Mortality and Air Pollution for Santa Clara County, California, 1989–1996

David Fairley

ABSTRACT

The data are reanalyzed using generalized additive models (GAMs*) with stricter convergence criteria. The results are essentially unchanged with coefficients differing by at most ± 0.5 standard error (SE) from the original and relative risks differing by at most ± 0.01 SE. Various inference methods are compared. The method comparing GAM coefficients with standard errors generated from generalized linear models (GLMs) produces *P* values similar to simulations. The S-Plus ANOVA (analysis of variance) feature often gave conservative results. As found previously, a statistically significant relation existed between daily nonaccidental mortality and every criteria pollutant, either on the same day or lag one. Particulate matter (PM) less than 2.5 µm in diameter ($PM_{2.5}$) and nitrate (NO_3) predominate when included in models with carbon monoxide (CO), nitrogen dioxide (NO_2), and sulfate (SO_4). Coarse-fraction PM, less than 10 µm in diameter (PM_{10}), was not statistically significant. A new ozone (O_3) variable—the daily number of parts per billion (ppb)-hours greater than a 60-ppb threshold (o3ppbgt60)—was found to have a statistically significant relation with nonaccidental mortality even when included in a regression jointly with $PM_{2.5}$ or NO_3 and was also significantly related to cardiovascular mortality.

INTRODUCTION

> Ockham's Razor: A rule in science and philosophy that... the simplest of two or more competing theories is preferable and that an explanation for unknown phenomena should be first attempted in terms of what is already known.
>
> —*The American Heritage Dictionary,* 2nd Edition

Problems have recently been reported in use of the GAM function in the S-Plus statistics package (Dominici et al 2002). The GAM procedure uses an iterative algorithm to find the solution. Further analysis found that the default number of iterations and the stopping criterion sometimes produced results that were far from those obtained after the algorithm converged. Furthermore, the standard errors provided in S-Plus were shown to underestimate the true standard errors. Questions about the results using GAM are of particular concern because the current reevaluation of the national particulate standards are based on time-series studies, many of which used this function (for example, Dominici et al 2000; Kelsall et al 1997; Moolgavkar 2000; Saez et al 2002; Samet et al 2000; Schwartz 1994).

Because of questions raised by these problems, we decided to reanalyze the data from our previous study (Fairley 1999). In our previous analysis, we used GAM but with a more stringent stopping criterion than the default so that the results in this reanalysis changed only in a minor way. Moreover, the previous analysis did not use the standard errors provided in the S-Plus summary GLM function. Instead we used an F test from the ANOVA function, which is based on the change in −2 log (likelihood), to determine whether the addition of a variable was statistically significant. Thus, conclusions about statistical significance are also essentially unchanged.

This analysis contained some innovations. One was to augment the $PM_{2.5}$ data using other variables. A second was to use an alterative ozone variable—o3ppbgt60.

DATA AND METHODS

The same data and variables were used as in the previous study with two additions (o3ppbgt60 and pm2.5aug [$PM_{2.5}$ measurements augmented with $PM_{2.5}$ predicted from COH and PM_{10}]): 24-hour $PM_{2.5}$, PM_{10}, $PM_{10-2.5}$, NO_3 and SO_4; 24-hour averaged coefficient of haze (COH) and NO_2; and 8-hour averaged CO and O_3, for 1989 through 1996 from the

* A list of abbreviations and other terms appears at the end of the section.

This short communication report is part of an HEI Special Report, which also includes 20 other reports, a section on NMMAPS II, two HEI Commentaries, and an HEI Statement. Please address correspondence about this section to Dr David Fairley, Bay Area Air Quality Management District, 939 Ellis Street, San Francisco CA 94109.

4th Street monitoring site in San Jose Santa Clara County, California. Regressions were performed for all nonaccidental deaths to residents who died in the county: *International Classification of Diseases* (ICD) codes 0–799; respiratory mortality (ICD codes 11, 35, 472–519, 710.0, 710.2, 710.4); and cardiovascular mortality (ICD codes 390–459).

The same methods were used as in the previous study except for increasing the stringency of the stopping rule. Briefly, a Poisson regression model was fit with GAM terms for day of year and trend with the best fit determined by minimizing Akaike information criterion (AIC). GAM terms for minimum and maximum temperatures were added to this model and again the model with the minimum AIC was determined. The degrees of the smooths were the same as found previously. Various pollutant terms were then added to this best fitting model.

The stopping rule was tightened with epsilon and backfit epsilon set to 10^{-12}, whereas previously it had been 10^{-4} (already been more stringent than the S-Plus default, 10^{-3}) in the previous analysis. The maximum number of iterations was raised from the default, 10, to 10^7. These were chosen based on trial and error so that there was no change if the stopping rule was made more stringent.

For comparison, a parallel analysis was done using GLM regression. Here, natural splines were used in place of smoothing splines. AIC was again used to find the optimal degree of freedom (*df*) for day of year, trend, and minimum and maximum temperatures. The same strict convergence criteria were used as with the GAM regression.

The reanalysis updates Tables 4, 5 and 6 from Fairley (1999). As mentioned previously, inference on whether a variable included in the fit was statistically significant was based on the deviance [change in -2 log(likelihood)] from adding that variable, using the ANOVA feature in S-Plus.

S-Plus does not provide standard errors for GAM regressions and advises using a suggestion from Chambers and Hastie (1992): "In practice, one can always approximate the nonparametric term parametrically (and even conservatively) using functions such as bs () [B spline] or ns () [natural spline], and use the inexpensive parametric standard-error curves." Following this advice, standard error estimates were obtained by running GLM regressions, replacing smoothing splines with natural splines of the same degree. These standard errors were used to compare the differences in coefficients between this and the original analysis and between GAM and GLM regressions.

To check the reliability of these inferences, simulations were performed from the model fit without pollutant variables (that is, assuming that the null-model fitted parameters were the true parameters), generating Poisson variates from this model and then refitting the model including a pollutant variable. This simulated the null distribution of the pollutant coefficient, allowing comparison of the pollutant coefficient estimated using the actual data. A more general simulation was also rerun for $PM_{2.5}$ and pm2.5aug. Here the whole model-building process is simulated—finding the model with the best AIC for time/season terms, then finding the model with the best AIC for temperature, and finally adding the PM term. See Fairley (1999) for more details. The only difference with the 1999 simulations was using epsilons of 10^{-12} and 10^3 iterations.

As a third comparison, inferences were based on the ratio of the fitted GAM coefficients to the GLM standard errors mentioned previously.

Note that some errors were found in the original analysis. One was the miscoding of missing ozone values. The other two were mistaken entries. One mistake was the relative risk of COH with respiratory mortality. The relative risk (RR) 1.07 should have been 1.10 (and highly significant). The second mistake was the RR for $PM_{2.5}$ for spring seasonal regression, which was listed in the original document as 1.05 when it should have been 0.92.

The original dataset had only 408 $PM_{2.5}$ observations so the power was borderline for detecting an effect. We found that $PM_{2.5}$ could be well predicted from PM_{10} and COH, more than doubling of the number of $PM_{2.5}$ values. In particular, a linear regression yielded a fit of $PM_{2.5} = 0.392 \times PM_{10} + 14.0 \times COH - 4.02$, with a regression standard error of 4.49 $\mu g/m^3$ and a multivariate coefficient of determination (R^2) of 0.88. The total number of pm2.5aug observations was 835.

A second and perhaps more significant innovation was to use an alternate variable for ozone. Previously, 8-hour maximum ozone (oz8hr) was used because it corresponded to the national standard. However, the 8-hour averages often contain very low ozone values; clinical studies have not found ozone health effects below about 80 ppb and the natural background is approximately 40 ppb. Thus, it seemed reasonable to consider the daily number of ppb-hours above a threshold, with thresholds between 40 ppb and 80 ppb. The highest correlations with daily mortality were found using a threshold of 60 ppb, so those results are reported here. Figure 1 shows this variable plotted against daily mortality. This variable, denoted *o3ppbgt60*, is highly skewed. The largest values are clearly influential observations.

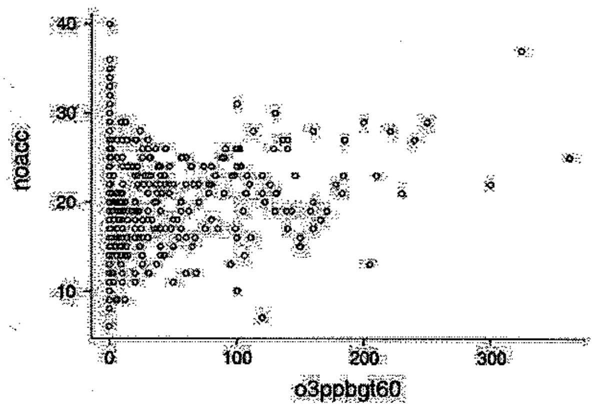

Figure 1. Daily nonaccidental deaths versus daily ozone ppb-hours greater than 60 ppb.

RESULTS

COMPARISON OF RESULTS FROM STRICT AND DEFAULT GAM ANALYSIS

Tables 1a and 1b present RRs from the old and new GAM regressions for all nonaccidental cause mortality. To compare the magnitude of the changes in coefficients, the GLM standard errors were used to normalize the differences (that is, taking the ratio of the difference between the coefficient from the old and new analyses to the standard error). There were 34 analyses that were repeated in Tables 1a and 1b. The range of normalized differences was −0.34 to +0.40, but 40 were between −0.09 and +0.05.

The relative risk estimates remained virtually identical, changing by at most ± 0.01. The statistical significance changed in only one case: same-day oz8hr was not found to be statistically significant in the new analysis (although the estimated coefficient increased somewhat). The difference was not due to GAM but to the miscoding of ozone data in the original analysis (as mentioned earlier).

The results for cardiovascular and respiratory mortality in Table 2 are almost identical to those in the original analysis. The one exception, as mentioned previously, is that the RR for lag COH was 1.10, which is highly significant; the previously reported RR of 1.07 was an error. The pm2.5aug data were statistically significant in the regression with cardiovascular mortality, whereas the coefficient for the nonaugmented $PM_{2.5}$ was not.

The RRs for Table 3 are identical except, as mentioned before, for $PM_{2.5}$ in summer; $pm_{2.5}$aug is statistically significant in the winter months.

COMPARISON OF RESULTS FROM DEFAULT GAM AND GLM ANALYSES

The same 34 values from Table 1 were also computed using a GLM regression. The normalized differences were greater on the whole than between the old and new GAM analyses, with a median absolute difference of 0.15. The differences ranged from −0.18 to +0.37. The GLM coefficients were statistically smaller, with an average normalized difference of 0.081. The GLM results are very similar to the GAM approach, with no RRs differing by more than 0.02 and all but five differing by 0.01 or less. The same coefficients were statistically significant.

Only PM_{10}, $PM_{2.5}$, pm2.5aug, and o3ppbgt60 GLM analyses are shown in Table 2. The results are again essentially unchanged.

Table 3 presents RRs by season. Here the results are unchanged with two exceptions. Under GLM, NO_3 is no longer statistically significant for the summer months, and lag COH becomes statistically significant for the winter months.

RESULTS FOR DAILY NUMBER OF HOURS WITH O_3 GREATER THAN OR EQUAL TO 60 ppb

For nonaccidental mortality, o3ppbgt60 is statistically significant alone and paired with other pollutants. The inclusion of o3ppbgt60 in a model with other pollutants produces almost no change either in the coefficient estimates or the statistical significance of the other pollutants. In contrast, the inclusion of $PM_{2.5}$ or NO_3 with another variable (except for o3ppbgt60 and oz8hr) reduces the magnitude and statistical significance of the coefficient of the other variable. The o3ppbgt60 was statistically significant in the regression with cardiovascular mortality but not respiratory mortality. For the seasonal regressions in Table 3, o3ppbgt60 is statistically significant in summer but not other seasons.

COMPARISON OF *P* VALUES

Because confidence intervals have been used to compare the results from various studies, it is worthwhile to examine the relation between inferences derived directly from the F tests of one model versus another with the indirect inferences derived from *t* tests comparing the coefficients with their estimated standard errors. Table 4 shows a comparison of *P* values for these tests for the inclusion of

Table 1a. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Nonaccidental Causes with Pollutant Alone and Lagged[c]

	Same-Day Variables			1-Day Lagged Variables		
Pollutant	Original GAM	New GAM	New GLM	Original GAM	New GAM	New GLM
PM_{10}	0.080	0.078 (0.028,0.131)	0.083 (0.029,0.139)	−0.009	−0.011 (−0.055,0.035)	−0.012 (−0.059,0.038)
$PM_{2.5}$	0.093	0.092 (0.018,0.172)	0.080 (0.016,0.148)	−0.041	−0.042 (−0.101,0.020)	−0.043 (−0.102,0.021)
pm2.5aug	NA	0.083 (0.030,0.138)	0.094 (0.034,0.158)	NA	−0.002 (−0.076,0.077)	−0.004 (−0.129,0.138)
$PM_{10-2.5}$	0.024	0.023 (−0.040,0.091)	0.017 (−0.028,0.064)	−0.022	−0.023 (−0.080,0.038)	−0.029 (−0.100,0.049)
COH	0.026	0.022 (−0.010,0.054)	0.024 (−0.012,0.061)	0.047	0.043 (0.016,0.071)	0.045 (0.017,0.073)
NO_3	0.074	0.074 (0.025,0.124)	0.070 (0.024,0.117)	−0.024	−0.025 (−0.070,0.023)	−0.025 (−0.071,0.023)
SO_4	0.053	0.053 (0.007,0.101)	0.052 (0.007,0.100)	−0.017	−0.017 (−0.060,0.027)	−0.014 (−0.049,0.022)
NO_2	0.035	0.030 (−0.002,0.062)	0.026 (−0.002,0.055)	0.032	0.028 (0.003,0.055)	0.027 (0.002,0.052)
CO	0.017	0.015 (−0.005,0.035)	0.014 (−0.005,0.033)	0.042	0.039 (0.016,0.064)	0.039 (0.016,0.063)
oz8hr	0.030	0.031 (−0.003,0.066)	0.033 (−0.003,0.070)	-0.008	−0.006 (−0.023,0.012)	−0.005 (−0.022,0.011)
o3ppbgt60	NA	0.038 (0.014,0.063)	0.041 (0.015,0.067)	NA	0.005 (−0.019,0.029)	0.002 (−0.007,0.010)

[a] Relative risks calculated by exp(b* Δp) − 1, where b is the pollutant coefficient from the Poisson regression, and Δp = 50 for PM_{10}, corresponding to the increment used in the 1997 Criteria Document for particulate matter. For other pollutants, p, the increment was 50 * sd(p)/sd(pm10). For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28.

[b] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[c] GAMs include a 7 *df* smoothing spline for trend, 12 *df* smoothing spline for season, 3 *df* smoothing spline for minimum temperature and 2 *df* smoothing spline for maximum temperature. GLMs include 7 *df* natural spline for trend, 4 *df* natural spline for day of year, 3 *df* natural spline for minimum temperature and 3 *df* natural spline for maximum temperature. All models include 1 pollutant term entered linearly.

Table 1b. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Nonaccidental Causes for $PM_{2.5}$ and pm2.5aug with Other Pollutants

Pollutant	Original GAM	New GAM	New GLM	New GAM	New GLM
		$PM_{2.5}$ with Other Pollutant		**pm2.5aug with Other Pollutant**	
PM_{10}	0.126	0.125 (−0.055,0.340)	0.129 (−0.057,0.352)	0.101 (−0.054,0.283)	0.123 (−0.065,0.348)
$PM_{10-2.5}$	0.111	0.109 (0.021,0.205)	0.099 (0.019,0.185)	0.135 (0.026,0.255)	0.122 (0.023,0.229)
lag COH	0.108	0.107 (0.008,0.217)	0.090 (0.006,0.182)	0.101 (0.027,0.181)	0.105 (0.027,0.188)
NO_3	−0.001	−0.003 (−0.046,0.042)	−0.010 (−0.153,0.157)	0.045 (−0.069,0.173)	0.062 (−0.093,0.244)
SO_4	0.099	0.097 (0.008,0.194)	0.092 (0.007,0.184)	0.082 (−0.053,0.236)	0.087 (−0.055,0.251)
NO_2	0.119	0.119 (0.030,0.216)	0.108 (0.027,0.196)	0.103 (0.037,0.173)	0.111 (0.039,0.187)
lag CO	0.107	0.107 (0.024,0.195)	0.092 (0.021,0.167)	0.091 (0.028,0.159)	0.095 (0.029,0.166)
oz8hr	0.100	0.100 (0.024,0.181)	0.088 (0.021,0.159)	0.087 (0.032,0.145)	0.088 (0.033,0.147)
o3ppbgt60	NA	0.092 (0.008,0.182)	0.070 (0.006,0.139)	0.079 (0.024,0.137)	0.080 (0.024,0.139)
		Other Pollutant with $PM_{2.5}$		**Other Pollutant with pm2.5aug**	
PM_{10}	−0.038	−0.037 (−0.276,0.280)	−0.056 (−0.385,0.450)	−0.015 (−0.136,0.122)	−0.029 (−0.242,0.244)
$PM_{10-2.5}$	−0.030	−0.029 (−0.110,0.059)	−0.031 (−0.117,0.063)	−0.029 (−0.110,0.059)	−0.031 (−0.117,0.063)
lag COH	−0.009	−0.010 (−0.103,0.091)	−0.006 (−0.062,0.053)	−0.017 (−0.074,0.044)	−0.017 (−0.076,0.045)
NO_3	0.089	0.090 (−0.009,0.200)	0.086 (−0.009,0.190)	0.052 (−0.015,0.124)	0.041 (−0.012,0.096)
SO_4	0.001	0.002 (−0.059,0.066)	−0.007 (−0.234,0.287)	0.029 (−0.020,0.081)	0.028 (−0.019,0.076)
NO_2	−0.040	−0.041 (−0.114,0.037)	−0.042 (−0.116,0.038)	−0.022 (−0.067,0.026)	−0.027 (−0.083,0.032)
lag CO	−0.024	−0.025 (−0.082,0.036)	−0.020 (−0.066,0.029)	−0.011 (−0.054,0.034)	−0.012 (−0.058,0.036)
oz8hr	0.038	0.057 (−0.046,0.171)	0.054 (−0.044,0.162)	0.030 (−0.054,0.121)	0.018 (−0.033,0.072)
o3ppbgt60	NA	0.054 (0.005,0.107)	0.073 (0.006,0.144)	0.050 (0.005,0.097)	0.052 (0.005,0.101)

[a] Relative risks calculated by $\exp(b^* \Delta p) - 1$, where b is the pollutant coefficient from the Poisson regression, and $\Delta p = 50$ for PM_{10}, corresponding to the increment used in the 1997 Criteria Document for particulate matter. For other pollutants, p, the increment was 50 * sd(p)/sd(pm10). For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28.

[b] The confidence intervals are $c \pm 2^*s$, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

Table 2. Relative Risks[a] (95% Confidence Intervals[b]) for Daily Mortality from Cardiovascular and Respiratory Causes with Pollutant Alone[c]

Pollutant	Cardiovascular Mortality[d]			Respiratory Mortality[e]		
	Original GAM[f]	New GAM[f]	New GLM	Original GAM[f]	New GAM[f]	New GLM
PM_{10}	0.086	0.085 (0.006,0.170)	0.089 (0.013,0.170)	0.108	0.107 (−0.037,0.272)	0.108 (−0.034,0.271)
$PM_{2.5}$	0.073	0.072 (−0.046,0.205)	0.076 (−0.028,0.191)	0.133	0.133 (−0.110,0.442)	0.154 (−0.041,0.389)
pm2.5aug		0.103 (0.018,0.196)	0.104 (0.025,0.190)		0.102 (−0.033,0.255)	0.108 (−0.038,0.277)
$PM_{10-2.5}$	0.026	0.026 (−0.072,0.134)		0.156	0.156 (−0.065,0.428)	
COH	0.030	0.027 (−0.015,0.072)		0.068[g]	0.097 (0.020,0.180)	
NO_3	0.093	0.092 (0.017,0.173)		0.100	0.100 (−0.048,0.272)	
SO_4	0.040	0.040 (−0.028,0.113)		0.147	0.147 (0.011,0.301)	
NO_2	0.023	0.019 (−0.024,0.065)		0.072	0.069 (−0.004,0.147)	
CO	0.041	0.039 (0.001,0.077)		0.082	0.079 (0.012,0.151)	
oz8hr	0.024	0.026 (−0.023,0.078)		−0.043	−0.047 (−0.122,0.034)	
o3ppbgt60		0.043 (0.004,0.083)	0.047 (0.009,0.086)		−0.018 (−0.071,0.038)	−0.027 (−0.097,0.049)

[a] Relative risks calculated by exp(b* Δp) −1, where b is the pollutant coefficient from the Poisson regression, and Δp = 50 for PM_{10} and 50 * sd(p)/sd(pm10) for other pollutants, p. For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28. Single asterisks indicate statistical significance at the 0.05 level, double asterisks at the 0.01 level.

[b] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[c] GAMs include a 7 *df* smoothing spline for trend, 12 *df* smoothing spline for season, 3 *df* smoothing spline for minimum temperature and 2 *df* smoothing spline for maximum temperature. GLMs have 7 *df* natural spline for trend, 4 *df* natural spline for season, 3 *df* natural spline for minimum temperature and 3 *df* natural spline for maximum temperature.

[d] ICD categories 390–459.

[e] ICD categories 11, 35, 472-519, 710.0, 710.2, 710.4.

[f] Original GAM used epsilon and bf.epsilon of 10^{-4} with default # of iterations. New GAM used epsilons of 10^{-12}, and iterations and bf.iterations of 10^7.

[g] Error in the original article; this value should have been 0.100.

Table 3. Relative Risks (95% Confidence Intervals[a]) for Daily Mortality from Nonaccidental Causes by Season[b,c]

	Old GAM	New GAM	New GLM
Spring[d]			
PM_{10}	0.076	0.076 (−0.086,0.267)	0.100 (−0.045,0.267)
$PM_{2.5}$	0.071	0.071 (−0.132,0.322)	0.047 (−0.132,0.263)
pm2.5aug		0.065 (−0.124,0.293)	0.108 (−0.053,0.295)
lag COH	0.023	0.023 (−0.054,0.105)	0.027 (−0.045,0.105)
NO_3	0.072	0.072 (−0.046,0.205)	0.062 (−0.047,0.182)
SO_4	0.058	0.058 (−0.049,0.179)	0.067 (−0.027,0.170)
o3ppbgt60		0.018 (−0.183,0.269)	0.024 (−0.141,0.220)
Summer[e]			
PM_{10}	0.100	0.100 (−0.161,0.444)	0.088 (−0.170,0.427)
$PM_{2.5}$	−0.077	-0.076 (−0.326,0.266)	−0.087 (−0.453,0.524)
pm2.5aug		0.192 (−0.172,0.716)	0.173 (−0.202,0.723)
lag COH	0.128	0.128 (−0.085,0.391)	0.116 (−0.094,0.376)
NO_3	0.316	0.316 (0.000,0.731)	0.296 (−0.019,0.712)
SO_4	0.107	0.108 (0.002,0.224)	0.108 (−0.009,0.238)
o3ppbgt60	0.107	0.074 (0.015,0.137)	0.094 (0.022,0.171)
Fall[f]			
PM_{10}	0.071	0.071 (−0.096,0.269)	0.095 (−0.074,0.295)
$PM_{2.5}$	0.039	0.039 (−0.191,0.335)	0.043 (−0.250,0.448)
pm2.5aug		0.152 (−0.052,0.399)	0.191 (−0.027,0.459)
lag COH	0.082	0.081 (−0.018,0.191)	0.077 (−0.021,0.185)
NO_3	−0.131	−0.131 (−0.269,0.033)	−0.099 (−0.239,0.067)
SO_4	0.030	0.030 (−0.058,0.126)	0.036 (−0.060,0.142)
o3ppbgt60		0.036 (−0.031,0.108)	0.038 (−0.029,0.109)
Winter[g]			
PM_{10}	0.065	0.065 (−0.004,0.138)	0.057 (−0.006,0.124)
$PM_{2.5}$	0.046	0.046 (−0.054,0.155)	0.034 (−0.057,0.135)
pm2.5aug		0.075 (0.000,0.154)	0.064 (−0.004,0.137)
lag COH	0.036	0.036 (−0.001,0.074)	0.036 (0.004,0.070)
NO_3	0.073	0.073 (0.006,0.145)	0.062 (0.002,0.126)
SO_4	−0.004	-0.004 (−0.088,0.088)	−0.011 (−0.101,0.088)
o3ppbgt60		-0.299 (−0.892,3.534)	−0.337 (−0.917,4.301)

[a] The confidence intervals are c ± 2*s, where for GAM, s = abs(c)/sqrt(F), the F obtained from applying ANOVA to the models with and without the pollutant variable, and for GLM, s = standard error provided by S-Plus.

[b] Relative risks calculated by exp(b*Δp) − 1, where b is the pollutant coefficient from the GAM or GLM regression, and Δp = 50 for PM_{10} and 50 * sd(p)/sd(pm10) for other pollutants, p. For example, sd(pm2.5) = 13, sd(pm10) = 23, so for pm2.5, Δp = 50 * 13/23 = 28. Single asterisks indicate statistical significance at the 0.05 level, double asterisks at the 0.01 level.

[c] All models include a 7 *df* smoothing spline for trend, 12 *df* smoothing spline for season, 3 *df* smoothing spline for minimum temperature, and 2 *df* smoothing spline for maximum temperature.

[d] February, March, April.

[e] May, June, July.

[f] August, September, October.

[g] November, December, January.

Table 4. Comparison of *P* Values for Models Evaluating Daily Mortality from Nonaccidental Causes with Same-Day Pollutant Alone Derived from Different Methods

	t values[a]	F test[b]	Deviance[c]	Simulation[d]	Confidence Limit Lower 95%	Confidence Limit Upper 95%	Simulation Size
PM_{10}	0.0017	0.0017	0.0016	0.0015	0.0005	0.0039	2000
$PM_{2.5}$	0.0105	0.0122	0.0145	0.0063	0.0046	0.0087	6000
pm2.5aug	0.0016	0.0017	0.0015	0.0015	0.0005	0.0039	2000
COH	0.1557	0.1786	0.1729	0.1380	0.1180	0.1610	1000
NO_3	0.0016	0.0022	0.0021	0.0020	0.0012	0.0035	6000
SO_4	0.0195	0.0207	0.0196	0.0210	0.0014	0.0320	1000
NO_2	0.0391	0.0599	0.0571	0.0417	0.0351	0.0495	3000
CO	0.1127	0.1410	0.1365	0.1130	0.0990	0.1280	2000
O_3	0.0389	0.0654	0.0624	0.0390	0.0314	0.0485	2000
$PM_{10-2.5}$	0.4912	0.4682	0.4783	0.4990	0.4680	0.5300	1000
o3ppbgt60	0.0005	0.0013	0.0012	0.0008	0.0003	0.0022	4000

[a] Ratio of estimated coefficient to standard error estimate from Table A3 in Fairley 1999.

[b] F test from ANOVA comparing GAMs with and without the pollutant variable.

[c] Assuming that the deviance has a chi-squared distribution with 1 *df.*

[d] Simulations of the null distribution of the pollutant coefficient. Confidence limits based on the binomial distribution.

same-day, single pollutant variables (the same-day, New GAM column of Table 1a). Also included are *P* values using the simple assumption that the change in deviance has a chi-squared distribution with 1 *df.* In addition, the simulation results are shown along with simulation size and confidence intervals based on the binomial.

Table 4 shows that the *t* tests gave results similar to the simulations. The F tests and chi-squared tests gave effectively the same results but appear conservative. In the case of same-day NO_2 and oz8hr, this small difference between the two tests meant the difference between significance and nonsignificance. In the case of the $PM_{2.5}$ coefficient, all the tests appear conservative: the upper confidence bound for the *P* value from the simulation is 0.0087, less than any of the other *P* value estimates. The more extensive simulation, where the entire model-fitting process was simulated, resulted in 8 of 1000 runs in which the simulated $PM_{2.5}$ coefficient was greater than the observed $PM_{2.5}$ coefficient in absolute value. The extensive simulation also suggests that the $PM_{2.5}$ *P* value is, if anything, overestimated using the F test, deviance or *t* value approaches. In the case of the pm2.5aug coefficient, all 4 methods gave similar *P* values, between 0.0015 and 0.0017. The more extensive simulation yielded 2 values out of 1100 greater in absolute value than the observed for an approximate *P* value of 0.0018, very similar to the other results.

SENSITIVITY ANALYSIS: USE OF SINGLE TREND/SEASON SMOOTH

An analysis was performed to determine the effect of using a single smooth for season/trend instead of two. Again, smoothing splines were used and AIC was used as the stopping rule. The minimum AIC of 3104.6 was achieved with 75 *df*, compared with an AIC of 3079.8 using 7 *df* for trend and 12 *df* for season. When temperature terms and $PM_{2.5}$ were added to this model, the estimated $PM_{2.5}$ RR increase dropped from 0.092 to 0.071 and was no longer statistically significant. The estimated RR for pm2.5aug dropped from 0.083 to 0.065 but remained statistically significant.

Figure 2 shows the crude partial ACFs for nonaccidental mortality, the residuals from the 7 *df* trend/12 *df* season fit, and the residuals from the 75 *df* trend/season fit, respectively. Figure 2A shows highly significant autocorrelation in the daily nonaccidental mortality data, with a first order coefficient of about 0.20. Figure 2B shows a first order ACF of just over 0.04, borderline statistically significant, but otherwise an apparently random pattern of positive and negative ACF. Figure 2C shows almost all negative coefficients. The 7 *df* trend/12 *df* season fit had a deviance of 3038.5, similar to the deviance from fitting a single series with 37 *df* (deviance 3039.4), close to the 4 *df* per year suggested by Joel Schwartz at the recent US Environmental Protection Agency (EPA) Workshop on GAM-

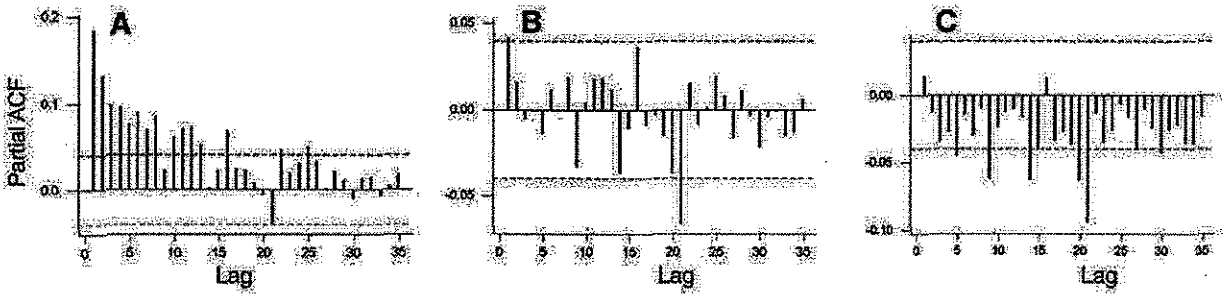

Figure 2. Partial autocorrelation coefficients (partial ACFs). *A.* Partial ACF of nonaccidental mortality. *B.* Partial ACF of residual *df*: 12 season, 7 time. *C.* Partial ACF of residuals from fitting time with 75 *df*.

Related Statistical Issues in PM Epidemiology, November 4–6, 2002 (Research Triangle Park NC).

DISCUSSION

The basic conclusions of the previous analysis were unchanged: Considered individually, every criteria pollutant (either same-day or lagged by one day) was significantly related to daily mortality. $PM_{2.5}$ and NO_3 continued to be significantly related to mortality in conjunction with the other pollutants.

The fact that the GAM results were virtually identical to the GLM results underscores the conclusion reached in the EPA Particulate Matter Criteria Document (1996), which considered a range of studies analyzing the effects of model choice in short-term mortality studies:

> Differences in model specification may produce important differences in estimates of effects. The general concordance of PM effects estimates, particularly in the analysis of short-term mortality studies, is a consequence of certain appropriate choices in modeling strategy that most investigators have adopted using several different types of standardized models (GLM, LOESS, etc.).

The inferences used in this analysis may be conservative. Although there is no gold standard, the simulations should be the closest to a gold standard because they do not rely on an asymptotic distribution. The only limitation is that the simulations did not account for the uncertainty in the underlying parameters (ie, it did not include simulation of fitting the null model). However, the more extensive simulation of the model-building process of selecting the *df* for trend, season and weather using AIC gave similar results to the simpler simulation. Using the ANOVA feature of S-Plus or assuming the change in the deviance has a chi-square distribution that gave more conservative results than those from the simulations. Inferences using the GLM-based standard errors agreed closely with the simulations in most cases, suggesting that they do not always produce conservative results. This finding suggests that the GLM-based standard errors may be acceptable for drawing inferences. Perhaps these could be incorporated as defaults or options in statistics packages that offer GAM.

The one modeling choice that does make a substantial difference is using two time smooths—one for trend and one for season—rather than a single time smooth. Using a single time smooth results in considerably lower PM coefficient estimates and lower statistical significance. However, the single time smooth model that minimized the AIC had 75 *df*, which is more than 9 per year from 1989 to 1996. This, according to the discussion at the GAM workshop, may constitute overfitting. In fact, Figure 2C shows evidence of overfitting in that almost all the autocorrelation coefficients are negative. Also, the AIC for the single time-term model is 3104.6, considerably poorer than the AIC for the 7 *df* trend/12 *df* season model, 3079.8.

Parsimony is an established scientific rule. By Ockham's Razor, if there is a choice between similar-fitting models, the one with fewer terms should be favored. And a previously suspected cause (in this case PM) should be favored over an unknown cause (for which time is a surrogate). In practical terms, the question to address is how much of the short-term mortality variation to attribute to PM vis-à-vis some other covariate represented by additional degrees of freedom in the time smooth. The fitting process begins with a fit of time/season only. This tends to add more degrees of freedom for these terms because every day is included, not the 1 in 6 days where $PM_{2.5}$ data are available. Secondly, AIC is a liberal criterion, which includes variables that are not statistically significant. Thus, the modeling approach allows a reasonable chance for the other covariate to demonstrate its existence. Yet pm2.5aug remains statistically significant, even in the 75 *df* trend/season model, and the estimated PM effect drops by about a quarter, rather than disappearing entirely. Thus, the two-smooths model with fewer parameters and higher PM coefficients appears preferable according to the Ockham's Razor.

The new ozone variable, o3ppbgt60, was found to be statistically significant even in a regression jointly with $PM_{2.5}$ or NO_3. Two other thresholds were tried—40 ppb and 80 ppb—both of which produced statistically significant results in one-pollutant models but somewhat smaller RRs. Although o3ppbgt60 is quite skewed, the results do not appear to be a function of a few outliers; a regression using log(o3ppbgt60) also yielded statistically significant results.

REFERENCES*

Chambers JM and Hastie TJ. Statistical Models in S. Pacific Grove: Wadsworth & Brooks/Cole. (Chapter 7.4.4, page 303)

Cox DR and Hinkley DV 1974. Theoretical statistics. London: Chapman and Hall.

Dominici F, Zeger SL, and Samet JM. 2000. Combining evidences on air pollution and daily mortality from the largest 20 US cities: A hierarchical modeling strategy. Royal Statistical Society, Series A. 163:263-302

Dominici F, McDermott A, Zeger SL, Samet JM. 2002. On the use of Generalized Additive Models in time series of air pollution and health. Am J Epidemiol 156:193-203

Environmental Protection Agency (US). 1996. Air quality criteria for particulate matter, Vol 3.pp 13-53:13-54, EPA/600/P-95/001cF, April 1996.

Fairley D. 1999. Daily mortality and air pollution in Santa Clara County, California: 1989–1996. Environ Health Perspect 107:637–641.

Kelsall JE, Samet JM, Zeger SL, and Xu J. Air pollution and daily mortality in Philadelphia, 1974-1988. Am J Epidemiol 146:750-762.

Moolgavkar S. 2000. Air pollution and daily mortality in three U.S. counties. Environ Health Perspect 108:777-784.

Saez M, Ballester F, Barelo MA, Perez-Hoyos S, Bellido J, Tenias JM, Ocaña R, Figueiras A, Arribas F, Aragones N, Tobias A, Cirera L, Cañada A. "A combined analysis of the short-term effects of photochemical air pollutants on mortality within the EMECAM Project," Environ Health Perspect. 110:221-228 (2002).

Samet JM, Dominici F, Curriero F, Coursac I, and Zeger SL. 2000. Particulate air pollution and mortality: Findings from 20 U.S. cities. New England Journal of Medicine, 343:1742-1757.

Schwartz J. 1994. Nonparametric smoothing in the analysis of air pollution and respiratory illness. Can J Stat 22:471-488.

* Bold type identifies publication containing the original analyses revised in this short communication report.

ABBREVIATIONS AND OTHER TERMS

ACF	autocorrelation coefficient
AIC	Akaike information criterion
ANOVA	analysis of variance
CL	confidence limit
CO	carbon monoxide
COH	coefficient of haze
df	degree of freedom
EPA	Environmental Protection Agency (US)
GAM	generalized additive model
GLM	generalized linear model
ICD	*International Classification of Diseases*
NO_2	nitrogen dioxide
NO_3	nitrate
ns	natural spline
O_3	ozone
o3ppbgt60	ozone ppb-hours greater than 60 ppb summed for each day
oz8hr	8-hour maximum ozone
PM	particulate matter
PM_{10}	particulate matter < 10 µm in diameter
$PM_{2.5}$	particulate matter < 2.5 µm in diameter
pm2.5aug	$PM_{2.5}$ measurements augmented with $PM_{2.5}$ predicted from COH and PM_{10}
ppb	parts per billion
R^2	multivariate coefficient of variation
RR	relative risk
SO_4	sulfate