American Journal of Epidemiology
Copyright © 2003 by the Johns Hopkins Bloomberg School of Public Health
All rights reserved

Vol. 157, No. 12
*Printed in U.S.A.*
DOI: 10.1093/aje/kwg089

# Response to Dr. Smith: Timescale-dependent Mortality Effects of Air Pollution

**Francesca Dominici[1], Aidan McDermott[1], Scott L. Zeger[1], and Jonathan M. Samet[2]**

[1] Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD.
[2] Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD.

*Received for publication January 2, 2003; accepted for publication February 5, 2003.*

Time-series studies have indicated that there is an association between day-to-day variation in ambient air pollution concentrations and day-to-day variation in numbers of deaths, after data are controlled for more slowly varying confounding factors such as weather, seasonality, and long-term trends. However, the estimated coefficients in the time-series studies are uninformative as to the amount of life lost due to pollution exposure, particularly within susceptible populations (1).

If individuals who were severely ill and were expected to die shortly were the only people affected by current levels of air pollution, reducing ambient concentrations would not necessarily increase life expectancy significantly. This phenomenon of only brief advancement of the timing of death has been referred to as "short-term mortality displacement," as well as by the unfortunate term "short-term harvesting." While no lives should be shortened by air pollution, society suffers a much smaller loss if air pollution affects only frail persons without great loss of life expectancy. Our paper (2) was motivated by the need to find methods of assessing short-term harvesting for studies of air pollution and other environmental agents.

In this rejoinder to the commentary of Dr. Richard Smith (3), we briefly 1) review the conceptual framework under which short-term harvesting would occur, 2) illustrate how our timescale model would detect short-term harvesting, and 3) summarize the statistical evidence supporting short-term harvesting.

A compartmental model (4–6) sets a biomedical stage for approaching the assessment of short-term harvesting. Suppose that the population can be divided into two groups according to susceptibility to an air pollution episode: low-risk and high-risk. On any given day, people in the low-risk pool can become frail and move into the high-risk pool (T1) and people in the high-risk pool can become healthier and move into the low-risk pool (T2) or can exit the high-risk pool by dying (T3). Assuming a steady-state condition, T3 = T1 – T2; that is, there is equilibrium between the number of people who die (T3) and the number of people who enter the susceptible pool (T1), net the number of people who recover (T2). We assume that under short-term harvesting, an air pollution episode would affect only transition out of the high-risk pool (T3), without increasing net recruitment into the high-risk pool (T1 – T2). Therefore, for some days after an air pollution episode, the susceptible pool would be depleted, and the daily death count would be diminished. (See Schwartz (6) for further details.)

This phenomenon can be further described by a distributed lag model (7–9) that includes several lags of the pollution variables:

$$\log E\left|Y_t\right| = \sum_{l=1}^{L_1} \theta_l x_t + \sum_{l=L_1+1}^{L_2} \theta_l x_{t-l} + \text{confounders}, \quad (1)$$

where $\theta_l$ represents the percentage increase in mortality associated with a 10-unit increase in the air pollution level $l$ days after an air pollution episode. Under short-term harvesting, we would expect to see an increase in deaths above the baseline level for $L_1$ days after the air pollution episode, followed by another $L_2$ days of decrease owing to the depletion of the pool. If the air pollution episode affected "only" the high-risk pool, as in the case considered here, the area above the baseline (the number of deaths attributable to the episode) would be roughly equal to the area below the baseline (the number of deaths necessary to replenish the high-risk pool). Figure 1 illustrates a hypothetical sinusoidal time course of lagged air pollution effects with $L_1 = L_2 = L$ and with rebounds of $L = 1$, 3, and 7 days, respectively.

To illustrate that our timescale regression approach can detect short-term harvesting, we implemented a simulation study as follows. First, we reanalyzed the 1987–1994 data from Philadelphia, Pennsylvania, to obtain estimates of the meteorologic (current day) and long-term trend effects, omitting total suspended particulates (as in Dr. Smith's commentary). Second, we simulated Poisson count data from model 1 with $\theta_l = \beta_1/L$ for $l = 1, \ldots, L$ and $\theta_l = -\beta_1/L$ for $l = L + 1, \ldots, 2L$, and we assumed that $\beta_1 = 0.001$—that is, a

Correspondence to Dr. Francesca Dominici, Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, MD 21205 (e-mail: fdominic@jhsph.edu).

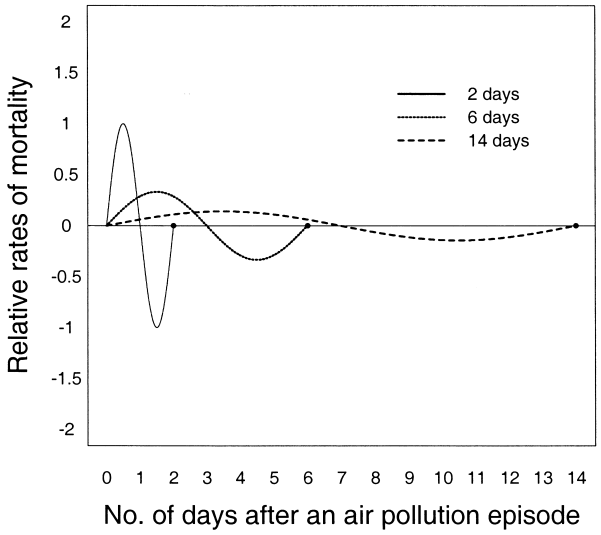


**FIGURE 1.** Hypothetical time courses of lagged air pollution effects on mortality under short-term harvesting. Mortality relative rates initially increase for $L$ days after the air pollution episode, followed by another $L$ days of immediate rebound, owing to the depletion of the pool. The three lines indicate time courses of wavelengths of 2, 6, and 14 days, respectively.

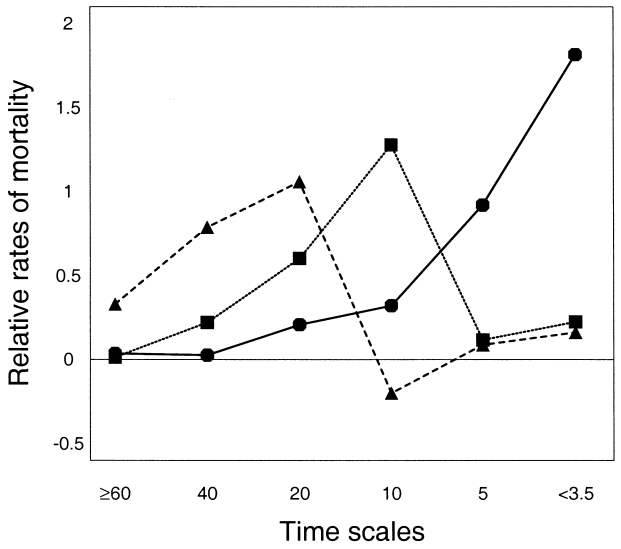


**FIGURE 2.** Timescale estimates obtained under the true distributed lag models illustrated in figure 1. The estimates are plotted at ≥60, 40, 20, 10, 5, and <3.5 days, denoting approximate midpoints of the intervals ≥60 days, 30–59 days, 14–29 days, 7–13 days, 3.5–6 days, and <3.5 days, respectively. The circles, squares, and triangles indicate timescale estimates obtained under a distributed lag model with wavelengths of 2, 6, and 14 days, respectively.

1 percent increase in mortality for a 10-μg/m$^3$ increase in total suspended particulates. We considered three hypothetical time courses of the lagged air pollution effects (wavelengths of $2L$ for $L$ = 1, 3, and 7 days), representing short-term harvesting.

Figure 2 shows the timescale estimates of particulate matter effects made with our approach for the three cases. We found that under a distributed lag model with $\theta_l$ having wavelengths of 2, 6, and 14 days, we detect air pollution effects at timescales of <3.5 days, 5–10 days, and 10–20 days, respectively. This indicates that the greatest air pollution effect occurs at the timescales determined by the true distributed lag model.

Note that in Dr. Smith's simulation studies (cases A, B, and C), data are also generated from model 1, where $\theta_l = \beta_l / L$ for $l = 1, \ldots, L$, but with $\theta_l = 0$ for $l = L + 1, \ldots, 2L$. Therefore, the time course of the $\theta_l$ does not have a rebound, indicating that an air pollution episode might affect both net recruitment (Tl – T2) and transition into the high-risk pool (T3). Although this is a sensible model that estimates air pollution effects in correspondence with average past exposures, it represents a different compartment model than the one used here. More specifically, it does not represent the harvesting case in which the distributed lag coefficients become negative due to the depletion of the high-risk pool. Dr. Smith's model D in his figure 3 is a "harvesting model" but with longer residence in the pool of susceptible people (30 days). Zeger et al. (4) performed a similar simulation study and showed that when data are generated from a compartment model, the association between deaths and air pollution becomes nonnegligible only at timescales shorter than the mean time of residence in the pool of frail persons.

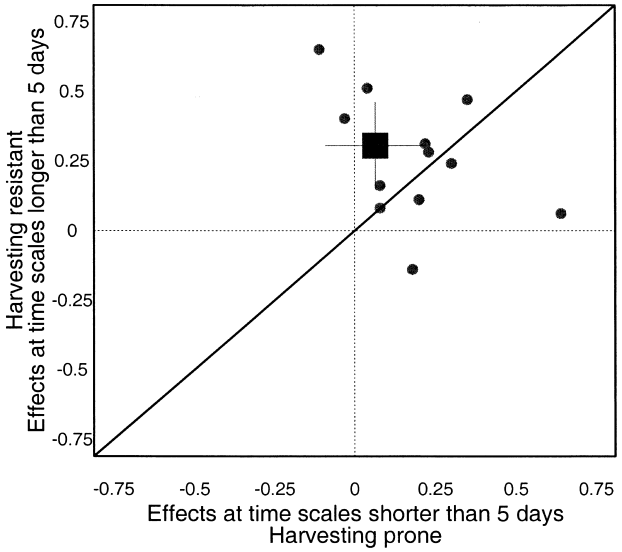


**FIGURE 3.** Estimated relative rates of mortality due to air pollution at timescales shorter than 5 days (harvesting-prone) versus relative rate estimates obtained at timescales longer than 5 days (harvesting-resistant) for four US cities (Pittsburgh, Pennsylvania; Minneapolis, Minnesota; Seattle, Washington; and Chicago, Illinois) and three cause-specific mortality outcomes (total mortality, cardiovascular and respiratory mortality, and mortality due to other causes) during the period 1987–1994. The square is placed at the averages of the 12 timescale estimates, with segments representing their 95% confidence intervals.

Consistently with this result, the timescale coefficients in part D of Dr. Smith's figure 3 are statistically significant for all timescales shorter than 30 days. Additional simulation studies in which data are generated from several compartment models and reanalyzed using our timescale approach are detailed in a recent report by Fung et al. (10).

In summary, in response to Dr. Smith's first conclusion, we interpret our figures 1 and 2 as showing a correspondence between the true time-lagged response to an air pollution event and our timescale estimates and as distinguishing short-term harvesting from other forms of time-lagged exposure-response relations with an air pollution event (such as averages of past exposures).

Is there evidence to support the existence of short-term harvesting? To address this question, we need to test the hypothesis that there exists an association between air pollution and mortality at the shortest timescales. Looking at table 1 of our manuscript (2), we have found that the pooled estimates at timescales shorter than 3.5 days are equal to –0.07 percent (95 percent confidence interval: –0.40, 0.27) for total mortality, 0.06 percent (95 percent confidence interval: –0.39, 0.51) for cardiovascular and respiratory mortality, and –0.26 (95 percent confidence interval: –0.75, 0.24) for other-causes mortality, providing evidence contrary to the hypothesis that the pollution-mortality association is largely or entirely due to short-term harvesting.

A related question is: Are air pollution effects at medium and long timescales (harvesting-resistant) smaller than the effects of air pollution at shorter timescales (harvesting-prone)? Figure 3 shows the 12 relative rates of mortality (for four cities and the three specific mortality groups) at timescales less than 5 days ($\hat{\beta}^{c}_{<5}$) versus estimates calculated at timescales greater than 5 days ($\hat{\beta}^{c}_{>5}$). The square is placed at the weighted averages of the 12 coefficients ($\hat{\beta}^{*}_{<5}$, $\hat{\beta}^{*}_{>5}$), with the weights being equal to the inverse of their statistical variances. The two segments represent the 95 percent confidence intervals of the weighted averages. Note that almost all of the coefficients are above the diagonal. The posterior probability that $\hat{\beta}^{*}_{<5}$ is larger than $\hat{\beta}^{*}_{>5}$ is 0.01, and it remains small when we average the relative rates using a random-effect model with a substantial amount of heterogeneity. In addition, this posterior probability is still very small when we choose as cutoff points timescales smaller than 3.5 days or timescales smaller than 10 days (equal to 0.006 and 0.12, respectively). In summary, this data analysis suggests that "harvesting-resistant" estimates are larger than "harvesting-prone" estimates; this is inconsistent with a phenomenon of short-term harvesting only. Note that these results are not suggesting that air pollution affects only healthy people. Rather, our results are consistent with all persons' being affected by air pollution, not only the very frail.

Could a different method of adjusting for weather qualitatively change our conclusions? As in any time-series study of air pollution and health using one city or a few cities, timescale effects can be sensitive to adjustment for confounding factors. However, the sensitivity analyses depicted in figure 4 (bottom) of our paper and Dr. Smith's figure 2 suggest that the basic findings obtained by our timescale method do not change with the different models for temperature. In any case, additional work on this topic would be valuable.

Thus, our timescale approach is able to detect short-term harvesting within the conceptual framework described by a compartment model with high-risk and low-risk pools of individuals. Although this is a reasonable starting point and makes the problem identifiable, the model must be a simplistic approximation of reality. Additionally, the sensitivity of the results to the degree of adjustment for confounding factors, although investigated in many sensitivity analyses, remains an issue. Because the information in time-series analyses comes from the variability across time in exposure and outcome, air pollution effects corresponding to slowly varying exposures (that is, air pollution components at timescales longer than 30 days and average past exposure of at least 30 days) are very unstable. Finally, the limited number of cities for which daily data are available further increases uncertainty. In summary, we certainly agree with Dr. Smith's general message: Results of any model of the relation between air pollution and mortality require careful interpretation, with consideration of the assumptions made and the sensitivity of the findings to those assumptions.

## ACKNOWLEDGMENTS

The authors thank Dr. Smith for his thoughtful comments and for carrying out additional analyses that provided further insights into their approach.

## REFERENCES

1. Kunzli N, Medina S, Kaiser R, et al. Assessment of deaths attributable to air pollution: should we use risk estimates based on time series or on cohort studies? Am J Epidemiol 2001;153:1050–5.
2. Dominici F, McDermott A, Zeger SL, et al. Airborne particulate matter and mortality: timescale effects in four US cities. Am J Epidemiol 2003;157:1055–65.
3. Smith RL. Invited commentary: timescale-dependent mortality effects of air pollution. Am J Epidemiol 2003;157:1066–70.
4. Zeger SL, Dominici F, Samet JM. Harvesting-resistant estimates of pollution effects on mortality. Epidemiology 1999;89:171–5.
5. Smith R, Davis J, Sacks J, et al. Regression models for air pollution and daily mortality: analysis of data from Birmingham, Alabama. Environmetrics 2000;100:719–45.
6. Schwartz J. Is there harvesting in the association of airborne particles with daily deaths and hospital admission? Epidemiology 2001;12:55–61.
7. Almon S. The distributed lag between capital appropriations and expenditures. Econometrica 1965;33:178–96.
8. Zanobetti A, Wand M, Schwartz J, et al. Generalized additive distributed lag models. Biostatistics 2000;1:279–92.
9. Zanobetti A, Schwartz J, Samoli E, et al. The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. Epidemiology 2002;13:87–93.
10. Fung K, Krewski D, Burnett R, et al. Testing the harvesting hypothesis by time domain analysis. (Technical report). Ottawa, Ontario, Canada: Institute of Population Health, University of Ottawa, 2003.