13 November 2009 Preliminary Review Comments from CASAC Carbon Monoxide Review Panel on EPA's *Integrated Science Assessment for Carbon Monoxide: Second External Review Draft (September 2009)*

Comments received:

Dr. Beate Ritz **2**
Dr. Paul T. Roberts **5**
CQ3 consensus response (draft) **10**

Dr. Beate Ritz

1. *The framework for causal determination presented in Chapter 1 was developed and refined in other ISAs (e.g., the PM ISA). During previous reviews, CASAC generally endorsed this framework in judging the overall weight of the evidence for health effects. Please comment on the extent to which Chapter 1 provides necessary and sufficient background information for review of the subsequent chapters of the CO ISA.*

This chapter has improved but still does not adequately address and present methodologic concepts in epidemiology and, thus, lacks clarity in how epidemiologic studies are evaluated and determined to be "high or low quality studies" as necessary for applying the criteria listed in Table 1.2 (i.e. for assessing the weight of evidence for causal determination).

A minor point: I previously recommended using the more appropriate term '***effect measure modification'*** instead of 'effect modification' but the wording in this chapter has not been corrected. More importantly, however, while there is some improvement, the authors of this chapter still seem to not be fully understanding nor formulating adequately some of the issues involved in confounding and confounder control. They claim on page 1-15 that "deciding which variables to control for in a statistical analysis of the association between exposure and disease or health outcome depends on knowledge about possible ***mechanisms*** and the distribution of these factors in the population under study. Identifying these ***mechanisms …***". Knowledge of 'mechanisms' may help, but such knowledge is not needed to decide whether a covariate is a potential confounder neither is it necessary to know mechanisms to assess confounding. It is furthermore completely obscure what the authors mean by the following sentence on page 1-15 "adjustment for potential confounders can be influenced by differential exposure measurement error"; here they seem to confuse error in measuring confounding variables with error in exposure assessment? Finally on page 1-13 the second sentence under 1.6.3. "Uncertainty can be defined…." seems to confuse precision and validity or at least does not acknowledge that these are two different concepts that have a different place in judging study results. These confusions of concepts does not instill much confidence in the ability of staff who wrote this chapter to judge epidemiologic studies adequately according to established criteria for study validity and precision (both contributing to accuracy); this is further confirmed by the chapter 5 qualitative reviews that are still grossly lacking in consistency and interpretation of epidemiologic results.

The criteria for causal determination detailed in table 1-2 are similar to those used by the IOM and the International Agencies for Research on Cancer. Yet, they leave open what the criteria are for deciding that a study is high quality (for example, confounding is a bias, so why list bias and confounding apart from chance?) and it is also unclear what is meant by "replicated" results and why this would be a criterion. Again, without a standardized approach to the review of epidemiologic studies or a quantitative meta-analysis based review, these criteria remain ambiguous. Since the epidemiologic literature on criteria air pollution health effects has multiplied greatly in the past decade, it would be appropriate if EPA staff abandoned qualitative reviews in favor of quantitative effect estimates based on meta-analytic procedures to draw inferences about the scientific literature and used standardized and transparent rules for data abstraction. Such a systematic and quantitative procedure requires making the authors'

assumptions explicit rather than allowing authors to emphasize studies they agree or disagree with and to pick the results they like to emphasize over others. Such quantitative reviews could be contracted out to entities that are able to conduct meta- or pooled analyses.

5. *Chapter 5 presents information on cardiovascular, central nervous system, developmental, respiratory, and mortality outcomes following exposure to CO. To what extent are the discussion and integration of toxicological, clinical, and epidemiologic evidence for these health effects scientifically sound, appropriately balanced, and clearly communicated? Are the tables and figures presented in Chapter 5 appropriate, adequate, and effective in advancing the interpretation of these health studies?*

In Chapter 5, the qualitative description of epidemiologic studies improved somewhat but is still inadequate; the level of detail devoted to each study in the text seems still arbitrary and the information provided in tables and figures selective without being systematic; for example why did the authors decide to present in Figure 5-8 the citywide and negative associations for the Australian CO study of PTB (Jalaludin) and not the positive associations for births within a 5km radius of a monitor. The review of birth weight and air pollution is lacking a discussion of the difference between LBW and term LBW sorely needed since LBW includes preterm birth outcomes that are then discussed separately and studies examining LBW are possibly more comparable in their results to those examining PTB; only term LBW is a mutually exclusive outcome. Also the measure of birthweight as a continuous outcome compared to the dichotomous variables LBW and PTB deserve some more general introduction about their general value (similar to the discussion of SGA versus IUGR), i.e. do we really expect the whole birthweight distribution to shift according to ambient air pollution exposures or only the most susceptibly infants to be affected.

Surprisingly, there is still a lot of information I requested in my first review missing from this new draft. This includes the following: no information is provided in the tables concerning the type of study design employed (e.g. Table 5-12). I also already mentioned previously that many of tables report mean CO levels and mention 24 hrs or 8 hrs in brackets; this is misleading for pregnancy outcome studies in which the averages are for trimesters, weeks, or months (e.g. the Ritz et al. (2000) study of PTB is listed in table 5-12 as having a Mean CO of 2.7 ppm for the 6-9 am period – however this mean represents a mean over the whole *first month* of pregnancy and the Wilhelm and Ritz (2005) study mentions a 1.4 ppm mean for 24 hrs but this is in fact a *first trimester* mean of 24 daily measurements; the way this data is shown now the bracketed 24 hour mention seems to imply similar averaging period and comparability in effect estimates. I also mentioned already previously that while the Ritz et al 2007 study is listed in table 5-12 no results for this study are presented in figure 5-6. I had also recommended to rescale quartiles to a continuous estimates rather than leaving results from important papers out of a figure that gives an overview over all study results.

Also I mentioned previously that according to the text accompanying the figures, the estimated increase in CO presented have been ‘standardized’, however, how this might have been done across so many different study types and averages for differing exposure periods (rather than 24 hour averages as the authors of these chapters seem to imply) has not been

explained. Also, in figure 5.1 the title says that the effect estimates have been standardized to a 1ppm increase in ambient CO for 1-hr max CO concentrations, 0.75 ppm for 8-h max CO concentrations and 0.5 ppm for 24 hrs avg CO concentrations, but the figure does not tell us which scale has originally been used in which study and it might be questionable whether effect estimate sizes based on these different scales and based on different length lag periods are comparable to each other (indicating which study used which scale would be informative.

There are also sentences in this review chapter that are plainly wrong, e.g. on page 5-71 and OR of 1 (95% CI 0.96-1.04) is called a positive association.

Dr. Paul T. Roberts

Preliminary Comments on 2nd draft ISA by for CASAC CO Panel, November 12, 2009

ISA Charge Question 2. Chapter 3 has been revised and expanded in response to Panel comments regarding climate, monitoring, spatial variability, and exposure.

2b. Additional detail has been provided regarding ..... Please comment on the usefulness of these revisions in characterizing the information provided by the CO monitoring network.

In general, the expanded discussions in the 2nd draft ISA Chapter 3 on CO detection limits, monitoring details, and spatial CO characteristics are very useful in characterizing the information provided by the CO monitoring network and in qualifying the data for use in exposure estimations. My detailed comments are provided below.

In discussion of non-anthropogenic CO emissions on page 3-5: it is confusing in the first paragraph (starting on line 3) to have fire emissions of about 13% (14.5 MT) shown in Figure 3-1, but biogenic emissions of about 5% not shown in Figure 3-1, and the text implies that the geogenic emissions are included (in the miscellaneous category?). This is confusing to the reader and makes it difficult to compare these smaller, but still important sources. Please add biogenics to Figure 3-1 and make it clear what is included.

Comments on discussion of Hudman et al and Figures 3-3 and 3-4, starting on page 3-5 at line 21: First, I suggest that this be a new paragraph; it is a different topic from the non-anthropogenic emissions. In addition, the CO from oxidation of VOCs, of isoprene, and of other biogenic VOCs (see lines 27-29), which are apparently huge in this simulation, relative to the anthropogenic CO emissions, have not been discussed before. These secondary emissions sources needed to be discussed in the overall context of CO emissions first (or put the general discussion in with the Climate text, Chapter 3.3.1. In addition, the potential influence of this huge source of CO on the results of the simulation and the conclusion needs to be discussed. I also suggest that Figure 3-3 be dropped, since I find it hard to compare these colored spatial plots. In contrast, Figure 3-4, as an example of the results, gets the point across that reducing the anthropogenic emissions by 60 % made significantly better comparisons with measurements. Maybe also add a statement saying that other results in the paper support this general conclusion.

Additional comment on the paragraph on page 3-5, lines 21-29: I suggest that a sentence be added translating the Tg amounts to MT, so that these emissions can be placed in context with the rest of the emissions discussion.

Comments on page 3-7, lines 13-15: I suggest that a comment be added (either in the text or the figure caption) about general transport winds being from west to east at this latitude, thus carrying the emissions from the Alaska fires across Canada and the northern US and into the north Atlantic, as shown in Figure 3-5 (assuming the data support this).

Page 3-11, lines 9-10: The comment on the significant quantities of aromatics in gasoline is likely no longer true, since regulations have significantly reduced aromatics in gasoline (and these references are old). I suggest that these comments be modified to say that this used to be the case, but less so with current fuel content.

Page 3-12, lines 8-9: The comments on limited mixing between the hemispheres would benefit from an additional comment that northern hemisphere CO emissions are significantly larger than the emissions in the southern hemisphere, plus a representative reference.

In general, the expanded discussions on monitor detection limits and monitoring locations in Chapter 3.4 (pages 3-18 to 3-31 and associated Annex figures and tables) are critical to understanding CO concentrations and are an important addition to the ISA. However, these limitations are still often left out of the discussions on CO concentrations in Chapter 3.5. See my detailed comments below.

Page 3-19, lines 20-23: This discussion on the needs for trace-level CO measurements should include the use of low-level CO data for improved exposure estimates at current ambient concentrations in many locations in the US.

Table 3-2: Please fix the first row of the table. I think there should be a header row labeling the columns as "Parameter" and "Specification", for example, plus the current first row should be part of the body of the table and left justified in each column.

Page 3-20 lines 3, 12, and 18: The LODs listed on these lines are not the same as listed in the referenced Table A-1 in the Annex, which lists the LOD of the trace-level monitor as 0.02, not 0.04. Thus, the value listed as 50% of the LOD on line 6 should be 0.01.

Page 3-20, line 5: I suggest that this line start with "When the monitored value is below the LOD, some states...

Page 3-20, line 14, page 3-21 and Figure 3-8: This discussion on a comparison of older and newer monitors with specific quantifications is a good one, and important to include here. However, the last sentence of the paragraph (lines 6-8) does not make sense to me. Also, the CO axis of the figure needs to be labeled and the units of the time axis needs to be added (hours since some start time, or?).

The limitation of the LOD issues discussed in Chapter 3.4 need to be added in several places in Chapter 3.5, including at the beginning of the sentence that starts on line 11 of page 3-36, in Figures 3-7, 3-8, and 3-9. In particular, statements similar to the paragraph at lines 13-18 on page 3-43 (good job there!) could be added to address this issue at these locations in the text and in the conclusions (3.7.3) and the summary of conclusions (2.1).

Page 3-41, lines 1-3: What was the cause of the 10.9 ppm CO measured at the Newkirk, OK site? This seems like an unusual concentration.

Page 3-43, lines 13-18: This is a good qualifying paragraph which is needed here and other places. On line 15, I suggest the following wording: "...in large part very near or below the detection ..."

Figure 3-18 on page 3-45 (and Figure 3-20): How is the data below LOD treated for this and similar figures? Might this influence the lower ends of the box-whisker plots? I suggest that a comment on this be added.

Chapter 3.5.1.2: Add a note that all of these monitors in LA and Denver are older, higher LOD, monitors, and add some comments similar to the comments at lines 13-18 on page 3-43.

Page 3-60, lines 14-15: I don't see how the data shown in Figure 3-24 can lead to a statement that includes the following words "...near-road CO concentrations.." .

Page 3-76, lines 5-6: I suggest this sentence be moved to the end of the paragraph at lines 12-21 on page 3-77. In addition, the sentence should read "....analogous to Figures 3-36 and 3-37 for ....in Annex A, Figures A.44 to A.48."

Page 3-77, line 11: I suggest these are meteorological, not micrometeorological factors.

2c. The section on exposure assessment has been reorganized to provide information on ..... Does the Panel consider that the sources of exposure error have been appropriately characterized, and agree with the revised conclusions regarding the impact of exposure error due to spatial variability and the presence of CO as part of a combustion-related mixture on health effect estimates from time-series epidemiological studies?

In general, the expanded discussions in the 2nd draft ISA Chapter 3.6 on sources of exposure and exposure assessment are a great improvement and are very useful in characterizing the potential impacts of exposure error. My detailed (minor) comments are provided below.

Page 3-93, lines 18 and 29: Please explain or modify the terms "driven cavity" and "posterior probability distribution function".

Page 3-94, lines 6-23, section on Land Use Regression Models: This section is still very limited in scope and does not represent the wide range of results from the literature. Admittedly, much of the LUR work in the literature is on pollutants other than CO, but the types of conclusions regarding what methods work and how they relate to estimating pollutant concentrations are directly applicable to CO. See the list of references I suggested last time (re-listed at the end of my comments), plus there must be many more than I could easily find.

Page 3-98, line 11: How can a regression coefficient be 1.99 (greater than 1.0)? Also, the results in lines 9-11 and in lines 13-14, although from the same reference, seem inconsistent; please explain how they are consistent or different.

Page 3-114, line 26: Table A-1 says the LOD for trace-level FRMs is 0.02, not 0.04 as stated here.

Page 3-115, lines 13-17: Please add the limitation statement on LOD to this section regarding characteristic concentrations.

Page 3-117, line 15: I suggest that the word "nearby" be added, so that the sentence would read "...at a location with few nearby CO sources could...".

In summary, Chapter 3 of this 2nd External Review Draft of the ISA clearly conveys and appropriately characterizes the atmospheric science and air quality analyses. The information provided regarding CO source characteristics, CO chemistry, policy-relevant background CO, and spatial and temporal patterns of CO concentrations accurate are relevant to the review of the CO NAAQS.

Minor edits and typos in the 2nd draft ISA:

- page 3-2, line 22: word near end of the line should be "inherent"
- Page 3-41, line 16: suggest that "medians" be replaced with "median correlation coefficients (r)"
- Figures 3-17 and 3-19: I can barely make out the lines for the highways (whereas the ones in the Annex are fine); lease make darker.
- Make bolder the lines separating the scales in Tables 3-10 and 3-11; it is currently difficult to read.
- Page 3-72, line 2 should read: "...as shown in Figure 3-6." not in Figure 3-32.
- It is very hard to see the 95th and 5th percentile lines in Figures 3-33 and 3-34; please make darker or bolder.
- Page 3-85, lines 17 and 24-25: I suggest that you use words other than "fidelity" in line 17 (and line 4 of page 3-86) and "fraught" in line 25; maybe "accurately" and "are difficult". Also, add a comma after troposphere in line 24.

Selected, easy for me to find, references for Land Use Regression and spatial mapping (see above discussion in Chapter 3.6.3):

Gauderman, Avol, Lurmann, Kuenzli, Filliland, Peters, and McConnell "Childhood Asthma and Exposure to Traffic and Nitrogen Dioxide, Epidemiology 2005; 16, 737-743.

Ross, Jerrett, Ito, Tempalski, and Thurston "A land use Regression for predicting fine particulate matter concentrations in the New York City region", Atmospheric Environment 41 (2007) 2255-2269.

Hoek, Beelen, Hoogh, Vienneau, Gulliver, Fischer, and Briggs "A review of land-use regression models to assess spatial variation of outdoor air pollution" Atmospheric Environment 42 (2008) 7561-7578.

Henderson, Beckerman, Jerrett, and Brauer "Application of Land Use Regression to Estimate Long-Term Concentrations of Traffic-Related Nitrogen Oxides and Fine Particulate Matter ES&T 2007, 41, 2422-2428.

Molitor, Jerrett, Chang, Molitor, Gauderman, Berhane, McConnel, Lurmann, Wu, Winer, and Thomas “Assessing Uncertainty in Spatial Exposure Models for Air Pollution Health Effects Assessment EHP vol 115,no 8, August 2007.

Popawski, Gould, Setton, Allen, Su, Larson, Henderson, Brauer, Hystad, Lightowlers, Keller, Cohen, Silva, and Buzzelli “Intercity transferability of land use regression models for estimating ambient concentrations of nitrogen dioxide” J Exposure Science & Environmental Epidemiology (2008), 1-11.

CQ3 consensus response (draft)

by Drs. T. Dahms, M. Kleinman and M. Hazucha

Dosimetry and Pharmacokinetics of Carbon Monoxide of the Second External Review Draft of the ISA for Carbon Monoxide

Charge: *"In response to comments from the CASAC CO Panel, material has been added to Chapter 4 describing comparisons among predictive COHb models, the relative influence of differing exposure scenarios on COHb concentration, and endogenous CO production rates in individuals with various diseases and conditions. Please comment on the usefulness of this information in illustrating the factors influencing COHb kinetics and potential COHb levels under various scenarios".*

Generally, we found the revised and substantially expanded Chapter 4 of the second draft quite comprehensive and certainly very useful in illustrating various physiologic factors and disease states that can influence blood level of COHb.

Identification of pollutant-specific associations with health endpoints in epidemiologic studies requires complex analyses. Utilization of COHb predictive models can help identifying such associations and provide further insight into pathophysiology of responses. We suggest exploring the usefulness of this approach, possibly including a discussion of how having a limited number of monitoring stations may impact on exposure estimations.

The section 4.2 describes in adequate detail various COHB predictive models. However, despite addition of section 4.2.3 Model Comparison, which discusses some limitations or advantages of the respective models, it is still unclear to us (1) how these different models will perform under the same temporal exposure scenario of 30- 60 min duration with occasional peak CO concentrations and (2) which one is the most accurate in predicting venous COHb. Several models, e.g., Smith et al, 1994, Bruce and Bruce, 2008, Gosselin et al., 2009 as well as the non-linear CFKE used by EPA in the APEX model seem to be most suitable for such comparison and evaluation.

Besides CVD there are other large population groups of "potentially" at-risk individuals from CO exposure such as patients with anemias and COPD. Although there are no laboratory studies on the effects of CO exposure on these groups at ambient concentration, these individuals may be more vulnerable to CO because their disease impairs the key sites of action of CO. The application of COHb predictive models to such disease conditions would be helpful in determining the extent of risk in such populations.

There are some inconsistencies between tables, figures and text presented in the discussion of the QCP model (section 4.2.4). They need to be reconciled and the section shortened by dropping less relevant material.

Addition of section 4.3.4 COHb Analysis Methods in this draft is very helpful by pointing to limitations and inaccuracy of some of the instruments used to measure COHb. Since the differences in COHb determination between methods may be substantial we suggest indicating the method/instruments used to determine COHb in tabulated studies as well as other key studies discussed in the text.