

<http://wmbriggs.com/blog/?p=4587>

Oct 30 2011

# WILLIAM M. BRIGGS

*Statistician to the Stars!*

## A Case Of Failed Peer Review: Dust And Death

---

Published by [Briggs](#) at 10:57 am under [Climatology](#), [Statistics](#)

The distance between what civilians think peer review is and what it actually is suffers from the same failing as that evinced by Han Solo—rare pop culture reference!—when he boasted to Obi Wan Kenobi that the Millennium Falcon could do “the Kessel run in less than twelve parsecs.” Let him that readeth understand.

Peer review—an institution a bare century old, and arising solely to control the page count of proprietary journals—is the weakest filter of truth that scientists have. Yet civilians frequently believe that any work that has passed peer review has received a sort of scientific imprimatur. Working scientists rarely make this mistake in thinking.

Here is an example of how the peer review process works—or rather, does not work.

The California Air Resources Board (CARB) met on 28 October to discuss the [Jerrett report](#) “Spatiotemporal Analysis of Air Pollution and Mortality in California Based on the American Cancer Society Cohort: Final Report (as revised)” by Michael Jerrett, Richard T. Burnett, and a host of others.

This is a study which claims to have found a statistical—not actual—relationship between dust (PM<sup>2.5</sup>) and premature death for (at least part-time) California residents. I reviewed this paper and found several significant flaws in the use and interpretation of statistical methods. Here is the most significant: “I find further that the summary in the abstract—and therefore the only part of the report liable to be read by most—to be the result of either poor work or deliberate bias toward a predefined conclusion.”

The authors prepared and intensely investigated a series of complex statistical models. There were nine models in total, each having particular strengths and weaknesses. Each had several subjective “knobs” and “dials” to twist. Only *one* model of the nine (p. 108) showed a “statistically significant” relationship between mortality and PM<sup>2.5</sup>, and that only barely; and in that model, only one sub-model showed “significance.” The other eight models showed no relationship. Some models even hinted that PM<sup>2.5</sup> *reduced* the probability of early mortality. With such a large number of tests and “tweaks”, the authors were practically guaranteed to find at least one “significant” result, even in the absence of any effect. Nowhere did the authors control for the multiplicity of testing, even though such controls are routine in statistical analyses of these sort.

You may even listen to the CARB meeting ([mp3: 80 minutes](#)).

There were seven separate critiques presented, all by peers with significant and lengthy expertise in relevant areas. Comments were provided by: [Jim Enstrom](#), [Matthew Malkan](#), [John Dunn](#), [Frederick Lipfert](#), [Gordon Fulks](#), [Skip Brown \(Delta\)](#), and [yours truly \(updated\)](#). A quote from each:

Enstrom: "The results in the Jerrett Report do not support the authors' claim."

Malkan: "[The] Abstract, Key Results, Key Findings, and Conclusion sections which do not accurately reflect, and are even contradicted by, the actual data analysis presented in this report."

Dunn: "[W]e have a modeling paper that looks a lot like the nonsense put out on global warming modeling, and it has the taint of data torturing in its presentation."

Lipfert: "I find that the consistent and overwhelming defect in this report is its arbitrary selectivity:...Selecting heart disease as the most important cause of death, while ignoring the apparently significant **beneficial** relationships with cancer."

Fulks: "With the apparent approval of the agency staff, the authors have refused to correct or even address mistakes."

Brown: "This 'new' report, whose entire purpose is to justify previously passed regulation, does not address the many scientific comments made rebutting the conclusions reached in the original report."

Briggs: "I find further that the summary in the abstract and therefore the only part of the report liable to be read by most to be the result of either poor work or deliberate bias toward a predefined conclusion."

CARB had earlier implemented regulations based on the assumption that particulates kill. [The story of how](#) they came by that assumption is odd, but is not relevant here. The Jerrett report was meant to bolster the research that led to the regulations that were already in force. Therefore CARB was to decide only whether to accept or reject the Jerrett report. Despite the numerous flaws and objections given by Jerrett's peers, after a few minutes discussion CARB voted to accept the report. In one sense, this was fine, because without this acceptance Jerrett could not claim that he fulfilled his contractual obligations.

But in the sense of approving the findings themselves, this peer review process clearly failed. This is true even if the large numbers of criticisms were wrong or inconclusive. This is because rebutting serious criticism takes time, thought, and effort. CARB did not attempt to rebut any of the criticism beyond saying because what Jerrett claimed was also claimed by other authors, therefore Jerrett's findings should be accepted. Another commentator said that because science is imperfect, we may as well accept Jerrett's findings.

Then the criticisms were not wrong, especially the "cherry-picking" critique cited above. The statistical mistake (choosing only the significant model which showed "significance" and ignoring the ones that did not, and for not correcting for multiple tests) made by Jerrett is enormous, and if addressed would have caused the claim of "statistical significance" to disappear. It is thus more likely that what Jerrett claims is false.

This is a (not at all unusual) failure of peer review.