Key Issues in the February 2014 Draft of the Health Risk and Exposure Assessment for Ozone
– Advance Written Materials for CASAC Meeting –

Prepared on behalf of UARG
by
Anne E. Smith, Ph.D.
NERA Economic Consulting
March 13, 2014

The Long-Term Mortality Risk Estimates Are Subject to Very Large Sensitivity Not Highlighted in the Draft HREA

The present draft of the Health Risk and Exposure Assessment (HREA) contains, for the first time, estimates of respiratory mortality risks based on epidemiological associations reported in Jerrett *et al*. (2009).[1] The HREA's core risk estimates for long-term respiratory mortality are based on one of several ozone-mortality associations estimated in that paper: a model that also controls for $PM_{2.5}$ (the "2-P" model). At a national level, that risk estimate is 44,820 deaths, based on observed ozone levels from 2006 to 2008, fused with modeled 2007 concentrations.[2]

Later in the HREA, two sensitivity analyses to the above core long-term mortality estimate are briefly mentioned. One uses the paper's "1-P" model; the HREA provides a table indicating that its city-specific O_3-attributable risks are about 30% lower than the core estimate from the 2-P model.[3] The second sensitivity analysis uses region-specific 1-P coefficients reported in Jerrett *et al*. This analysis finds substantial city to city sensitivity. However, rather than discussing the implications of this large sensitivity when risks are considered at the regional level, the HREA dismisses them with the following words: "while the results of this sensitivity analysis point to the potential for regional heterogeneity in the long-term O_3-attributable mortality effect estimate, we do not have significant confidence in the regionally-based risk estimates themselves given the relatively large confidence intervals associated with those estimates."[4]

A third and far more interesting sensitivity analysis could have and should have been performed using yet other results in Jerrett *et al*.: the effect of estimating the risks using results of a model that detected an effects threshold within the data. The HREA does not even suggest that a

[1] Jerrett *et al*. 2009. "Long-Term Ozone Exposure and Mortality." *The New England Journal of Medicine*, 360(11):1085-1095.

[2] The HREA reports this estimate in rounded form as 45,000 deaths (see HREA, Table 8-1, p. 8-7). The more precise value reported here is from NERA's own BenMAP model runs using EPA's input data, provided at NERA's request.

[3] HREA, Table 7-15, p. 7-83. We confirmed this ~30% difference by making our own estimate at the national level using the 1-P coefficient; the 44,820 deaths under the 2-P model were reduced to 33,598 under the 1-P model.

[4] HREA, p. 7-79.

sensitivity analysis might be warranted regarding the threshold, even though that was a specific issue addressed in the paper. In fact, with respect to the question of a potential threshold, the HREA only states that evidence of a threshold is "limited" in the epidemiological literature generally, and does not mention that evidence of one was reported in Jerrett *et al.*[5] In fact, Jerrett *et al.* find that a model with a threshold of 56 ppb provides a better fit to the data than the no-threshold model (using a 1-P model in both cases). The HREA should have provided analyses showing how very sensitive its long-term risk estimates are to this better-fit model.

Using BenMAP, we find that the alternative threshold model makes an enormous difference to risks estimates. National risks computed using the 56 ppb/1-P threshold model are 95% less than the no-threshold 1-P result, and 97% less than the no-threshold 2-P result that is the core estimate emphasized in the HREA. Figure 1 provides a map showing the dramatic differences in estimates of elevated risk of respiratory mortality across the nation projected by each of these three alternative models. The 2-P and 1-P no-threshold models (Figures 1A and 1B, respectively) imply that over 6% of respiratory deaths were hastened by ozone levels in 2006-2008 across the entire nation. In stark contrast, the threshold model indicates that risk is elevated by more than 6% only in a part of Southern California, and that risk is zero in the majority of the U.S.

It is also noteworthy that under the 56 ppb threshold model, 10 of the 12 cities studied in the HREA are projected to have zero risk for long-term respiratory mortality at the current standard of 75 ppb.

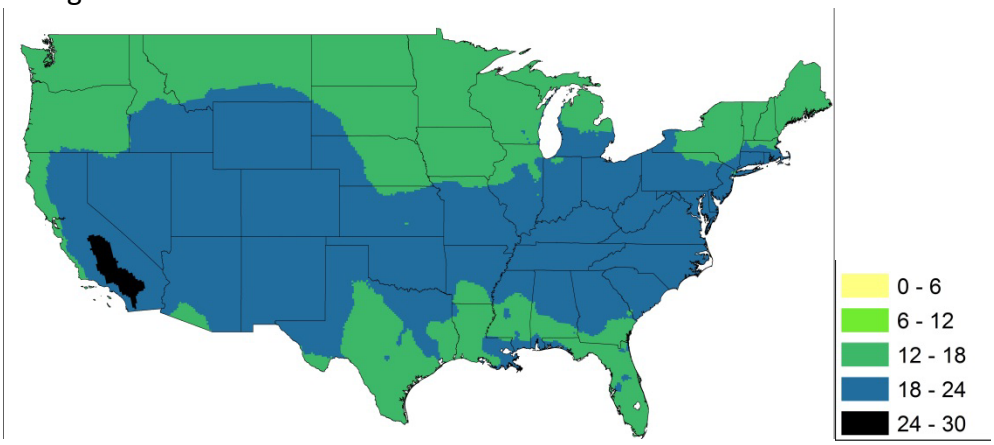
Jerrett *et al*. notes that the 56-ppb threshold model is not statistically significantly different from the zero-threshold model, but we note that the threshold model is nevertheless *the best fit*. In fact, all thresholds tested in that paper from 45 to 57 provided a better fit (*i.e.,* a higher log-likelihood function value) than the zero-threshold model, and the 56 ppb case had the best fit of all. The paper dismisses this best-fit threshold estimate as not statistically significant because its *p* value is 0.06. While this *p* value means that the no-threshold model assumption cannot be rejected with 95% confidence, *it can be rejected with 94% confidence*. When the question of significance is so borderline, it is not reasonable to ignore the implications of the alternative models in one's *risk* calculations.[6]

[5] HREA, p. 2-13.

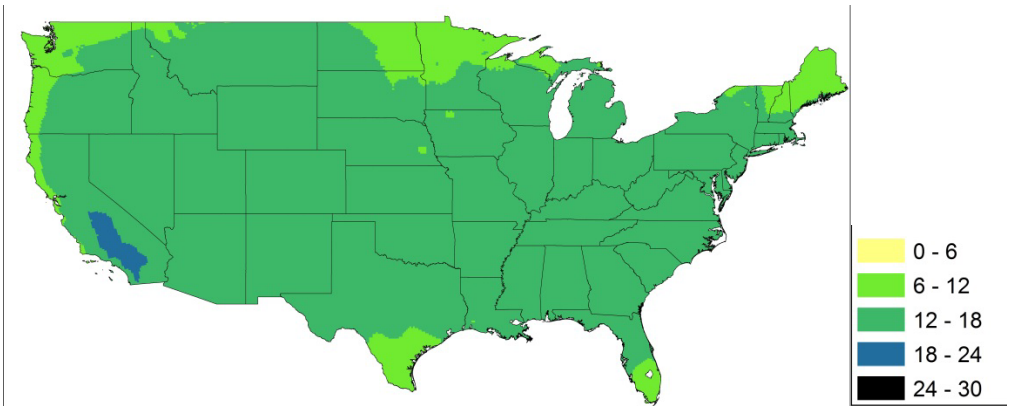
[6] Jerrett *et al.* provides no information to assess the statistical merits of any of the 1-P models compared to the 2-P model. It also provides no information on whether a threshold model might be a best fit in the 2-P case as well.

Figure 1. Sensitivity of Long-Term Mortality Estimates to Alternative Models in Jerrett *et al*.
(Maps Showing Percent of U.S. Respiratory Mortality Attributed to Ambient Ozone 2006-2008)

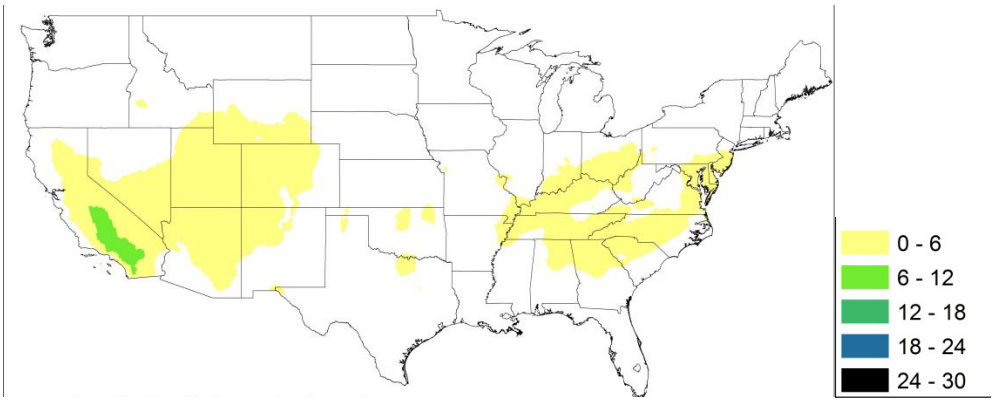
A. Using Jerrett *et al*.'s 2-P Model:



Using Jerrett *et al*.'s 1-P Model:



B. Using 56 ppb Threshold Model



Despite the 56 ppb threshold model being a better fit, the HREA has adopted the no-threshold model for its core analysis. The only statistical principle that would support adopting the zero-threshold model as the core result rather than a better-fitting threshold model would be if medical professionals held an *a priori* belief is that no threshold exists. In fact, the opposite is the case: medical professionals widely expect a diminution of risk at lower concentrations, and the surprise has been that air pollution epidemiological studies have not generally identified thresholds. The finding of a non-linearity in the association at lower ozone levels by Jerrett *et al*. is consistent with *a priori* expectations, and it should not be dismissed simply because one has "only" 94% confidence that a threshold exists in these data. Nevertheless, even if the confidence level were much lower but the threshold model still provided a better fit, the appropriate action for the risk analyst would be to highlight the huge policy significance of this very fine line among the alternative models. As is shown above, the alternative models reported in Jerrett *et al.* produce estimates of as-is risk that differ by over a factor of 25, and range qualitatively from a localized and modest risk increase to a universally large risk. This is an abject degree of uncertainty about what this paper's findings tell us about long-term respiratory risk from ozone in the U.S. As there are no other publications to support any particular conclusion out of this paper, long-term ozone mortality risk estimates based on Jerrett *et al.* (2009) do not merit attention in the HREA. If they are to be included at all, the above sensitivity across the threshold assumptions should be the <u>core</u> result that is presented.

Short-Term Mortality Risk Estimates Based on Smith *et al*. (2009) Were Not Reproducible Due Errors in HREA

The HREA's core short-term mortality risks are based on Smith *et al*. (2009).[7] We were not been able to replicate any of the twelve city-specific or the national core short-term mortality risks reported in the HREA when using EPA's ozone concentration data (provided to NERA by EPA) with the Smith *et al.* risk coefficients and baseline incidence rates as documented in the draft HREA. After describing this replicability problem to EPA, EPA determined that the HREA table documenting the short-term risk coefficients for the urban area risk estimates (*i.e.*, the table in Appendix 7A) was incorrect. On March 12 (the day before this written material for CASAC must be submitted), EPA staff emailed NERA a corrected version of the beta coefficients, which finally allowed us to replicate EPA's short-term city-specific risk estimates. However, these risk coefficients cannot be obtained from the paper itself, and EPA communications with the paper's author(s) to obtain them do not appear to be available in the docket or among posted Technical Support Materials to confirm that the new values are now correct. We can only confirm that the values in the revised Appendix 7A of March 12, 2014 were indeed used by EPA to produce its core city-specific short-term risk estimates. (The city-specific coefficients used in the HREA's national calculations also cannot be obtained from the original paper.) This experience highlights the tenuous quality of the massive number of risk estimates in the HREA when it relies on input assumptions not in any published document.

[7] Smith *et al*. 2009. "Reassessing the relationship between ozone and short-term mortality in U.S. urban communities." *Inhalation Toxicology*, 21(S2):37-61.

With respect to the national estimate of short-term mortality based on Smith *et al.*, EPA concluded that NERA's estimate was correct, and that the estimate reported in HREA Table 8-1 is incorrect. (Specifically, the estimate of 16,000 deaths in Table 8-1 should have been reported as 15,000, given that the precise estimate from BenMAP is 14,556.)

Short-Term Mortality Risk Estimates Based on Smith *et al*. (2009) Misrepresent Those Authors' Conclusions

Concerns with replicability aside, we note that the use of coefficients from Smith *et al.* (2009) without extensive sensitivity analysis is at odds with the conclusions of that paper's authors.

The 2012 preliminary draft of the ozone HREA did not mention the possibility of relying on Smith *et al*. (2009), and in fact provided estimates based on Zanobetti and Schwartz (2008)[8] and Bell *et al*.[9] This draft of the HREA not only introduces use of Smith *et al.* for the first time, but also now relies on this paper for its core risk estimates. Risk estimates from Zanobetti and Schwartz appear only in an appendix, and none based on Bell *et al*. are provided. The HREA does not provide any reason for this change in its plans. We find it useful to quote the last paragraph from Smith *et al*.:

> *In summary, it is our view that estimates of the associations between ozone and mortality, based on time-series epidemiologic analyses of daily data from multiple cities, reveal important still-unexplained inconsistences and show sensitivity to modeling choices and data selection. These inconsistencies and sensitivities contribute to serious uncertainties when epidemiological results are used to discern the nature and magnitude of possible ozone-mortality relationships or are applied to risk assessment.*[10]

In reaching this conclusion, the paper discusses a number of issues that could dramatically affect risk estimates that might be based on such epidemiological evidence. For example, it notes that there is apparent non-linearity in the ozone relationship, with the risk coefficient greater above a city's median temperature[11]; but on the other hand that ozone-mortality association does not appear to be increased at higher ozone concentrations.[12] It notes that the ozone-mortality association seems to exist only in the northeast and industrial Midwest.[13] It

[8] Zanobetti A. and Schwartz J. 2008. "Mortality Displacement in the Association of Ozone with Mortality: An Analysis of 48 Cities in the United States." *American Journal of Respiratory and Critical Care Medicine*, 177:184-189.

[9] Bell M. *et al*. 2004. "O_3 and Short-term Mortality in 95 U.S. Urban Communities, 1987-2000." *JAMA*, 292:2372-2378.

[10] Smith *et al*. at p. 59.

[11] Smith *et al*. at p. 56.

[12] Smith *et al*. at p. 52.

[13] Smith *et al*. at p. 48.

also notes that the finding of a stronger statistical association at the lowest ozone levels is "mysterious" given that clinical studies do not find effects at those levels.[14]

These and other issues lead the authors to state:

> *[W]e believe that the heterogeneity and sensitivity of ozone effect estimates to a variety of covariates leaves open the issue of whether or not ozone is causally related to mortality. Consequently, the question arises whether any particular ozone-mortality effect estimate can reliably be used to predict mortality reductions that would ensue from specific ozone reductions.*[15]

The HREA, on the other hand, takes just one set of city-specific results from the paper and provides "core" risk estimates based solely on those. The fact that those city-specific results are statistically insignificant in all but four of the twelve cities is not highlighted, and none of the non-linearities noted in the paper are explored at all in the HREA. Furthermore, the HREA provides a national risk estimate from this paper, despite the author's position that "any national summary, even a population-weighted average, will necessarily conceal the still-unexplained heterogeneities"[16] and "the presence of several statistically significant effect modifiers…throws further into question the validity of computing national risk estimates."[17] EPA not only reports a national risk estimate in Chapter 8 of the HREA (*i.e.*, 16,000 premature deaths), but provides a confidence range suggesting that it is statistically significant (*i.e.*, 7,200 to 22,000).[18]

The Majority of EPA's Risk Estimates – Even for "As-Is" Ozone Levels – Are Irreducible

Despite the HREA's overstatements regarding the quantity and certainty of its mortality risk estimates (in light of the concerns noted above), the percentage reductions that EPA projects for tightening the current standard are modest. For example, the estimated percentage reductions in mortality risks are in the ballpark of 10% to 15% even if the current standard were to be tightened to 60 ppb. That small percentage risk reduction occurs despite estimated reductions of *all* U.S. manmade NOx emissions ranging from 60% to over 90%.[19] The percentage reductions in risk are so small despite such enormous reductions in U.S. emissions because ozone levels are so close to the level that is predicted if all U.S. manmade emissions (or even all North American emissions) were to be eliminated. A majority of each of EPA's many

---

[14] Smith *et al*. at p. 58.

[15] Smith et al., 2009, p. 54.

[16] Smith *et al*. at p. 54.

[17] Smith *et al*. at p. 58.

[18] HREA, Table 8-1, p. 8-7. (As noted above, the estimate of 16,000 should actually read 15,000.)

[19] These are reductions from emissions during the periods 2006-2008 and 2008-2010 and come from the HDDM analysis for the 60 ppb standard. That analysis could not simulate attainment with 60 ppb for New York City. For the specific percentage emissions reductions, see Table 2 of Appendix 4 of the HREA, pp. 26-27.

risk estimates – even for as-is conditions – are irreducible; they would remain even if every ton of U.S. manmade emissions were eliminated. The HREA has not provided any estimates of risk remaining in its US zero-out case (which is described in Chapter 2 of the PAD), but it should do so, in order that HREA readers can more readily realize this fact without studying its many appendices.