bounding evaluation" or "plausibility analysis" to determine whether any more complex or sophisticated evaluation is even warranted. In the Proposed Framework, EPA states that "the impact … to total blood lead and bone lead from P&CB renovation scenarios is expected to be quite small." Given this anticipated outcome, EPA has failed to provide a robust technical basis to justify undertaking a highly uncertain and labor-intensive modeling effort in this context.

References

Abt Associates, Inc. 2014. Developing a concentration-response function for Pb exposure and cardiovascular disease-related mortality. Prepared for National Center for Environmental Economics, Office of Policy, U.S. Environmental Protection Agency. Abt Associates, Inc., Bethesda, MD.

Bellman R. 1961. Adaptive control processes: A guided tour. Princeton University Press, Princeton, NJ.

Bogen KT. 1995. Methods to approximate joint uncertainty and variability in risk. Risk Anal 15(3):411−419.

Bogen KT. 2003. Recommendations for SZ/TSPA model uncertainty analysis concerning the Yucca Mountain Project. UCRL-TR-201447, October 10, 2003. Lawrence Livermore National Laboratory, Livermore, CA. Available at: https://e-reports-ext.llnl.gov/pdf/302344.pdf

Bogen KT, Spear RC. 1987. Integrating uncertainty and inter-individual variability in environmental risk assessment. Risk Anal 7:427−436.

Bogen KT, Cullen AC, Frey HC, Price PS. 2009. Probabilistic exposure analysis for chemical risk characterization. Toxicol Sci 109(1):4–17.

Bogen KT, Gouveia FJ. 2008. Impact of spatiotemporal fluctuations in airborne chemical concentration on toxic hazard assessment. J Hazard Mater A 152(1):228–240.

CDC. 2013. Fourth national report on human exposure to environmental chemicals. Updated Tables, September 2013. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, GA.

Chisolm JJ, Jr. 1970. Childhood lead intoxication. Diagnosis, management and prevention. Med Times 98(9):92–106.

Feller W. 1970. An introduction to probability theory and it applications, Vol. I, 3rd ed. (revised printing). John Wiley & Sons, New York, p. 106.

Hara a, Thijs L, Asayama K, Gu Y, Jacobs L, Zhang Z, Liu Y, Nawrot T, Staessen J. in press. Blood Pressure in Relation to Environmental Lead Exposure in the National Health and Nutrition Examination Survey 2003–2010. Hypertension (in press)

Harriss LR, Ajani AE, Hunt D, Shaw J, Chambers B, Dewey H, Frayne J, Beauchamp A, Duvé K, Giles GG, Harrap S, Magliano DJ, Liew D, McNeil J, Peeters A, Stebbing M, Wolfe R, Tonkin A. 2011. Accuracy of national mortality codes in identifying adjudicated cardiovascular deaths. Aust N Z J Public Health35(5):466–476.

Herrett E, Shah AD, Boggon R, Denaxas S, Smeeth L, van Staa T, Timmis A, Hemingway H. 2013. Completeness and diagnostic validity of recording acute myocardial infarction events in

primary care, hospital care, disease registry, and national mortality records: Cohort study. BMJ 346:f2350.

Iman RL, Conover WJ. 1982. A distribution-free approach to inducing rank correlation among input variables. Commun Statist-Simulat 11(3):311–334.

Jennrich RI. 1970. An asymptotic Chi-2 test for the equality of two correlation matrices. J Am Statist Assoc 65(330):904–912.

Khalil N, Wilson JW, Talbott EO, Morrow LA, Hochberg MC, Hillier TA, Muldoon SB, Cummings SR, Cauley JA. 2009. Association of blood lead concentrations with mortality in older women: a prospective cohort study. Environ Health 8:15.

Mage DT, Ott WR. 1996. Accounting for nonuniform mixing and human exposure in indoor environments. In: Tichenor BA, Ed. Characterizing sources of indoor air pollution and related sink effects, ASTM STP 1287. American Society for Testing and Materials, pp. 263–278.

Menke A, Muntner P, Batuman V, Silbergeld EK, Guallar E. 2006. Blood lead below 0.48 micromol/L (10 microg/dL) and mortality among US adults. Circulation 114(13):1388–1394.

Møller L, Kristensen TS. 1992. Blood lead as a cardiovascular risk factor. Am J Epidemiol 136(9):1091–1100.

Menzie C, Kashuba R. 2013. A Conceptual Model for Cumulative Risk Analysis Using CVD as an Example. December 10, 2013 Symposium: Proposed Methods for U.S. EPA's CRA Guidelines Society for Risk Analysis 2013 Annual Meeting Baltimore, MD.

rvland

NCHS. 1984. Blood lead levels for persons ages 6 months–74 years. United States, 1976–80. Vital and health statistics. Series 11, No. 233. U.S. Department of Health and Human Services, Public Health Service, National Center for Health Statistics, Washington, DC.

NRC. 1994. Science and judgment in risk assessment. National Research Council), National Academy Press, Washington, DC.

NTP. 2012. NTP monograph on health effects of low-level lead. U.S. Department of Health and Human Services, National Institute of Environmental Health Sciences, National Toxicology Program, Research Triangle Park, NC.

Pirkle JL, Brody DJ, Gunter EW, Kramer RA, Paschal DC, Flegal KM, Matte TD. 1994. The decline in blood lead levels in the United States. The National Health and Nutrition Examination Surveys (NHANES). JAMA 272(4):284–291.

Schober SE, Mirel LB, Graubard BI, Brody DJ, Flegal KM. 2006. Blood lead levels and death from all causes, cardiovascular disease, and cancer: Results from the NHANES III mortality study. Environ Health Perspect 114(10):1538–1541.

U.S. EPA. 2011. Exposure factors handbook, 11th ed. EPA/600/R-090/052F, September. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

U.S. EPA. 2013. Integrated science assessment for lead. EPA/600/R-10/075F. U.S. Environmental Protection Agency, Washington, DC.

U.S. EPA. 2013. Integrated science assessment for lead. EPA/600/R-10/075F. U.S. Environmental Protection Agency, National Center for Environmental Assessment – RTP Division, Office of Research and Development, Washington, DC.

U.S. EPA. 2014. Framework for identifying and evaluating lead-based paint hazards from renovation, repair, and painting activities in public and commercial buildings. U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxics, Washington, DC. Advanced Notice of Proposed Rule. Fed Regist 79(104):31072–31074.

Weisskopf MG, 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.

Weisskopf MG, Hu H, Sparrow D, Power M. 2013. What you don't see can hurt you: Selection bias in cohort studies—Are environmental studies at particular risk? Abstract Number 4606. ID: O-3-39-03. Environment and Health, Basel, Switzerland, 19–23 August. Available at: http://ehp.niehs.nih.gov/ehbasel13/o-3-39-03/.

Wexelman BA, Eden E, Rose KM. 2010. Survey of New York City resident physicians on cause-of-death reporting, Prev Chronic Dis 2013;10:E76.

Woo G. 1991. A quitting rule for Monte Carlo Simulation of extreme risks. Reliability Engineering and System Safety 31:179–189.