Comments on “PM$_{2.5}$ Mortality in Long-term Prospective Cohort Studies: Cause-Effect or Statistical Association?”

Most models of how science works propose that competition between ideas contributes to the advancement of knowledge. Criticism of scientific work plays a part in facilitating such competition by exposing the strengths and weaknesses of rival explanations, encouraging debate, and suggesting alternatives. Nevertheless, not all criticism has equal value to the scientific process.

In a review of epidemiologic studies on fine particles and mortality that appeared recently in EHP, Gamble (1) charged that the two major studies on this topic (2,3) may have been compromised by bias, yet he offered no serious effort to evaluate the alleged errors with the same standards of rigor demanded of the original studies.

For example, in Gamble’s (1) claim that the study findings are compromised by the ecologic fallacy, he failed to address two important issues. First, the major prospective studies of fine particles and mortality are not classical ecologic designs because only air pollution exposures are measured on the aggregate level; the outcome and potential confounders are based on individual-level measurements. Thus the biases stemming from the ecologic fallacy do not apply to these studies. Instead, these should be viewed as individual level studies in which exposure is measured with error (4). The calculations in Gamble’s Table 2 (1) are erroneously portrayed as demonstrating the ecologic fallacy. Instead they appear to present strong evidence of a supralinear dose–response relationship between particles and mortality. This type of dose response would be expected if there existed a subset with much greater susceptibility, e.g., a bimodal distribution of susceptibility. Second, the inference that the U.S. Environmental Protection Agency (EPA) and others have drawn from the studies results is logically consistent with evidence based on exposure measured at the group level. The correct inference is that individuals living in communities with high air pollution levels have a higher risk of dying than people living in communities with low pollution levels and, therefore, that lowering community-wide air pollution levels should reduce community mortality rates. Such a policy is a logical and efficient means of minimizing the health impact of a widespread exposure (5).

Even when the criticism focuses on specific factors that might explain the results of the studies, it does not consistently address the potential magnitude and direction of alleged biases. For example, Gamble (1) concluded that lung function is a probable confounder of the observed relative risk for PM$_{2.5}$ because the average lung function of the Six Cities cohort differs by city and because reduced lung function is a risk factor for mortality. However, neither the text nor Gamble’s Figure 3 (1) explicitly identify possible ranges of confounding that this variable may have produced in the relative risk estimate of 1.26 derived in the original Six Cities Study. Gamble’s Figure 3C shows that the average forced expiratory volume in 1 sec (FEV$_1$) differs by approximately 0.1 L between the cohort members in the driest and cleanest cities. In light of the cited 1.52 relative risk estimate for total mortality associated with a 1-L decrease in FEV$_1$, it is difficult to imagine how adjustment for a FEV$_1$ difference one order of magnitude smaller could explain the observed association (6). Perhaps more importantly, reduced lung function could be on the causal pathway between chronic exposure to particles and mortality, such that any adjustment for FEV$_1$ would introduce bias. Sedentary lifestyle was postulated as another potential confounder; however, Gamble (1) presented no evidence to suggest that sufficient differences in sedentary lifestyle among the six cities could account for the observed particle/mortality association.

Gamble (1) raised further doubts about cohort studies of mortality and fine particles by reviewing lists of criteria for epidemiologic studies (7,8). Hill never intended that his standards be used to exclude evidence. In his famous 1965 paper (7), Hill wrote,

I do not believe ... that we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we accept cause and effect.

Hill emphasized that guidelines can be helpful at the margins of epidemiologic interpretation, but the contribution of any particular study must be evaluated through careful assessment of the individual facts. He wrote (7).

None of my nine viewpoints ... can be required as a sine qua non. What they can do, with greater or less strength, is help us to make up our minds on the fundamental question....

Hill’s criteria appropriately would require a careful review not of one study design, but rather of the entire body of literature pertaining to the hypothesis of a causal association between air pollution exposure and health. This literature includes experimental studies in animals and clinical and epidemiologic studies in humans, examining outcomes ranging from direct measures of lung function to self-reported symptoms to hospital admissions and death.

Hertz-Picciotto’s (8) criteria on the use of epidemiology in quantitative risk assessment are also cited by Gamble (1), with emphasis on their application for setting air quality standards. In fact, Hertz-Picciotto’s framework was developed for the specific purpose of classifying “individual epidemiologic studies as to their adequacy for use in dose–response extrapolation” (8), not for assessing weight of evidence regarding causation. Gamble (1) applied these criteria to a study design rather than individual studies; because he characterized them as having an ecologic design, he wrongly concluded that the alleged design weaknesses threaten the evidential base of the EPA’s new PM$_{2.5}$ standard. This use of Hertz-Picciotto’s criteria (8) does not reflect the spirit in which they were intended, that is, “to have a reliable process for making the best use of available data ...” for dose–response extrapolation.

It is appropriate to criticize epidemiologic findings, especially when they have major implications for public policy, as the studies of fine particles do. However, it is much more helpful to evaluate the evidence that a given type of bias did (or did not) occur and to quantify its direction and magnitude than merely to suggest it is a possibility. Sensitivity analysis is an excellent tool for quantitative exploration of potential biases that can be used to gauge which of the many biases that can be envisioned are plausible explanations for a set of results and which are not. The magnitude of confounding can be calculated from several parameters related to the univariate distribution of the postulated confounder and bivariate distributions involving the confounder, the exposure of interest, and the outcome (9). The kind of detailed analysis that these techniques encourage is fundamental to serious scientific criticism. Without these elements, Gamble’s review (1) is largely an expression of opinion, more appropriately published as a commentary.

Dana Loomis
Irva Hertz-Picciotto
Marie O’Neill
School of Public Health
University of North Carolina-Chapel Hill
Chapel Hill, North Carolina
E-mail: dana.loomis@unc.edu

REFERENCES AND NOTES

1. Gamble JF. PM$_{2.5}$ and mortality in long-term prospective cohort studies: cause-effect or statistical association? Environ Health Perspect 106:535-548 (1998).
2. Dockery DW, Pope CA II, Xu X, Spengler JD, Ware JH, Fay ME, Ferris BG Jr, Speizer FE. An association between air pollution and mortality in six US cities. N Engl J Med 329:1753-1763 (1993).
3. Pope CA III, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, Heath CW Jr. Particulate air pollution as a predictor of mortality in a prospective study of US adults. Am J Respir Crit Care Med 151:669-674 (1995).

4. Morgenstern H. Ecologic studies. In: Modern Epidemiology (Rothman KJ, Greenland S, eds). 2nd ed. Philadelphia, PA: Lippincott Raven Publishers, 1998:459-480.

5. Rose G. Sick individuals and sick populations. Int J Epidemiol 14(1):32-38 (1985).

6. Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Sosin M. Smoking and lung cancer: recent evidence and a discussion of some questions. J Natl Cancer Inst 22:173-203 (1959).

7. Hill AB. The environment and disease: association or causation? Proc R Soc Med 58:295-300 (1965).

8. Hertz-Picciotto I. Epidemiology and quantitative risk assessment: a bridge from science to policy. Am J Public Health 85:494-491 (1995).

9. Flanders DW, Khoury MJ. Indirect assessment of confounding: graphic description and limits on effect of adjusting for covariates. Epidemiology 1:239-246 (1990).

Response to Loomis et al.

I would like to address comments of Loomis et al. about inferences drawn from studies using group-level exposure variables, the use of the tobacco analogy, the application of Hill’s criteria for causality (7), and the use of the Hertz-Picciotto criteria for evaluating studies (2). Whether the hybrid studies under discussion (3-5) are considered partly ecological (6,7) or individual level with exposure misclassification, bias (ecological or otherwise) is possible and should be checked. I hope that these discussions will lead to more discussions of the interplay between outcomes and confounders measured at the individual level and exposure measured at the group level.

Loomis et al. suggest that the biases stemming from the “ecologic fallacy” do not apply to the PM$_{2.5}$ air pollution studies because they are individual-level studies where exposure is measured with error. That is, by implication there is one PM exposure variable. But as indicated by Morgenstern (8,9), ecologic bias can arise when the mean of a group-level exposure variable has an effect on the individual-level exposure. By this definition there will be ecologic bias whenever the ecologic exposure variable has an effect, and when there is also an individual-level exposure effect in addition to the ecologic exposure effect. Unmeasured individual-level exposure to PM$_{2.5}$ from all sources can be several orders of magnitude higher than ambient PM$_{2.5}$ concentrations (16) because of extensive exposure to unmeasured sources such as tobacco and indoor combustion. These individual-level exposures vary for individuals within the group and contribute to the individual-level risk. The additional effect of ambient exposure provides the group-level component that leads to ecologic bias.

The American Cancer Society (ACS) Study (4) and the Six Cities Study (5) suggest that an increase of about 20 µg/m$^3$ PM$_{2.5}$ results in a 20-30% increase in total mortality. I sought to test the consistency of these findings by comparing risk estimates based on group-level exposure estimates to those based on individual-level exposure to a similar but more thoroughly studied particular (i.e., tobacco smoke). Applying the models developed in these studies to tobacco smoke, one can predict that a 20-µg/m$^3$ difference in ambient PM$_{2.5}$ between cities is too small to result in a measurable difference in overall mortality (6). If this is true, the differences in mortality between cities may be due to causes other than differences in PM. Whether there is ecologic bias, exposure misclassification bias operating at the individual level, or uncontrolled bias from other sources, the tobacco analogy suggests that bias away from the null may be operating in these studies.

Loomis et al. suggest that the tobacco analogy presents “strong evidence of a supralinear dose–response relationship between particles and mortality.” In order to fit the data, the degree of supralinearity would have to be enormous. In fact, an increase of 19.6 µg/m$^3$ in ambient PM$_{2.5}$ and an increase of 16,000 µg/m$^3$ from smoking would have to result in a similar 20-30% increased risk (Figure 1). It is not plausible that two increases in exposure, which differ by almost three orders of magnitude, would both produce the same response. A more plausible inference is that either the PM$_{2.5}$ or the smoking risk estimates are in error. However, I would place more credence in the smoking relative risks (RRs) because smoking is measured at the individual rather than the group level, and the smoking RRs are compatible with a large body of literature.

It is not necessarily correct to infer, as Loomis et al. do, that lowering community-wide air pollution below existing levels will reduce community mortality rates. In making this inference, one assumes there is independent evidence for a causal relationship between ambient PM$_{2.5}$ and mortality. These studies (4,6) showed that there were differences in total mortality, but did not show why mortality was higher in cities with higher PM$_{2.5}$ concentrations. IF PM$_{2.5}$ is the reason for increased mortality, all important individual risk factors must be taken into account to a reasonable degree. Total mortality has a large number of risk factors. It is speculative therefore to assume, as Loomis et al. do, that lowering PM$_{2.5}$ concentrations beyond existing levels will provide a “logical and efficient means of minimizing the health impact of a widespread exposure,” and that the proposed cure will produce the desired effect. The tobacco analogy provides evidence against such an effect.

Loomis et al. state that the effects of potential confounders are too small to explain the observed associations in the PM studies. In my paper (6) I assessed differences in lung function and sedentary living as two examples of possible confounders because some evidence was available to me. Even in these cases, individual-level data were not available to adequately estimate or adjust for these effects (6). There are undoubtedly many other examples such as personal lifestyle factors or other inadequately controlled variables that are correlated with dirty versus clean cities or with geography. Based on the tobacco analogy, it appears that whatever biases are operating resulted in a large overestimate of the PM$_{2.5}$ risk.

I agree with Loomis et al. to apply Hill’s criteria appropriately requires a “careful review not of one study design, but rather of the entire body of literature pertaining to the hypothesis of a causal

---

**Figure 1.** Association of total mortality with group level ambient PM$_{2.5}$ exposure and individual level tobacco smoke exposure. Data from the Six Cities Study (5).

*Approximately 16,000 µg/m$^3$.**