

CARB

truckbus14: diesel rules,

S. Stanley Young
National Institute of Statistical Sciences

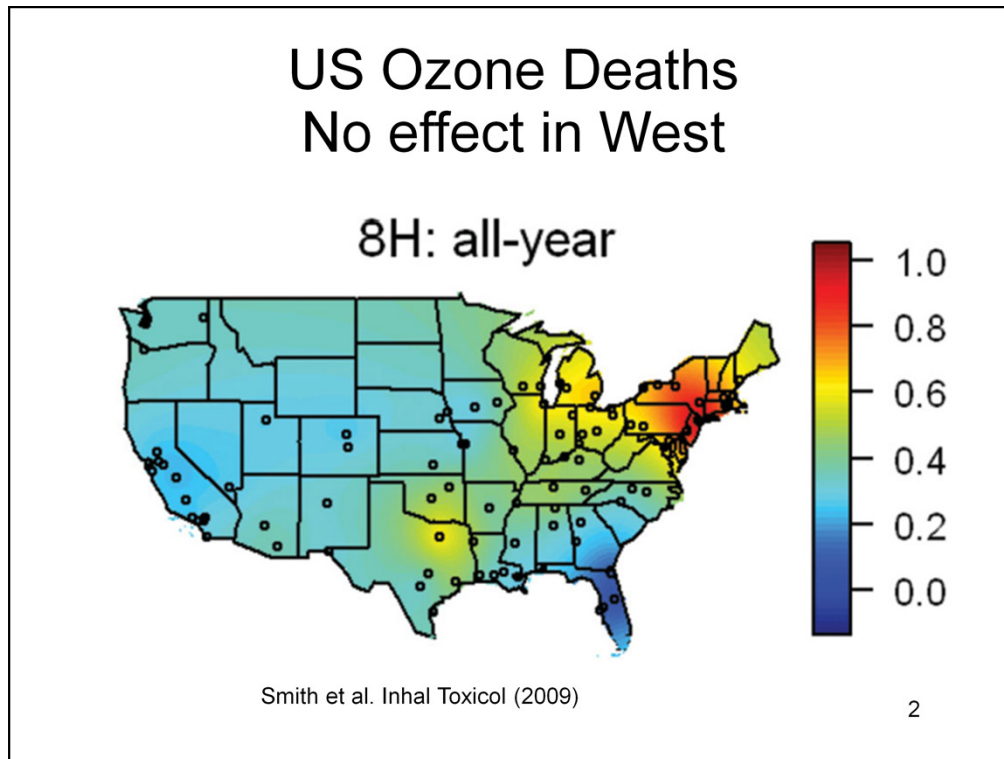
April 24, 2014

1

Policy that is based on science needs to satisfy basic rules of science. For example, data sets used in papers cited should be publicly available. It should be noted that “peer review” does not mean that the claims made in the papers are true and reproducible. Peer review only means that the normal standards in that scientific area were followed.

Any CARB/EPA report on a question should cite research both for and against the question at issue.

As a citizen I think one main purpose of Board is to ensure scientific oversight to the CARB on their use of science.



The risk of mortality from ozone varies by location. Geographic heterogeneity of effects of air pollution were known as early as 2000, Krewski. Heterogeneity was observed by Smith et al. (2009). It has been repeatedly observed that there is no effect of ozone or PM2.5 on deaths in California. See Enstrom (2005), Young and Xia, Statistical Analysis and Data Mining (2013). Experts have known about geographic heterogeneity since 2000, but the CARB/EPA has not taken about geographic heterogeneity into account in rule making.

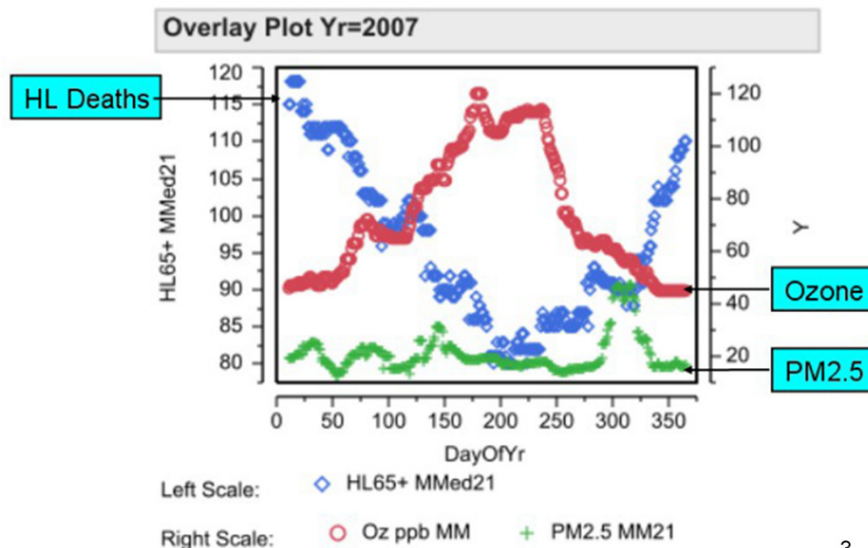
D. Krewski, R. T. Burnett, M. S. Goldberg, K. Hoover, J. Siemiatycki, M. Jerrett, M. Abrahamowicz, and W. H. White, Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality. Part II: Sensitivity Analysis, HEI Publications. <http://pubs.healtheffects.org/view.php?id=6>, See Figure 21, in particular, **2000**.

J. E. Enstrom, Fine particulate air pollution and total mortality among elderly Californians, 1973–2002, Inhalation Toxicology 17 (**2005**), 803–816.

R. L. Smith, B. Xu, and P. Paul Switzer, Reassessing the relationship between ozone and short-term mortality in U.S. urban communities, Inhal Toxicol 29(S2) (**2009**), 37–61.

Young SS, Xia JQ. Assessing geographic heterogeneity and variable importance in an air pollution data set. *Statistical analysis and data mining*. 6, (**2013**) 375-386.

Ozone/PM2.5 Acute Deaths LA



3

Los Angeles California is reported to have high air pollution that is acutely lethal to its citizens, if you believe the EPA and the California Air Resources Board, CARB. Data and analysis does not support that claim.

Here we plot heart and lung deaths for people 65 and older, the blue diamonds, the ozone levels, the red circles, and PM2.5 levels, the green plus signs, for the year 2007. We use moving medians of 21 days. We start with daily air pollution levels and for the 1st 21 days we compute a median, that value where half the values are below the selected number and half are above. The median is a measure of central tendency that is minimally affected by spurious data. We moved over by one day and recomputed the median. Etc.

First, follow heart and lung deaths for people 65 and older, the blue diamonds. The deaths start high at about 120 per day in the winter and decline to the summer reaching a low of about 80 deaths at about day 200. It is well-known that daily deaths peak in the winter and reach a low point during the summer. Follow the red circles across the figure that track ozone levels. The median ozone level starts at 40 parts per billion (using the right axis) during the winter and goes to a level of about 120 ppb during the summer. It makes minor wiggles and ends the year at about 40 ppb near where it started. **Note well that daily deaths and ozone levels are inversely related.** Deaths go down as ozone goes up. Now follow PM2.5, the green plus signs. PM2.5 wiggles its way from left to right, starting at about 20 micrograms per cubic meter ($\mu\text{g}/\text{m}^3$), reaching a high point at about day 312 of about 50. Now let's track deaths and PM2.5 levels. The dramatic increase in PM2.5 from about day 280 to about day 305 corresponds to a decline in daily deaths. There appears to be no increase in daily heart and lung deaths associated with the air pollutants ozone and PM2.5. The lack of any apparent relationship between death and ozone and PM2.5 levels, as shown here, holds for 8 California air basins over 4 years.

The data used for this slide is publicly available. The EPA and CARB should make their research funded papers' data public to help resolve questions.

Ozone/PM2.5 Asthma Sacramento Valley Air Basin

Pollutant	Pearson's correlation
O ₃ (No lag)	-0.07
O ₃ (1-day lag)	-0.11
O ₃ (2-day lag)	-0.13
O ₃ (3-day lag)	-0.13
PM _{2.5} (No lag)	-0.02
PM _{2.5} (1-day lag)	0.02
PM _{2.5} (2-day lag)	0.02
PM _{2.5} (3-day lag)	0.03

Jan 1, 2010 to Dec 31, 2012. Milloy (personal communication). Data available.

4

It is often contended that ozone and PM2.5 increase hospital admission for asthma. Average ground-level ozone (O₃) and fine particulate matter (PM_{2.5}) measurements were not correlated with 19,327 patient admissions for asthma at the University of California-Davis Medical Center (UCDMC), which serves 33 counties with a population of 6M, during 2010-2012.

See

<http://junkscience.com/2013/09/03/study-ozone-not-linked-with-asthma-hospitalizations-in-major-california-hospital-system/>

for details. **Note that the data set used in this study is available.**

The following paper

Sahsuvaroglu T, **Jerrett M**, et al. Spatial analysis of air pollution and childhood asthma in Hamilton, Canada: comparing exposure methods in sensitive subgroups. *Environ Health*. 2009; 8: 14.

is of interest for two reasons. First it support no association: “There were no significant associations between any of the exposure estimates and asthma in the whole population, but ...” Second, having found no overall effect, the authors look through subgroups to claim an effect. There are countless ways to slice and dice a large, observational data set and it is well-known that any claim coming from such a process is unreliable.

One size does not fit all.

There is no reason to apply increased ozone/PM2.5 regulation to western US.

Ozone/PM2.5 do not appear to be causative for acute deaths or asthma.

5

The EPA asserts/claims that ozone (and PM2.5) are causing deaths in the US. Multiple studies and data sets find no correlation between ozone/PM2.5 and deaths in California/western US states. Therefore ozone/PM2.5 can not be causative as any negation defeats a general rule.

Ozone/PM2.5 should not be further regulated in the western US. One size does not fit all.

Also, any effects noted in the rest of the country should be considered associations, not causal.

CARB should reject the EPA position as not based on data and analysis for California .

Data Availability

I personally requested 50 air pollution data sets and got NONE.

Congress subpoenaed 2 data sets from EPA and got neither.

No public access to data sets = “Trust me Science”.

There should be no CARB/EPA regulation without

- a. Pointing to specific papers.
- b. Public access to data sets.

CARB and EPA policy is not to obtain data sets used in papers where they fund the research.

6

Peng et al. (2006) called for air pollution data sets to be public. The US National Academy published a report saying that authors should make their data sets public as a quid pro quo for their recognition as authors of a paper. The Royal Society issued a report, 2012, saying that scientists should make their data sets public.

The White House and OSTP in Feb 2013 issued a memorandum and executive order essentially requiring data sets used in papers supported by federal grants had to be made public.

Sadly all but a few researchers working in the area of air pollution do not make their data sets public.

R. D. Peng, F. Dominici, and S. L. Zeger, Commentary: Reproducible epidemiologic research. American Journal of Epidemiology 163 (2006), 783–789.

Board on Life Sciences, Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences (National Academies Press, Washington, DC, 2003) (see www.nap.edu/books/0309088593/html/).

Science as a open enterprise. <http://royalsociety.org/policy/projects/science-public-enterprise/report/>

Increasing Access to the Results of Federally Funded Scientific Research. OSTP February 22, 2013.

Paper, data, claim

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Long-Term Ozone Exposure and Mortality

Michael Jerrett, Ph.D., Richard T. Burnett, Ph.D., C. Arden Pope III, Ph.D.,
Kazuhiko Ito, Ph.D., George Thurston, Sc.D., Daniel Krewski, Ph.D.,
Yuanli Shi, M.D., Eugenia Calle, Ph.D., and Michael Thun, M.D.

American Cancer Society Cancer Prevention Study II

No association with CV deaths, corrected for PM2.5.

Ozone associated with respiratory deaths.

7

This Jerrett paper is cited 47 time in one of the EPA reports. How reliable are the claims in this paper?

First, note that the ACS Prevention Study II used in this data set is not publicly available. I've asked for the data set. Several other scientists have asked for the data set. All requests were turned down.

US House subpoenaed the data set and ACS has not provided the data set.

Jerrett et al. found no association of ozone with cardiovascular deaths, when the analysis was corrected for PM2.5 levels. They report that a subgroup analysis found an association with respiratory deaths. These positive claims appear to be the result of a "data dredge", keep asking questions until you find something positive

See following slides.

Large Search Space

Table 3. Relative Risk of Death Attributable to a 10-ppb Change in the Ambient Ozone Concentration.^a

Cause of Death	Single-Pollutant Model [†]			Two-Pollutant Model [‡]	
	Ozone (96 MSAs)	Ozone (86 MSAs)	PM _{2.5} (86 MSAs)	Ozone (86 MSAs)	PM _{2.5} (86 MSAs)
	<i>relative risk (95% CI)</i>				
Any cause	1.001 (0.996–1.007)	1.001 (0.996–1.007)	1.048 (1.024–1.071)	0.989 (0.981–0.996)	1.080 (1.048–1.113)
Cardiopulmonary	1.014 (1.007–1.022)	1.016 (1.008–1.024)	1.129 (1.094–1.071)	0.992 (0.982–1.003)	1.153 (1.104–1.204)
Respiratory	1.029 (1.010–1.048)	1.027 (1.007–1.046)	1.031 (0.955–1.113)	1.040 (1.013–1.067)	0.927 (0.836–1.029)
Cardiovascular	1.011 (1.003–1.023)	1.014 (1.005–1.023)	1.150 (1.111–1.191)	0.983 (0.971–0.994)	1.206 (1.150–1.264)
Ischemic heart disease	1.015 (1.003–1.026)	1.017 (1.006–1.029)	1.211 (1.156–1.268)	0.973 (0.958–0.988)	1.306 (1.226–1.390)

Multiplicity : 5 causes of death
 2 models
 2 subgroups
 2 air pollutants

$5 \times 2 \times 2 \times 2 = 40$ questions

8

Here is a key table from Jerrett et al. There are a number of things to note. There are 4 causes of death

under consideration. There are two statistical models, one pollutant and two pollutants. There are two

subgroups of cities, 96 and 86. There are two pollutants, ozone and PM2.5. There are 32 questions

under consideration. Given the number of questions asked, none of these associations are statistically

significant. **The Jerrett claim does not rise above chance.**

Also, note that all of the reported risk ratios are quite small, <2.0. The rules of evidence from Department of Justice say that any risk ratios less than 2.0 can not be used as evidence of causality.

[http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/\\$file/sciman00.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/sciman00.pdf/$file/sciman00.pdf)

Any small bias in the data processing or analysis could give rise to the risk ratios reported in Jerrett et al.

Ecological Covariate Adjustment

Variable

Ecologic risk factors||

Nonwhite race (%)

Home with air conditioning (%)

High-school education or greater (%)

Unemployment rate (%)

Gini coefficient of income inequality**

Proportion of population with income
<125% of poverty line

Annual household income (thousands
of dollars)††

$$2^6 = 64$$

$$2^7 = 128$$

9

It is typical in observational studies to use linear regression to “adjust out” effects from other factors. Jerrett et al. list 7 covariates under consideration. Note that Jerrett et al. may have considered other factors. With 7 factors there are 128 possible covariate adjusted models. It is customary in randomized clinical trials to say exactly how the analysis is to be adjusted before looking at the data. Usually researchers doing observational studies select how the data is adjusted by looking at the data.

Using linear regression for adjustments requires very strong statistical assumptions.

So not only might variable selection move the answers around, but the model be incorrect.

Large search space

$$40 \times 8,192 = \mathbf{327,680}$$

The data used in this paper is not available.

We are asked to trust that analysis decisions were good and claims are robust.

Any adjustment for multiple testing and/or multiple modeling renders p-values NS.

10

So the search space available, $40 \times 64 \times 128 = 327,680$ endpoints and models, is very large. Searching over a space this large is virtually guaranteed to find effects, unadjusted for searching, that are “statistically significant”/have risk ratios greater than one.

Any adjustment for multiple testing and/or multiple modeling would say that chance could produce the claimed effects in Jerrett et al.

An Observer's Comments

1. CARB should require data sets from papers cited to be public.
2. CARB should reject EPA "one size fits all" regulations.
3. CARB should provide real oversight. It has been known since at least 2005 that there are no excess deaths in California due to air pollution.

11

Quite simply, it is not science unless data sets are available. The CARB and EPA have a policy of not holding data sets produced by their funded grants. CARB has a policy of not obtaining and analyzing data sets that are used in key air pollution papers. A key product of any grant should be any data sets used in any paper where the research was funded by the CARB/EPA.

There is overwhelming evidence that the effects of ozone/PM2.5 are heterogeneous across the US. At a minimum California and the western US should be separately regulated.

Citizens expect real oversight by "independent" boards and congress.