Comments on the California Air Resources Board's Draft Report "Methodology for Estimating Premature Deaths Associated with Long-Term Exposures to Fine Airborne Particulate Matter in California"

Joel Schwartz
Visiting Fellow
American Enterprise Institute[1]

July 11, 2008

This paper provides comments on the California Air Resources Board's (CARB) draft report "Methodology for Estimating Premature Deaths Associated with Long-term Exposures to Fine Airborne Particulate Matter in California,"[2] which CARB released for public comment on May 22, 2008.

This update of CARB's methodology for assessing premature mortality due to exposure to fine particulate matter ($PM_{2.5}$) provides an opportunity for CARB to assess the weight of the evidence on the health effects of today's historically low air pollution levels. Unfortunately, rather than provide such an assessment, CARB has selected and structured information in ways that exaggerate harm from air pollution. CARB accepts uncritically the results of studies claiming to find a causal link between air pollution and mortality. On the other hand, CARB stretches for reasons to discount studies that fail to find harm from PM, often misrepresenting these studies in the process. CARB's selective marshalling of evidence creates a false appearance that harm from $PM_{2.5}$ is greater and more certain than is warranted by the actual weight of the underlying evidence from the scientific literature.

CARB's advisory and peer review process only exaggerates the shortcomings in CARB's substantive review of air pollution health science. Despite the wide range of scientific opinion on the validity of observational epidemiology studies and air pollution epidemiology in particular, CARB chose as peer reviewers and scientific advisors epidemiologists who believe strongly in the validity of the methods and results of air pollution epidemiology studies, who are supportive of CARB's regulatory goals, and who have published much of the research CARB and EPA rely on to justify the expansion of their regulatory powers. These selection biases and conflicts of interest ensured that CARB's PM mortality analysis did not receive a genuine critical review by independent experts.

In the remainder of these comments I provide evidence that $PM_{2.5}$ at current and recent levels is not a cause of premature mortality, and show how CARB exaggerated the case for harm from $PM_{2.5}$ and shielded itself from independent review.

---

[1] Email address: jschwartz@aei.org.

[2] Hereafter cited as "CARB Methodology."

The Big Picture: Observational Epidemiology Studies Give False Indications of Risk Where No Risk In Fact Exists

CARB's claim that air pollution at current U.S. levels is killing people rests almost solely on results of observational studies—that is, studies with non-randomly selected groups of people and non-randomly assigned exposures. A number of researchers have provided evidence that observational studies are simply not capable of providing reliable information on the existence of small risks, such as those claimed for air pollution.

The implicit assumption in an observational study is that after researchers have controlled for all known non-pollution factors that might be correlated with pollution levels and health outcomes (e.g., weather, smoking, diet, etc.) any remaining correlation between air pollution and health represents a genuine causal linkage between the two. A wide range of evidence shows that this assumption is false and that observational studies tend to "find" effects where no real effects exist.[3]

Indeed, many prominent epidemiologists are wringing their hands over the widespread problem and embarrassment of spurious health claims from observational epidemiology studies and are questioning whether observational studies are even capable of providing valid evidence on health risks.[4]

Unfortunately, this acknowledgement of the limits of observational studies in the wider community of epidemiologists has had little effect on the relatively insular world of air pollution epidemiologists and the regulators who fund them. Even so, there have been some critiques from within air pollution epidemiology. Here, for example, is one caution on the validity of observational studies of air pollution's health effects:

> estimation of very weak associations in the presence of measurement error and strong confounding is inherently challenging. In this situation, prudent epidemiologists should recognize that residual bias can dominate their results. Because the possible mechanisms of action and their latencies are uncertain, the biologically correct models are unknown. *This model selection problem is exacerbated by the common practice of screening*

---

[3] S. Begley, "New Journals Bet 'Negative Results' Save Time, Money," *Wall Street Journal*, September 15, 2006, B1, http://online.wsj.com/article_print/SB115827169620563571.html; J. P. Ioannidis, "Why Most Published Research Findings Are False," *PLoS Med* 2 (2005): e124; J. P. Ioannidis, "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research," *Journal of the American Medical Association* 294 (2005): 218-28; G. Taubes, "Epidemiology Faces Its Limits," *Science* 269 (1995): 164-69; G. Taubes, "Do We Really Know What Makes Us Healthy?" *New York Times*, September 16, 2007, http://www.nytimes.com/2007/09/16/magazine/16epidemiology-t.html?_r=3&ref=magazine&oref=slogin&oref=slogin&oref=slogin.

[4] S. Ebrahim and M. Clarke, "Strobe: New Standards for Reporting Observational Epidemiology, a Chance to Improve," *International Journal of Epidemiology* 36 (2007): 946-48; S. J. Pocock, T. J. Collier, K. J. Dandreo et al., "Issues in the Reporting of Epidemiological Studies: A Survey of Recent Practice," *BMJ* 329 (2004): 883; G. D. Smith and S. Ebrahim, "Epidemiology - Is It Time to Call It a Day?" *International Journal of Epidemiology* 30 (2001): 1-11; E. Von Elm and M. Egger, "The Scandal of Poor Epidemiological Research," *British Medical Journal* 329 (2004): 868-69.

> *multiple analyses and then selectively reporting only a few important results.*[5] (emphasis added)

The highlighted portion is key. Researchers make many subjective choices in developing statistical models relating air pollution to health. Furthermore, the studies are undertaken by regulatory agencies and air pollution health researchers with the explicit goal of finding harm from air pollution. In this environment, researchers tend to choose statistical models that maximize the effect they "expect" or "hope" to find—a problem known as data-mining. As a result, observational studies become statistical fishing expeditions that turn up chance correlations rather than real effects.

An additional bias is that researchers are more likely to seek publication of, and journal editors are more likely to accept for publication studies that find an effect, while studies that don't find any effects end up packed away into filing cabinets. The result is a problem known as "publication bias." The overall result is that the scientific literature includes lots of studies reporting "effects" that aren't real. Once again, even some air pollution epidemiologists have noted the problem:

> Publication bias arises because there are more rewards for publishing positive or at least statistically significant findings. It is a common if not universal problem in our research culture. In the case of time-series studies using routine data there are particular reasons why publication bias might occur. One is that the data are relatively cheap to obtain and analyse, so that there may be less determination to publish "uninteresting" findings. The other is that each study can generate a large number of results for various outcomes, pollutants and lags and there is quite possibly bias in the process of choosing amongst them for inclusion in a paper. In the field of air pollution epidemiology, the question of publication bias has only recently begun to be formally addressed.[6]

In many areas of health research, randomized trials—a gold standard methodology that reduces or eliminates the biases inherent in observational studies—can be conducted to test claims made based on observational studies. In such cases, observational studies are routinely contradicted when checked against randomized trials, confirming concerns about data-mining and publication bias.[7]

In the case of air pollution, ethical and practical concerns make it impossible to do a randomized trial to test whether today's historically low air pollution levels are deadly. But if observational studies are invalid in all other areas of health research, there's no reason to expect them to do any better on air pollution. In fact, we should expect observational air pollution studies to be even less likely to be valid, because the putative

---

[5] T. Lumley and L. Sheppard, "Time Series Analyses of Air Pollution and Health: Straining at Gnats and Swallowing Camels?" *Epidemiology* 14 (2003): 13-4.

[6] H. Anderson, R. Atkinson, J. Peacock et al., *Meta-Analysis of Time-Series Studies and Panel Studies of Particulate Matter (Pm) and Ozone* (World Health Organization, 2004), www.euro.who.int/document/e82792.pdf.

[7] Begley, "New Journals Bet 'Negative Results' Save Time, Money."; Ioannidis, "Why Most Published Research Findings Are False."; Taubes, "Do We Really Know What Makes Us Healthy?."

effects they claim to be uncovering are much smaller than in observational studies of medical interventions. As a consequence, the results of observational air pollution studies are at even greater risk of being statistical figments rather than real effects.

Readers of CARB's $PM_{2.5}$ methodology report would have no inkling that the report's conclusions are based on a discredited research methodology. Instead, CARB creates a false impression that observational studies are finding real cause-effect linkages.

Direct Evidence of Spurious Results from Air Pollution Cohort Studies

Although observational air pollution studies in humans can't be checked against randomized trials, we do have some direct evidence that observational studies are producing spurious indications of harm from $PM_{2.5}$. Unfortunately, CARB omits this evidence.

CARB considers the American Cancer Society (ACS) study (also known as the Pope Study) and the Harvard Six Cities (HSC) study to provide strong evidence that any amount of particulate matter in the air is deadly.[8] Both studies assessed the association between long-term $PM_{2.5}$ exposure and risk of death in different cities around the U.S. CARB states "the primary evidence for $PM_{2.5}$ mortality C-R [concentration-response] functions comes from multiple analyses from the Harvard Six Cities study…and the ACS cohort study.[9] In fact, based on the most recent reports from these two studies, CARB has *increased* the estimated risk from $PM_{2.5}$.[10]

However, reanalyses of the ACS and HSC data have demonstrated the extent to which observational studies can give spurious results when researchers leave out important confounding variables. For example, in a reanalysis by the Health Effects Institute (HEI), when migration rates into and out of various cities over time were added to the statistical model relating $PM_{2.5}$ and risk of death, the apparent effect of $PM_{2.5}$ dropped by two-thirds and became statistically insignificant.[11] Migration was just one of several confounding factors that diminished or erased the apparent harm from $PM_{2.5}$, but that were not accounted for by the original researchers.

Regulators and air pollution epidemiologists (including the HEI researchers who did the reanalysis as well as CARB's and EPA's scientific advisors) have ignored this refutation

---

[8] C. A. Pope, 3rd, M. J. Thun, M. M. Namboodiri et al., "Particulate Air Pollution as a Predictor of Mortality in a Prospective Study of U.S. Adults," *American Journal of Respiratory and Critical Care Medicine* 151 (1995): 669-74; C. A. Pope, 3rd, R. T. Burnett, M. J. Thun et al., "Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution," *Journal of the American Medical Association* 287 (2002): 1132-41.

[9] CARB Methodology, p. 22.

[10] M. Jerrett, R. T. Burnett, R. Ma et al., "Spatial Analysis of Air Pollution and Mortality in Los Angeles," *Epidemiology* 16 (2005): 727-36; F. Laden, J. Schwartz, F. E. Speizer et al., "Reduction in Fine Particulate Air Pollution and Mortality: Extended Follow-up of the Harvard Six Cities Study," *American Journal of Respiratory and Critical Care Medicine* 173 (2006): 667-72.

[11] D. Krewski, R. T. Burnett, M. S. Goldberg et al., *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality* (Cambridge, MA: Health Effects Institute, July, 2000).

of the ACS/Pope results and continue to claim the ACS/Pope study provides proof of harm from air pollution. When discussing the HEI reanalysis of the ACS/Pope study, CARB claims "the adjusted results did not differ substantively from the original findings. The reanalysis demonstrated the robustness of the PM-mortality risk estimates to many alternative model specifications." This claim is simply false, as several HEI sensitivity analyses showed that the original results suffered from confounding. Once the confounding was corrected, the $PM_{2.5}$ association went away.

CARB ignores other inconvenient results from the HEI reanalysis. For example, in addition to a national average association between $PM_{2.5}$ and mortality, HEI looked at regional variations and reported that $PM_{2.5}$ was not associated with any increase in mortality in California.[12]

The Harvard Six Cities results also turned out to be sensitive to changes in the statistical model. For example, the HSC did not account for differences in physical activity levels among the cities in the study. It later turned out that physical activity and $PM_{2.5}$ levels were inversely correlated, so the ostensible effect of $PM_{2.5}$ could instead have been due to confounding.[13]

The Big Picture: No Harm from $PM_{2.5}$ in Animal Studies

Animal studies provide a further check on the validity of observational epidemiology studies. If air pollution at today's low ambient levels is deadly to people, then we would expect that much higher levels of air pollution would kill at least some laboratory animals. However, researchers have been unable to kill various species of animals even with air pollution at levels many times greater than are ever found in ambient air. A recent review of particulate matter toxicology concluded, "It remains the case that no form of ambient PM—other than viruses, bacteria, and biochemical antigens—has been shown, experimentally or clinically, to cause disease or death at concentrations remotely close to US ambient levels."[14] If high levels of $PM_{2.5}$ can't kill several different species of animals, it's unlikely that low levels of $PM_{2.5}$ are killing people.

CARB implies that Sun et al. (2005) provides direct toxicological evidence of harm from $PM_{2.5}$ at real-world ambient levels.[15] Sun et al. claimed to have uncovered a direct cause-and-effect relationship between current $PM_{2.5}$ levels and heart disease, especially along with a high-fat diet, based on a study of mice. Both researchers and the media hailed this

---

[12] See Figure 21, p. 197 of the HEI report. Note that relative risks were below 1.0 (i.e., no increase in mortality due to $PM_{2.5}$) in all of California. Ibid.

[13] F. W. Lipfert, "Estimating Air Pollution-Mortality Risks from Cross-Sectional Studies: Prospective Vs. Ecologic Study Designs," Health and Regulatory Issues, Proceedings of the International Specialty Conference, Air and Waste Management Association, 1995.

[14] L. C. Green and S. R. Armstrong, "Particulate Matter in Ambient Air and Mortality: Toxicologic Perspectives," *Regulatory Toxicology and Pharmacology* 38 (2003): 326-35.

[15] CARB Methodology, p. 17. Q. Sun, A. Wang, X. Jin et al., "Long-Term Air Pollution Exposure and Acceleration of Atherosclerosis and Vascular Inflammation in an Animal Model," *Journal of the American Medical Association* 294 (2005): 3003-10.

study as providing proof, previously lacking in animal studies, that air pollution is causing heart disease, and therefore premature death, in humans.

In fact, the Sun et al.'s study had nothing to do even with real mice, much less with people. Sun et al. used mice genetically engineered to be lacking a blood lipid/cholesterol regulation system. These mice had 14 times the cholesterol levels of normal mice. For comparison, only about one in 500 American males has cholesterol of even twice the national average and virtually no human has cholesterol greater than four times the average. The very reason for using such unrealistic mice is that even massive $PM_{2.5}$ doses don't cause heart disease in normal mice. In addition, although the researchers claimed their $PM_{2.5}$ doses were similar to real world doses, the acute doses were in fact substantially higher than even the highest real world exposures.[16]

CARB Mischaracterizes Cohort Studies that Do Not Find Harm from $PM_{2.5}$

Two cohort studies did not find harm from $PM_{2.5.}$ Rather than contend with this evidence against harm from $PM_{2.5}$, CARB instead mischaracterizes the studies' methods and results, creating a false impression that the studies are irrelevant or invalid.

The Veterans study assessed the association between $PM_{2.5}$ and mortality risk from 1976-2001 in a cohort of 70,000 male U.S. veterans with high blood pressure.[17] The study reported that higher $PM_{2.5}$ was associated with a statistically significant *decrease* in risk of death.

CARB claims "Overall, in the VA analyses, effect estimates to various measures of PM were unstable and not robust to model selection, time windows used, or various other analytic decisions."[18] Even if this were true, the criticism applies equally to the ACS/Pope and Six Cities cohorts. As already noted, in the ACS/Pope cohort the ostensible effect of $PM_{2.5}$ disappeared when additional confounding factors were considered, including migration, sulfur dioxide, and several others. The ACS/Pope results also feature several biologically implausible results. For example, $PM_{2.5}$ appeared to kill men, but not women; those who said they were moderately active, but not those who said they were very active or sedentary.[19] These biologically implausible patterns suggest the correlation of $PM_{2.5}$ and mortality was a statistical figment rather than a real causal effect.

---

[16] For a more detailed demonstration of why Sun et al. is irrelevant for assessing health risks in mice or people see J. Schwartz, *Air Pollution and Health: Do Popular Portrayals Reflect the Scientific Evidence?* (Washington, D.C.: American Enterprise Institute, May 2006), http://www.joelschwartz.com/pdfs/AirPoll_Health_EPO_0506.pdf.

[17] F. W. Lipfert, J. D. Baty, J. P. Miller et al., "PM2.5 Constituents and Related Air Quality Variables as Predictors of Survival in a Cohort of U.S. Military Veterans," *Inhalation Toxicology* 18 (2006): 645-57; F. W. Lipfert, H. M. Perry, J. P. Miller et al., "The Washington University-EPRI Veterans' Cohort Mortality Study," *Inhalation Toxicology* 12 (suppl. 4) (2000): 41-73; F. W. Lipfert, R. E. Wyzga, J. D. Baty et al., "Traffic Density as a Surrogate Measure of Environmental Exposures in Studies of Air Pollution Health Effects: Long-Term Mortality in a Cohort of Us Veterans," *Atmospheric Environment* 40 (2006): 154-69.

[18] CARB Methodology, p. 8.

[19] Pope, Burnett, Thun et al., "Lung Cancer, Cardiopulmonary Mortality, and Long-Term Exposure to Fine Particulate Air Pollution."

The most recent report on the ACS/Pope cohort also reports results that are "not robust to model selection." CARB highlights Jerrett et al. (2005) because it ostensibly found greater risks from $PM_{2.5}$ than were reported in the original ACS/Pope studies.[20] CARB ignores the fact that the relative risk from $PM_{2.5}$ became statistically insignificant in the models that had the most extensive controls for confounding.

In any case, the Veterans results are not in fact unstable. The study has consistently found that higher $PM_{2.5}$ is associated with either no mortality or lower mortality. The Veterans study did however find that whatever the effects of $PM_{2.5}$, they are decreasing with time. Perhaps CARB mistook this decline for "instability." Ironically, the ACS/Pope study also suggests that $PM_{2.5}$ effects are decreasing with time (though Pope et al. never say so explicitly).

For the 1982-89 follow-up period, Pope et al. (1995) reported a nationwide average relative risk (RR) of 1.069 per 10 $\mu g/m^3$ $PM_{2.5}$. However, for the 1982-98 follow-up period Pope et al. (2002) reported an RR of 1.04. Although the authors never say so explicitly, this means that the RR declined between the 1982-89 and 1990-98 follow-up periods. Based on data provided in the two papers, one can calculate that the RR for 1990-98 was about 1.019, and is statistically insignificant. Thus, even on its own terms, the ACS/Pope study suggests that any harm from $PM_{2.5}$ that might have existed 20 or 30 years ago has now disappeared.

CARB also discounts the Veterans study based on the claim that the cohort is not representative of Californians. CARB states "As our objective is to derive a relative risk applicable to the general population of California, it is important to use studies that have a similar at-risk population. This criterion would eliminate direct application of studies like the Washington University-EPRI Veterans Cohort…which focused on male military veterans under treatment for hypertension, with 81 percent current or former smokers."[21]

CARB's objection is particularly ironic because the Veterans cohort has exactly the characteristics CARB would normally look for in an air pollution health study. It has the largest percentage of minorities of any cohort in an air pollution mortality study (35 percent African American). The high minority component dovetails with CARB's goal of ensuring that air pollution doesn't disproportionately harm minorities.

The men in the cohort also had high blood pressure, which should have made them *more* susceptible to any harm from air pollution, when compared with the general population. CARB's goal is to set standards that protect even the most "sensitive" groups, and the Veterans cohort *is* a sensitive group. Instead, CARB focuses on the mainly white, middle class ACS cohort and on the Harvard Six Cities cohort, which did not even include people in California.

CARB also gives short shrift to Enstrom (2005), which reported on the association of $PM_{2.5}$ and mortality in cohort of 36,000 elderly Californians from 1973-2002.[22] The study

---

[20] Jerrett, Burnett, Ma et al., "Spatial Analysis of Air Pollution and Mortality in Los Angeles."

[21] CARB Methodology, p. 21.

[22] J. E. Enstrom, "Fine Particulate Air Pollution and Total Mortality among Elderly Californians, 1973-2002," *Inhalation Toxicology* 17 (2005): 803-16.

found that $PM_{2.5}$ was not associated with any increase in mortality risk after the early 1980s. In discounting the Enstrom results, CARB states "the Enstrom (2005) study of elderly Californians neither adequately controlled for smoking nor adjusted for exposure to environmental tobacco smoke, two factors that could significantly alter the effect of PM exposures on premature death. Further, exposure misclassification is another issue of concern. In Entrom's [sic] study, $PM_{2.5}$ was assigned on the basis of data from just a few monitoring sites and at times on very few measurements (Brunekreef 2006). No discussion was provided as to the representativeness of sites."[23]

CARB cites a letter to the editor by Brunekreef and Hoek (2006) to support its discounting of Enstrom.[24] But Brunekreef and Hoek's claims are either mistaken or apply equally well to studies that claim to find harm from $PM_{2.5}$. For example, CARB says of Enstrom (2005) "$PM_{2.5}$ was assigned on the basis of data from just a few monitoring sites and at times on very few measurements." In fact, Enstrom used data from the Inhalable Particulate Network (IPN), a special $PM_{2.5}$ monitoring network EPA set up during 1979-83. These are the only data on $PM_{2.5}$ available from that time. Furthermore, *the Pope/ACS study used this exact same IPN data*. Thus, if Enstrom's study is invalid because of problems with the $PM_{2.5}$ data then the ACS/Pope study is likewise invalid. Indeed, one could level the same critique at the Harvard Six Cities study, which set up $PM_{2.5}$ monitors especially for the study, but only one monitor per city.

CARB is also incorrect in claiming that Enstrom did not adequately control for smoking. Enstrom included controls for smoking status both at study entry in 1959 and in 1972, just before the follow-up period began. Enstrom did not adjust for exposure to environmental tobacco smoke, but as he points out "No control for environmental tobacco smoke (ETS) was necessary because a separate study showed that ETS was not related to mortality among the never smokers in this cohort."[25]

CARB also asserts of Enstrom (2005) "Yet another issue is the long time passed since enrollment (1959) and follow-up (1973- 2002), which must have been associated with many changes in diet, smoking, occupation, etc., factors for which the authors could not adequately control." This claim is misleading. Smoking status was ascertained not only at entry to the study in 1959, but also in 1972 at the beginning of the follow-up period. Smoking is the single largest factor affecting health and CARB is simply mistaken in claiming that Enstrom did not control for it.

CARB's criticism also applies equally well to the cohort studies that CARB lauds. In the Pope/ACS study, the controls for smoking, diet, etc. were based on data collected at entry to the study in 1982. Thus, this study also fails to capture any changes in status or behavior that occurred after entry to the study.

---

[23] CARB Methodology, p. 22.

[24] B. Brunekreef and G. Hoek, "A Critique Of "Fine Particulate Air Pollution and Total Mortality among Elderly Californians, 1973-2002" By James E. Enstrom," *Inhalation Toxicology* 18 (2006): 507-8.

[25] J. E. Enstrom, "Response To "A Critique of 'Fine Particulate Air Pollution and Total Mortality among Elderly Californians, 1973-2002" by Bert Brunekreef, PhD, and Gerard Hoek, PhD," *Inhalation Toxicology* 18 (2006): 509-14.

CARB's critique of Enstrom is also an extreme case of selective citation. Although CARB cites Brunekreef and Hoek's (2006) critique of Enstrom (2005), CARB omits Enstrom's response, which refutes Brunekreef and Hoek's claims.[26]

Overall, CARB stretches for ways to discount the results of studies that fail to find harm from $PM_{2.5}$, while ignoring the shortcomings and inconsistencies of studies that do claim to find harm from $PM_{2.5}$. In the final version of the report, CARB must do a far better effort of providing realistic and honest reviews of the evidence, regardless of whether than evidence is congenial to CARB's bureaucratic interests.

EPA's Expert Elicitation Should Not Be Taken Seriously

EPA recently completed an "Expert Elicitation" on particulate matter health effects.[27] The Expert Elicitation included 12 scientists, most with expertise in air pollution epidemiology and others with expertise in toxicology and medicine. The experts reviewed a wide range of studies on air pollution and health and sat for several hours of interviews during which they gave their expert opinions on the health effects of particulate matter air pollution, and in particular their opinion on the exposure-response relationship between PM levels and risk of premature mortality. CARB places great weight on Expert Elicitation's results, concluding "In summary, it is appropriate to rely on the U.S. EPA experts' judgments for California's specific risk assessments."[28]

Both EPA and CARB create the impression that the scientists EPA chose for the expert elicitation provided an independent and unbiased evaluation of PM health effects. But the Expert Elicitation was in fact vitiated by selection biases and conflicts of interest. CARB states that the "Experts relied upon a core set of cohort epidemiology studies to derive their quantitative estimates, mainly those associated with the ACS [American Cancer Society] and [Harvard] Six Cities cohorts."[29] But of the 12 experts, six are co-authors of these studies, meaning they were giving their expert opinion on their own research.[30] One of the experts is the chief air pollution epidemiologist at the California Environmental Protection Agency. Most, perhaps all of the researchers are heavily funded by EPA and/or CARB to do the research that EPA and CARB then use to justify expansion of their regulatory authority. Among epidemiologists skeptical of a link between low-level and pollution and mortality, none were included in the Expert Elicitation.

---

[26] Ibid.

[27] Industrial Economics, *Expanded Expert Judgment Assessment of the Concentration-Response Relationship between $PM_{2.5}$ Exposure and Mortality, prepared for the U.S. Environmental Protection Agency* (Washingon, D.C.: September 21, 2006); H. A. Roman, K. D. Walker, T. L. Walsh et al., "Expert Judgment Assessment of the Mortality Impact of Changes in Ambient Fine Particulate Matter in the U.S," *Environmental Science and Technology* 42 (2008): 2268-74.

[28] CARB Methodology, p. 23.

[29] CARB Methodology, p. 17.

[30] For example, of the 12 experts, C. Arden Pope, Daniel Krewski, Kazuhiko Ito, and George Thurston authored papers on the ACS study. Joel Schwartz, Douglas Dockery, and Pope authored papers on the Harvard Six Cities study.

Clearly, the Expert Elicitation did not come close to providing independent and unbiased analysis. Rather, the study design inherently ensured that it would confirm EPA's preconceptions and regulatory goals, rather than uncover realistic information on air pollution health effects. The Expert Elicitation is not an appropriate guide upon which to make scientific judgments or inform regulatory policy and CARB should remove the Expert Elicitation's results from the Methodology report.

CARB Should Commission Genuinely Independent Reviews of Its Analysis

CARB's own advisory and peer review process suffers from a similar lack of independence and conflict of interest. To ensure that CARB's report receives a genuine critical evaluation before its release, CARB should include among its advisors and peer reviewers (1) epidemiologists who have provided evidence against the existence of a causal association between air pollution and mortality, (2) epidemiologists who have evaluated the validity of observational methods for assessing the existence of small risks, and (3) researchers who are not air pollution epidemiologists, but who are expert in the mathematical techniques used in air pollution epidemiology, and of assessing the real-world validity of causal inferences based on those techniques—for example, econometricians, statisticians, and researchers in other sub-fields of risk factor epidemiology.

EPA's Expert Elicitation and CARB's advisory and review process are cases of the emperor asking his tailors to judge the quality of his clothes. It is time for the emperors of air pollution regulation to expand their circle of advisors.

Additional Errors and Mischaracterizations

Dutch cohort study. According to CARB, even studies that find no harm from PM, nevertheless lend support to CARB's claim of a PM-mortality link. Here is CARB's description of results from a Dutch study of PM and mortality:

> A more recent study on the same [Dutch] Cohort, Beelen et al. (2008), reinforces the conclusions of the pilot study. The authors found a positive association between traffic intensity on the nearest roadway to the subject's residence and death rate. They also confirmed the link between interpolated BS [black smoke] concentrations and cardiopulmonary mortality. *While the associations between pollutants and mortality in this study were not statistically significant, the authors' methodology was very careful, and their results lend convincing support to the link between premature death and PM.* [31] [emphasis added]

In other words, Beelen et al. did not find a statistically significant association between particulate matter and mortality and CARB believes Beelen et al. used a "very careful" methodology. Yet CARB still claims that this study "lend[s] convincing support to the link between premature death and PM."

---

[31] CARB Methodology, p. 8.

CARB should adjust its Methodology report to reflect the lack of support for $PM_{2.5}$ effects in this study.

Dublin coal-ban study. The city of Dublin, Ireland in 1990 banned to use of soft (bituminous) coal for home heating and cooking, which resulted in a large drop in black smoke levels, particularly in winter. A study in the *Lancet* concluded that the coal ban caused a reduction in premature mortality.[32] CARB singles out this report as an intervention study that provides evidence that declines in $PM_{2.5}$ cause declines in mortality.

The Dublin study clams to demonstrate that premature mortality decreased due to PM reductions that resulted from Dublin's ban on the burning of bituminous coal on September 1, 1990. However, there was a large spike in mortality in winter 1990, just before the ban went into effect. This could have created the appearance that the drop in mortality after the coal ban was due to the reduction in black smoke due to the coal ban, even if the coal ban had nothing to do with it.

The authors controlled for flu outbreaks in their model using an indicator variable, and one of the five flu outbreaks during the study period did occur during winter 1990. But it is clear from the data that the mortality spike is not mainly a result of the flu outbreak. First, the winter 1990 mortality spike occurred for all causes of death, whereas only cardiovascular and respiratory deaths were anomalously high during other flu outbreaks. Second, the winter 1990 mortality anomaly was much greater than during other flu outbreaks.[33]

Even if the Dublin results are taken to have found a causal relationship between lower air pollution and fewer deaths, it's not clear that it has any lessons for air pollution in California. First, the study was based on black smoke levels in Dublin, which went from winter-average levels of 85 $\mu g/m^3$ before the coal ban down to 22 $\mu g/m^3$ after. Annual-average levels went from 50 $\mu g/m^3$ down to 15 $\mu g/m^3$. These are for black smoke alone, rather than total $PM_{2.5}$, so total $PM_{2.5}$ levels would have been even higher. The study is thus based on much higher average $PM_{2.5}$ levels than the levels of the federal or California $PM_{2.5}$ standards.

In addition, the study used outdoor black smoke levels as the exposure variable. But the coal was being used for home space and water heating. Indoor PM exposures would therefore have been much higher than even the already-large outdoor exposures, further increasing the exposure levels when compared with current U.S. standards, especially given that people spend most of their time indoors, especially during winter when indoor $PM_{2.5}$ levels would have been highest.

---

[32] L. Clancy, P. Goodman, H. Sinclair et al., "Effect of Air-Pollution Control on Death Rates in Dublin, Ireland: An Intervention Study," *Lancet* 360 (2002): 1210-4.

[33] It is also worth noting that the authors didn't actually have any data on flu outbreaks in Ireland. Instead, they assumed that a flu outbreak was occurring in any 14-day moving window in which the national mortality rate due to influenza or pneumonia was above the 95^{th} percentile.

Overall, the Dublin study isn't all CARB cracks it up to be and in any case is irrelevant for PM levels and routes of exposure in California. CARB's Methodology report should be adjusted to reflect this.

CARB Must Go Back to the Drawing Board

Before finalizing the Methodology report, CARB must consider the full weight and strength of the evidence, including evidence against causal associations of air pollution and mortality, weaknesses in the studies that purport to demonstrate a causal connection, and evidence on the fundamental validity of the methods used to make causal claims. Furthermore, CARB must take these steps within a framework that includes genuinely independent scientists both from within and outside air pollution epidemiology.

In order to ensure that CARB's estimate of $PM_{2.5}$ health effects reflects the real-world validity of $PM_{2.5}$ studies, and the real-world likelihood of harm from current, historically low levels of $PM_{2.5}$, I offer the following recommendations:

1. CARB should step back and assess whether observational epidemiology studies are capable of providing reliable information on the existence of small risks. Observational studies are the main justification for the claim of a causal association between air pollution and premature death, but they are also the weakest form evidence.

   As shown in these comments, there is good reason to discount the results observational studies, due to the inherent weaknesses and biases in the methods themselves, and due to the clear influence of data mining and publication bias. These factors work to inflate apparent harm from air pollution. In addition, experimental studies with both humans and animals don't support a causal air pollution-mortality association, contradicting the observational studies.

   There are thousands of observational studies claiming to provide support for a causal association between low-level air pollution and risk of death. But implementing invalid techniques over and over again doesn't improve their validity.

2. CARB should not to omit or mischaracterize contrary evidence, and should take a more critical look at studies claiming to support a causal association between air pollution and mortality. My comments provide a few examples of such omissions and mischaracterizations, but there are many more.

3. EPA's Expert Elicitation suffers from serious biases and conflicts of interest that render it's results invalid. CARB should not base its conclusions about the health effects of $PM_{2.5}$ on the Expert Elicitation and should not give the Expert Elicitation a prominent role in its Methodology report.

4. CARB's analysis suffers from biases and conflicts of interest similar to those of EPA's Expert Elicitation. To ensure that CARB's report receives a genuine critical evaluation before its release, CARB should include among its advisors and peer reviewers (1) epidemiologists who have provided evidence against the existence of a

causal association between air pollution and mortality,[34] (2) epidemiologists who have evaluated the validity of observational methods for assessing the existence of small risks,[35] and (3) researchers who are not air pollution epidemiologists, but who are expert in the mathematical techniques used in air pollution epidemiology, and of assessing the real-world validity of causal inferences based on those techniques—for example, econometricians, statisticians, and researchers in other sub-fields of risk factor epidemiology.[36] CARB needs a broad range of views and expertise to ensure that its results reflect the weight of the evidence rather than merely CARB's bureaucratic interests.

5. Putting aside the fundamental concerns about whether estimates based on observational studies represent real risks, it is important to find out why different researchers come up with such different results for $PM_{2.5}$ effects.

   To find out what is causing all of these different results, CARB should commission reanalyses to confirm that the original results can be replicated and to determine how robust and reliable the various results are to different specifications and approaches.

   At least two separate researchers should perform these reanalyses; at least one "skeptic" and at least one "believer." Having researchers with different points of view will ensure vigorous testing and review of the validity of any given approach to analyzing the data.

   Outside statisticians, econometricians and epidemiologists should also be part of the reanalyses themselves, as well as up-front reviewers of the reanalysis protocols as well as peer-reviewers of the results.

The Methodology report's errors and biases are too extensive and profound for the report to merely be tweaked and released in a few weeks, as the current schedule requires. Instead, CARB must go back to the drawing board by appointing genuinely independent scientific advisors and peer reviewers, commissioning genuinely independent reanalyses of key data, and rewriting the report from scratch.

---

[34] For example, Fred Lipfert, James Enstrom, Suresh Moolgavkar, Lise Tole, William Keatinge, and Richard L. Smith to name just a few.

[35] For example, George Davey Smith, John Ioannidis, or members of the STROBE team.

[36] For example, Michael Greenstone, David Freedman, Paul Switzer, Anne Smith, as well as researchers mentioned in the previous footnote.