Journal of Exposure Analysis and Environmental Epidemiology **(2001) 11**
© 2001 Nature Publishing Group All rights reserved 1053-4245/01/$17.00

www.nature.com/jea

Epidemiological studies of acute ozone exposures and mortality

GEORGE D. THURSTON AND KAZUHIKO ITO

Department of Environmental Medicine, New York University School of Medicine, Nelson Institute of Environmental Medicine, Tuxedo, New York

Many, but not all, observational epidemiological studies of ozone (O_3) air pollution have yielded significant associations between variations in daily ambient concentrations of this pollutant and a wide range of adverse health outcomes. We evaluate some past epidemiological studies that have assessed the short-term association of O_3 with mortality, and investigate one possible reason for variations in their O_3 effect estimate, i.e., differences in their approaches to the modeling of weather influences on mortality. For all of the total mortality–air pollution time-series studies considered, the combined analysis yielded a relative risk, RR = 1.036 per 100-ppb increase in daily 1-h maximum O_3 (95% CI: 1.023–1.050). However, the subset of studies that specified the nonlinear nature of the temperature–mortality association yielded a combined estimate of RR = 1.056 per 100 ppb (95% CI: 1.032–1.081). This indicates that past time-series studies using linear temperature–mortality specifications have underpredicted the premature mortality effects of O_3 air pollution. For Detroit, MI, an illustrative analysis of daily total mortality during 1986–1990 also indicated that the model weather specification choice can influence the O_3 health effects estimate. Results were intercompared for alternative weather specifications. Nonlinear specifications of temperature and relative humidity (RH) yielded lower intercorrelations with the O_3 coefficient, and larger O_3 RR estimates, than a base model employing a simple linear spline of hot and cold temperature. We conclude that, unlike for particulate matter (PM) mass, the mortality effect estimates derived by time-series analyses for O_3 can be sensitive to the way that weather is addressed in the model. The same may well also be true for other pollutants with largely temperature-dependent formation mechanisms, such as secondary aerosols. Generally, we find that the O_3–mortality effect estimate increases in size and statistical significance when the nonlinearity and the humidity interaction of the temperature–health effect association are incorporated into the model weather specification. We recommend that a minimization of the intercorrelations of model coefficients be considered (along with other critical factors such as goodness of fit, autocorrelation, and overdispersion) when specifying such a model, especially when individual coefficients are to be interpreted for risk estimation. *Journal of Exposure Analysis and Environmental Epidemiology* (2001) **11.**

Keywords: *modeling, mortality, ozone (O_3), particulate matter (PM), weather.*

Introduction

Observational epidemiological studies provide especially relevant evidence as to whether ambient environmental factors, such as air pollution, can adversely affect the general public. This is because these studies consider the "real-world" experiences of human populations as they are exposed to pollution in natural settings. Such observational epidemiological studies are termed "ecological" if they consider aggregate data for large groups of people (e.g., group counts of deaths per day or per year), rather than individuals. These studies often follow populations of people in a defined geographical area (e.g., a city) as they undergo varying everyday exposures to pollution over time, or from one place to another, and then statistically evaluate the variations in the total number of adverse health impacts that occur in these populations (e.g., city-wide respiratory hospital admissions counts) when higher (versus lower) concentrations of pollution are experienced.

In the case of ozone (O_3) air pollution, many, but not all, recent observational epidemiologic studies have yielded significant associations between variations in daily ambient concentrations of this pollutant and a wide range of adverse health outcomes. This paper primarily focuses on a quantitative evaluation of some past epidemiological studies that have assessed the short-term association of O_3 with mortality, and investigates one possible reason for variations in their O_3 effect estimate, i.e., differences in their approaches to the modeling of weather influences on mortality. Thus, this paper considers many of the recent observational epidemiological studies that have evaluated the possible effects of O_3 air pollution on human mortality to quantitatively evaluate the possible role of modeling choices regarding weather specification in the variability of past evaluations of O_3 and mortality. In addition, an illustrative example for Detroit, MI presents how various weather specifications affect the model intercorrelations and the O_3 effect estimate for mortality in that city. Collectively, these evaluations aim to give new insights into the extent to which O_3–mortality effect estimates can be influenced

1. Address all correspondence to: Dr. George D. Thurston, Department of Environmental Medicine, New York University School of Medicine, Nelson Institute of Environmental Medicine, Tuxedo, NY.
Received 4 March 2001.

by the choice of weather specification in time-series models.

Methodological issues in O_3 time-series modeling

The studies considered here are longitudinal time-series statistical analyses, or those studies that follow a single population over time. Although being semiecological in nature, as they do not estimate individual exposures for each person in the population, such time-series analyses present numerous statistical advantages over other types of studies. These include the fact that, by design, they obviate the need to control for individual-level confounding factors (e.g., education level, income, smoking, etc. in the population) that can confound other study designs (such as cross-sectional studies, which compare effects across different populations). This is the case because such intrinsic population characteristics are relatively constant from day-to-day. Thus, the population acts as its own "control" in such a time-series model, since a single population's health effects incidence on higher pollution days are being compared to that same population's incidence on lower pollution days.

Epidemiological studies also do not share the practical and ethical limitations of controlled human exposure studies, which are usually not able to study the most sensitive populations (e.g., those with severe preexisting disease) and must consider pollution concentrations that are expected to only result in relatively mild responses. Instead, time-series epidemiological studies consider a community's entire population, and the entire spectrum of ambient exposures experienced by that population. Thus, in addition to being very relevant to the general public, epidemiological studies are also extremely useful in capturing the most serious effects among the most sensitive members of the general population, and can have great statistical power when they consider extremely large populations (e.g., entire metropolitan areas) over multiple years.

Despite the above-noted strengths of time-series observational epidemiology, the application of such methods also present important statistical challenges of their own, as discussed in detail by Thurston and Kinney (1995). In particular, shared long-term cycles in the health outcome (e.g., mortality) and the pollutants being analyzed can, if not adequately addressed, cause misleading associations (e.g., due to shared winter to summer seasonal trends), and yield biased pollutant risk estimates. These cycles can also cause statistical autocorrelation and/or overdispersion in the model residuals that, if unaddressed, can bias pollutant significance tests. This problem is especially relevant to the time-series analysis of year-round and multiyear records of daily population counts of human morbidity and mortality, as these health outcome daily series usually exhibit strong seasonal variations over time. Moreover, it is difficult for such correlation-based models to separate the individual influences of other environmental factors that covary over time with the pollutant of interest in the study locale, potentially biasing the pollutant effect size estimate provided by such models. For example, O_3 is usually moderately to strongly associated with ambient temperature, representing a potential confounder to the elicitation of O_3 associations with morbidity and premature mortality. Thus, if not appropriately addressed, the influence of seasonal variations and temperature impacts on the incidence of health outcomes can confound such time-series models' evaluation of O_3's effects on human health.

Fortunately, a variety of time-series statistical modeling options are now available to analyze the short-wave "signal" of interest separately from the long-wave "noise" superimposed on day-to-day variations, including: Fourier techniques (i.e., fitting sine/cosine waves to the data), high pass prefiltering, autoregression methods, the fitting of a smooth of mortality over time, or the use of time-dependent (e.g., monthly) dummy variables in regressions. The model results do not seem to be especially sensitive to which of these approaches is employed (e.g., see Lipfert, 1994; Kinney et al., 1995), but one of these methods needs to be applied to avoid confounding by seasonality.

Similarly, how the known acute effects of temperature extremes on human morbidity and mortality are handled in a model can also affect the pollution–health effects association estimates. This is especially true for O_3, which tends to experience peak concentrations on high temperature days, when many O_3 precursors are emitted at higher rates (e.g., *via* the greater vaporization of hydrocarbons) and their conversion to O_3 is most rapid. Also, as noted in the U.S. EPA Ozone Criteria Document: "Ambient air temperature often exhibits a moderate to high correlation over time with O_3 in acute epidemiology studies due, in part, to the dependence of O_3 formation rate on light intensity. Among the studies reviewed..., correlations ranging from 0.06 to 0.90 (mean=0.51) have been reported" (U.S. EPA, 1996). Some older O_3–mortality studies ignored temperature, which may have led to an overestimation of O_3 effects. Conversely, modeling the associations between temperature and mortality as a linear relationship can cause the underestimation of the O_3 effect, due to the serial intercorrelation of O_3 and temperature over time.

Although temperature and O_3 are moderately to highly correlated over time, differences in their respective relationships with health outcomes allows their effects to be disentangled *via* advanced statistical methods. Whereas the temperature–health effect dose–response relationship is "U-shaped," with increased adverse effects at either extremely high or extremely low temperatures, the O_3–health effects relationship is more linear, with adverse O_3 effects increasing as concentration increases. There is also

an interaction between the effects of temperature and relative humidity (RH) on high temperature days, which also needs to be incorporated. Thus, if these factors are specified in the model, their respective effect coefficients (e.g., that between the "U-shaped" temperature specification and the monotonic O_3 term) should be less intercorrelated than their raw variables, allowing these model terms to be better discriminated (i.e., having lower intercorrelations between their coefficients), despite O_3 and raw temperature's moderate to high linear intercorrelation. The effect of weather specification on the O_3 effect estimate is an as yet undefined parameter, and was the focus of this work.

Another concern that has been raised regarding aggregate time-series analyses is that the collection of individual exposure data in such large populations is not practical, so central site monitor concentration data (either from a single site, or from the average of multiple sites) are usually used to provide indices of day-to-day variations in population exposures. However, the use of outdoor central site air pollution monitoring data as an index of average population exposure is apparently not as large a problem as once thought. Recent work by Mage and Buckley (1995) and Mage et al. (1999) indicates that, although individual exposures are not always well characterized by a central site monitor's data, such central site data are highly correlated with the average of individual exposures in a population, which is what is required for the purposes of time-series analyses of aggregate population counts of health effects. Moreover, because the purpose of such time-series epidemiologic analyses is usually to help set ambient standards that will ultimately be monitored at a central site, the use of central site data in the original epidemiological studies simplifies the standard setting process (i.e., thereby avoiding any extrapolation between individual exposures measured during research, and central site monitor concentrations to be employed in standards-compliance monitoring). Thus, any differences between personal and central site O_3 concentrations should not usually present a serious problem in such time-series studies of O_3 health effects as those considered here.

However, random errors that greatly reduce the correlation between measured and actual exposure (e.g., due to poor spatial correlations of a pollutant) will tend to reduce the significance of the pollutant effect and bias pollutant effect regression estimates toward zero (e.g., see Caroll et al., 1995). This may, in turn, reduce such time-series methods' ability to detect the effects of some pollutants. The use of multiple-site averages for a city or region in such aggregate epidemiological studies should reduce this problem, and the problem is less severe for more spatially correlated pollutants (e.g., see Ito et al., 1995, 2001). In the case of O_3 air pollution, which is highly correlated spatially across a metropolitan area, the chief concern is actually that the pollutant may not permeate equally in all buildings, for example, in air-conditioned versus non–air-conditioned buildings (Wechsler et al., 1989). In warm climates (e.g., the U.S. Southwest), the percentage of homes with air conditioning can exceed 90% (U.S. Bureau of the Census, 1983), which may reduce the ability of reactive pollutants like O_3 to reach residents for much of the day. This may, in turn, diminish the overall population adverse health implications of outdoor O_3 air pollution in those cities versus cities without extensive air conditioning. As a result, it is likely that, in cities with warmer climes that have extensive use of air conditioning (e.g., Houston, TX), the estimated O_3 health effect on exposed individuals (i.e., the RR per amount of pollution) may be underestimated compared to effects estimated in communities having cooler climates, and more limited use of air conditioning. The influence of such potential effect modifiers needs further investigation, but requires multicity analyses that is beyond the scope of the research presented here.

Ozone and premature mortality

Limitations of Early Studies

Whereas recent studies of mortality and acute O_3 exposures have usually attempted to address potential confounders such as seasonal, temperature, and other pollutants' influences in their analyses, older studies of the possible association of O_3 with human mortality were usually flawed in these regards. Unlike hospital admissions and ED visit studies, where daily visits are usually stratified by cause, most studies of air pollution and mortality have examined only total daily counts. This is possibly because of a lack of reliably recorded by-cause data, and also because the small numbers of daily respiratory deaths limit the statistical power of such studies to detect by-cause mortality effects, even if present. Most studies published in the 1950s and 1960s considered total daily mortality in Los Angeles, CA. Many of these early studies did not recognize and attempt to address seasonality in the data series, including the California Department of Public Health (1955) study, which was also weakened by the qualitative treatment of the air pollution data. Newer studies (especially those conducted during the past decade) have usually better addressed the various potential confounding factors noted above (to greater or lesser extents), but have still provided somewhat varying results. As discussed further below, most of those have addressed long-wave confounding satisfactorily, but not all have looked at the role of copollutants, considered the same O_3 averaging times (e.g., 1-h daily maximum value vs. 8-h or 24-h average O_3 values), or modeled weather influences similarly, complicating comparisons among studies.

Illustrative Example of Weather Specification Effects on O_3–Mortality Effect Estimates

Unlike the modeling of mortality–particulate matter (PM) air pollution associations, which have been indicated to be relatively insensitive to the weather–mortality modeling approach (e.g., see Pope and Kalkstein, 1996; Samet et al., 1997, 1998), the choice of time-series weather modeling approach may have a large effect on the O_3–mortality effect estimate. This is because elevated O_3 concentrations are much more highly correlated with high temperatures than PM mass (e.g., PM_{10}), and temperature has its own adverse effects on human health, especially when present in tandem with high humidity conditions (e.g., see Ellis, 1972). Thus, if both temperature and O_3 are simultaneously entered into a time-series as raw variables, they will compete with each other for the same mortality variations, resulting in unstable effect estimates.

However, despite the high correlation with one another, the differing relationships of weather and O_3 with mortality, respectively, can be exploited to separate their individual effects. As discussed previously, the relationship between mortality and each of these two variables is quite different, with mortality having a highly nonlinear (U-shaped) relationship with temperature (higher daily mortality at both hot or cold temperature extremes), whereas O_3 has a more linear relationship with daily mortality. This is shown in Figure 1 for the case of Detroit, MI, where both temperature and RH have a quite strong nonlinear relationship with mortality, along with an interaction between high temperature and high RH effects on mortality. In contrast to the nonlinear weather–mortality relationship, mortality in Detroit is shown in Figure 2 to exhibit a more monotonic relationship with O_3 concentration.

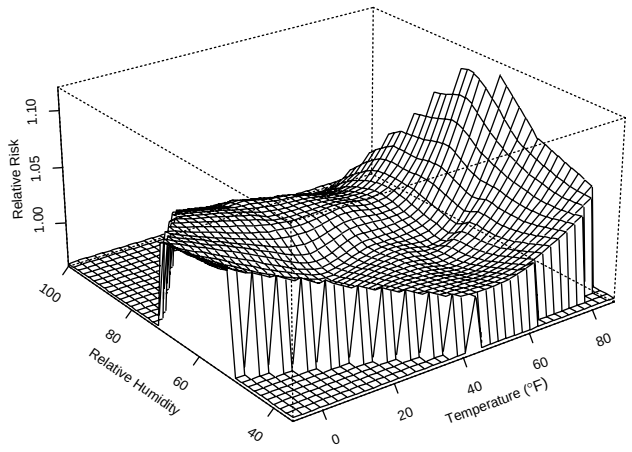

Figure 1. LOESS 3-D smooth surface of total mortality RR versus daily temperature and RH (adjusted for season and day-of-week) in Detroit, MI.

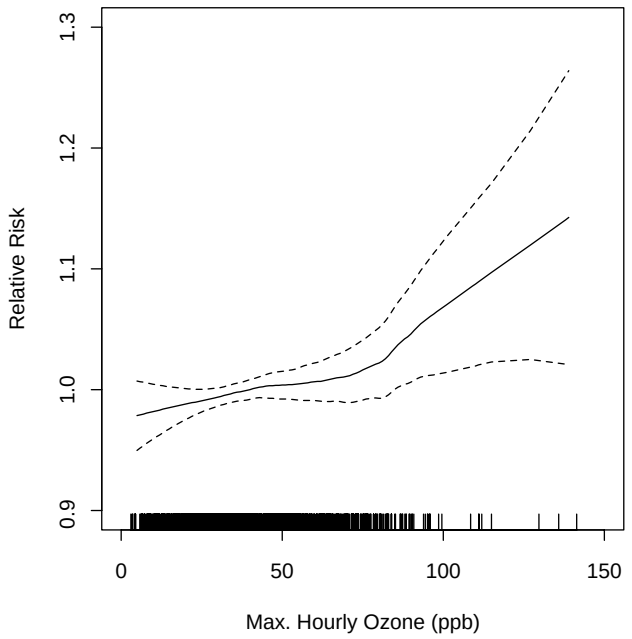

Figure 2. Total mortality RR versus O_3 concentration (adjusted for weather, season, day-of-week, and PM) in Detroit, MI.

As an illustrative example of the potential problems that can result from weather–O_3 intercorrelations, we modeled total daily mortality during the 1986–1990 period in Detroit (Wayne County), MI. The basic model was a Poisson generalized additive model (GAM) that accounted for overdispersion and included seasonal controls, day of week variables, and PM_{10} concentrations. As a semiparametric modeling technique, GAM allows for nonlinear variables to be included in regression analyses (Hastie and Tibshirani, 1990). Here, we applied a GAM using local weighted regression smoothing techniques, LOESS (Cleveland, 1979) to fit the long-wave mortality trends with a span that provided a smoothing periodicity of 30 days, which was selected to minimize overdispersion in the mortality series, while at the same time minimizing the induction of negative autocorrelations in model residuals.

To the basic GAM model, several reasonable alternative weather model specifications were added, along with daily maximum hourly O_3. These different weather models have differing implications to the O_3 RR effect estimate. As seen in Table 1, the simplest weather model (which assumes a linear relationship between heat and total daily mortality) yields the lowest O_3 effect estimate and (especially) the lowest t-statistic, and the highest intercorrelation with the O_3 coefficient. In this table, the risk estimates are presented as the percent excess risk (ER) per 100-ppb increment in 1-h maximum O_3, where $ER=(RR-1)\times 100$. One reason for the low significance level for O_3 in this first model may be an inflated standard

Table 1. Ozone excess risk (ER) estimates, their *t*-ratios, and the correlation of the O_3 ER beta estimates for models with different weather model specifications in Detroit, MI (Wayne County) during 1986–1990 ($n=1121$), after controlling for season, day of week, and PM_{10}.

Weather model	% ER per 100 ppb O_3	O_3 coefficient *t*-ratio[a]	Correlation of betas between O_3 and hot weather term
Model 1. Piecewise linear temperature terms for hot (same-day >=60°F) and cold (2-day lag <60°F) effects	7.4	1.95	−0.56
Model 2. Separate LOESS smooths of temperature and RH for both same-day lag (for heat) and 2-day lag (for cold) effects	10.6	2.87	−0.23
Model 3. Separate LOESS smooths of temperature and dew point for both same-day lag (for heat) and 2-day lag (for cold) effects	10.7	3.06	−0.22
Model 4. LOESS smooth interaction of temperature and humidity for same day lag and 2-day lag effects	13.1	4.23	−0.16

[a]After upward adjustment of coefficient standard error for constant overdispersion.

error of the coefficient, which is a common problem when correlated variables are entered into a regression simultaneously. In Model 2, as the weather term is specified to model the nonlinear nature of the temperature–mortality relationship, the O_3 effect estimate, and (most notably) its significance, is increased, as the intercorrelation of the betas decrease (to $r_\beta=-0.23$).

In Model 3, it is seen that the specification of a dewpoint locally estimated smooth (LOESS) term (span=0.5) gives slight further improvement vis-à-vis its intercorrelation with O_3 over the use of raw RH. This is likely because a very high dewpoint day (e.g., 69°F dewpoint) can result from either a comfortable 69-degree temperature on a rainy day (with RH=100%), or from a very uncomfortable 90-degree plus day with a high RH for such a hot day (e.g., RH>50%). Thus, a given dewpoint can result from very different types of days, and the use of dewpoint therefore might not uniquely identify the high temperature–humidity days in Model 3 as uniquely as does the weather specification in Model 4. The temperature and dewpoint LOESS model (span=0.5) also retains a moderate intercollinearity problem with O_3 in the mortality model ($r=-0.22$).

Finally, the weather model is specified to incorporate both the temperature nonlinearity and the temperature–RH interdependence of the hot weather–mortality relationship (Model 4), as depicted in Figure 1. The weather term now fits not only the nonlinearity of the temperature–mortality association, but also the greatly heightened impacts on days having both high RH and high temperature (LOESS span=0.2). This 3-D weather surface is now also less intercorrelated with the more linear O_3–mortality association ($r=-0.16$) than other models considered, and a larger and more significant O_3 association is indicated by this model (ER=13.1% per 100 ppb O_3, *t*-statistic=4.23). It is also notable, by way of comparison, that this Model 4 estimate of O_3–mortality ER is somewhat smaller than the effect estimate derived when no weather is included, a model that avoids the O_3–temperature intercorrelation altogether (O_3 ER=13.7%, $t=4.80$).

Model fit was also examined using the both the Akaike Information Criteria (AIC) (Akaike, 1974) and the *r*-squared measure of model fit (the latter derived using a similarly specified log-linear OLS model). The AIC values indicated that the overall model fit with the various specifications were similar, with the smallest AIC value actually being found for the simple linear model specification (i.e., Model 1). This result is likely due to the fact that, given reasonably similar overall fits of mortality, the Akaike Criterion's penalty for added degrees of freedom caused the more complex weather models to get weaker AIC values. The *r*-squared values (which measure overall goodness of fit without a penalty for added degrees of freedom) also indicated similar fits across specifications (ranging from 19.8% for the linear weather model, up to 20.9% for a nonlinear temperature model), indicating the simple linear model to now be indicated as the worst fitting model. In addition, in this case, the model with the lowest (i.e., best) AIC, also gave the least independent estimate of the individual model coefficients, due to high intercorrelations of the temperature and O_3 coefficients in Model 1. Overall, the goodness of fit did not vary appreciably across models, so the "best fit" conclusions varied depending on the measure of fit examined.

We also investigated the influence of co-pollutant lag choice on the ozone effect estimate. In the Table 1 models, we used the maximum individual effect lag for each pollutant in the combined pollutant regressions (i.e., 1-day for both in this city), but there was moderate intercorrelation found between the pollutant coefficients in these models ($r=-0.45$). We therefore also applied these four models for the case where the PM_{10} was for lag 0, keeping O_3 at lag 1. The trend across models was similar (i.e., weaker ozone effects in Model 1 than in other models), but there was a lowering of the intercorrelation between the pollutant

coefficients, and a lowering of the ozone effect estimate and its significance (to O_3 ER=8.1%, t=2.75 for Model 4). This indicated that co-pollutant intercorrelations, though not as high as with weather, also affected pollutant effect estimates when the two pollutants were entered with the same lag.

Although not the focus of this investigation, it is interesting to note that the ER effect estimate for 100 $\mu g/m^3$ PM_{10} in Model 1 (ER=6.5%, t=1.85) was not nearly so correlated with the weather term ($r=-0.13$) as was O_3 ($r=-0.56$) in this model. This is consistent with earlier analyses of this question for PM, as discussed previously. In Model 4, the PM effect estimate (not shown in Table 1) rises to only ER=7.9% (t=2.16). Thus, it seems clear, in this city, that the O_3 effect estimate is much more sensitive to the choice of weather specification than is the PM_{10} effect estimate.

Weather Specification Influences on Some Published O_3–Mortality Effect Estimates

We also have sought to evaluate whether the apparent influence of a model's weather (i.e., temperature and RH) specification on the resultant mortality effects estimate for O_3 is responsible for much of the variability in the published values of the O_3 RR. To test this, we present the results from the various recently published O_3–mortality studies in Figure 3 as a function of the weather modeling approach (in terms of effects per 100 ppb of daily maximum 1-h average O_3 concentration). All of these results are for a model specification where a PM concentration index has also been included, to reduce the chances that the reported O_3 RR incorporates the effects of another copollutant. A concentration of 100 ppb O_3 as a daily 1-h maximum was chosen for these calculations to provide consistency across all study results, and because 100 ppb is on the order of the difference between the annual average concentration and the daily maximum 1-h maximum concentration experienced in most urban areas. When the analysts used O_3 data averaged over another period (e.g., 8- or 24-h averages), this is noted in the figure. Also, an adjustment was made to convert the reported RR in terms of a 1-h maximum averaging period, based on the ratio of the mean concentrations experienced for each of the respective averaging times in that study area. If this concentration information was unavailable from a specific study, a conversion of 2.5 was used to convert 24-h average RRs, and a conversion of 1.33 was used to convert 8-h average RRs, based on past experience (e.g., see Schwartz, 1997). Thus, Figure 3 presents relative risks (RRs) for the various studies so that they are directly intercomparable, and in terms that should be more intuitively interpretable than other candidate pollution increments (e.g., for the difference between the 25th and the 75th quartiles), because the 100-ppb RRs provide an index of the increased risk associated with a typical high O_3 day versus an average day.

The results depicted in Figure 3 are consistent with the hypothesis that the choice of temperature specification employed in a total daily mortality time-series model can have a substantial influence on the estimated O_3 effect size. As would be expected, those studies that employed a linear temperature specification for year-round data (left-most group of studies) had the lowest random effects model pooled O_3 RR estimate (pooled RR=1.026, 95% CI: 1.016–1.036). To obtain this combined estimate, we used a two-stage random-effects model approach, as

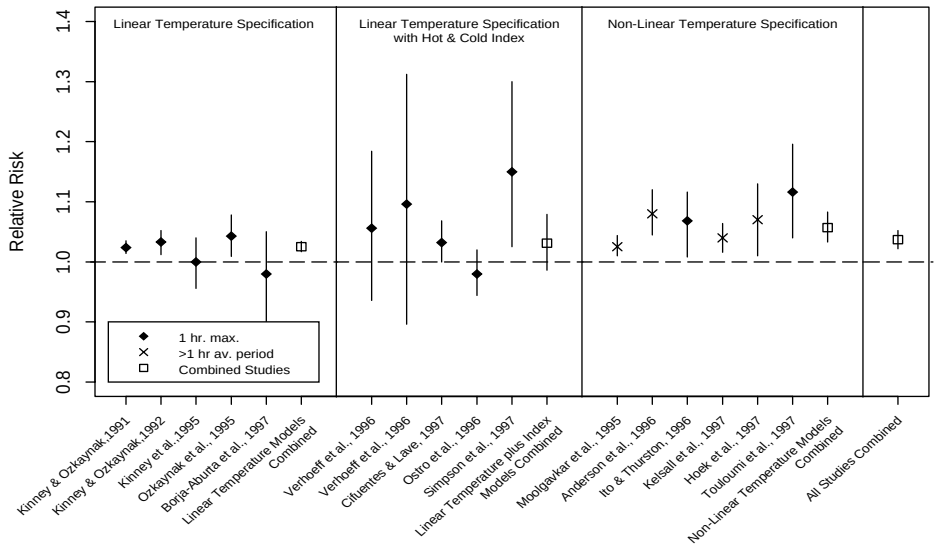

Figure 3. Reported RRs of mortality associated with a 100-ppb increase in 1-h maximum O_3 (with 95% CI).

suggested by Dersimonian and Laird (1986) to take into account the among-studies variance. However, as the O_3 models were specified in some studies (i.e., the center group of studies in Figure 3) to account for differing hot and cold extreme effects (e.g., by adding dummy variables for extreme days), the overall O_3 RR estimate increased somewhat (pooled RR=1.033, 95% CI: 0.985–1.084), although it is statistically nonsignificant in these studies. Finally, in those studies (to the right in Figure 3) in which the nonlinear nature of temperature effects was more fully addressed, statistically significant O_3 RRs greater than 1.0 were consistently found, irrespective of the averaging time employed (pooled RR=1.056, 95% CI: 1.032–1.081). Thus, the method of weather specification was shown to be an important factor affecting the O_3–mortality effect estimate, with the better (i.e., more realistic) weather–mortality models resulting in the largest O_3 RR estimate. Those studies that have incorporated nonlinear temperature specifications are therefore seen to be of the most interest in assessing the potential premature mortality effects of O_3.

As shown in Figure 3, the estimated RR associated with a 100-ppb increase in 1-h maximum O_3 (which is roughly equivalent to 75 ppb daily 8-h maximum, or 40 ppb 24-h average O_3) increases as the method of weather specification is improved. When all 15 studies are considered (after controlling for season, weather, day-of-week, and PM copollutant effects), the combined analysis yielded an overall 100-ppb effect size of RR=1.036 per 100-ppb increase in daily 1-h maximum O_3 (95% CI: 1.023–1.050). A pooling of the seven studies with nonlinear temperature terms (and simultaneously including a PM index term) reveals that the overall RR=1.056 per 100-ppb increase in daily 1-h maximum O_3 (95% CI: 1.032–1.081), and that the various RRs are not statistically different from each other. Note, however, that none of these studies have yet considered the likely interaction of temperature and percent RH in enhancing the model's fit of weather's influence on daily mortality (and making the weather specification less like O_3, further reducing intercorrelation between the pollutant and weather terms in the model). For example, in the illustrative case of Detroit presented here, the ER estimate was highest and most significant for the model that included a weather specification that incorporated the temperature–RH interaction. Thus, even those recent models discussed in this paper that include nonlinear temperature terms may yet be underestimating the acute effects of O_3 on premature mortality.

In Figure 4, the O_3 RR results from the studies using nonlinear fits of temperature are shown both for the "O_3-alone" model and for the same model with a PM air pollution index also included. As indicated by these results, the inclusion of the PM index tends to only very slightly change the combined estimate of the size of the O_3 RR. Thus, the influence of a copollutant on the O_3 effect estimate is seen to be small in these studies.

The size of the premature mortality effects by O_3 indicated by this synthesis is greater than that assumed in most past assessments of O_3–mortality effects, especially in the case where the combined estimate from the nonlinear

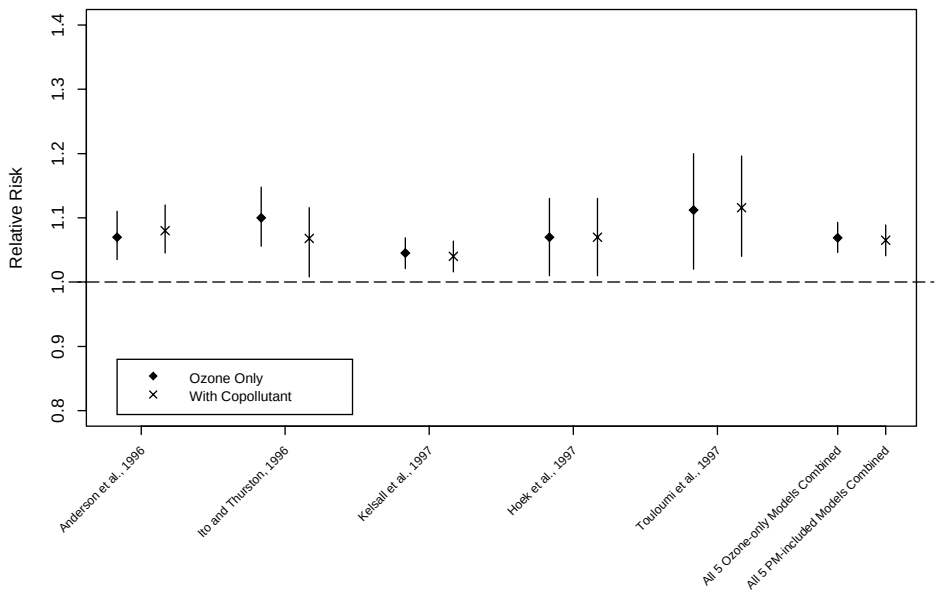

Figure 4. Effect of including a PM index on estimates of the total mortality RR of a 100-ppb increase in daily 1-h maximum O_3 in studies using a nonlinear temperature specification (with 95% CI).

temperature specification models are considered. For example, the recent U.S. EPA Regulatory Impact Analysis used an O_3–mortality effect estimate equivalent to RR=1.029 for a 100-ppb increase in 1-h maximum O_3 (U.S. EPA, 1997). This suggests that the U.S. EPA estimates of the premature mortality benefits of implementing the most recently promulgated O_3 standard may have been underestimated by as much as a factor of two.

Conclusions

Whereas older epidemiological studies of O_3 and mortality had severe limitations, especially regarding the way they handled confounding by weather, newer studies have used more complex weather specifications in their analyses, and the O_3–mortality relationship has been more clearly identifiable. Among some of the newer aggregate population time-series epidemiology studies examining the acute effects of ambient O_3, many have yielded significant associations with premature mortality. Our examination in this paper of an illustrative example for Detroit, MI indicates that the choice of specification of the weather terms can have an important effect on the O_3 health effects estimate, unlike the quite limited influences of weather variables for time-series mortality analyses of PM_{10}. The nonlinear nature of the weather–health effects dose–response relationship should, therefore, be carefully investigated and addressed when specifying a time-series model, so as to avoid an inappropriate accounting for O_3–mortality associations. This is especially true with regard to the nonlinearity of the heat–mortality relationship, the temperature–RH interaction, and the potentially high intercorrelation of a hot temperature variable with O_3. Indeed, when these nonlinearities and interactions are considered in the weather specification so as to avoid intercorrelation with the more linear O_3 association with health effects, it appears to be an important determinant as to whether the O_3 association can be distinctly discriminated by such regression analyses.

We conclude from our illustrative Detroit, MI analysis that, depending on the purpose of the analysis, optimizing a model by the goodness of fit (e.g., the AIC) alone will not necessarily yield the most useful model. Indeed, in this city, the consideration of AIC alone would have pointed to choosing the simplest weather model, but that model also had the greatest intercorrelation with O_3. This result was concluded to be a product of the degrees of freedom adjustment incorporated into the AIC, which penalizes more complex models for the added degrees of freedom used. If the objective of an analysis is solely to use the model to predict total mortality, then the AIC is likely to be a very important statistic for deciding between models. However, if the aim of an analysis is to interpret the resulting regression coefficients to derive individual variable's effect estimates, such as for O_3, then the intercorrelations of the model coefficients should also be considered in choosing the most useful model, as highly intercorrelated variables (e.g., $r \geq 0.5$) included simultaneously in a multiple regression model can greatly affect the estimation of their respective coefficients. As noted by Snedecor and Cochran (1980), "the meaning of a regression coefficient depends on the other x's in the model" and "it helps if the investigator can find a set in which the intercorrelations among the x's are small."

The conclusion that the goodness of fit (e.g, minimum AIC) should not be used as the sole criteria in selecting a "best" model has been made before, notably by Cakmak et al. (1998) regarding selection of the long-wave cycles to remove from the outcome variable in a time-series analyses. These authors found that overfiltering could result in the inducement of undesirable negative autocorrelation in a model. Therefore, the authors recommended three criteria in selecting the degree of smoothing in the outcome variable: "removal of temporal cycles, minimizing autocorrelation and optimizing goodness of fit." To these three criteria, we now argue that the intercorrelation of the other "independent" model terms (e.g., weather) with the pollutant terms should also be considered when choosing a model if the intent of the analysis is to utilize the resulting pollutant effect estimate coefficient(s) for risk-assessment purposes.

Finally, our quantitative synthesis of some recent epidemiological studies of O_3 has provided combined estimates of O_3 effect sizes for premature mortality in the studies considered here. For all of the total mortality–air pollution time-series studies considered here, the combined analysis RR=1.036 per 100-ppb increase in daily 1-h maximum O_3 (95% CI: 1.023–1.050), whereas the subset of studies that were judged to have better modeled the complex influences of weather yielded a combined estimate of RR=1.056 per 100-ppb increase in daily 1-h maximum O_3 (95% CI: 1.032–1.081). Generally, we found that the O_3–mortality effect estimate was largely independent of the PM mass–mortality association, and that the O_3 effect estimate increases in size and statistical significance when the nonlinearity and the humidity interaction of the temperature–health effect association are incorporated into a model's weather specification. Thus, apparently unlike the case for estimates of the mortality associations with PM mass, the estimates derived for O_3 seem to be sensitive to the way that seasonality and weather are handled. Such weather–pollutant intercorrelations should be considered and minimized as much as possible during the model development process. We hypothesize that this may well also apply to other pollutants with largely temperature-dependent formation mechanisms, such as secondary aerosol components of the PM mass (e.g., sulfates and acidic aerosols).

Acknowledgments

Funding for this study was provided by the National Institute of Environmental Health Sciences (NIEHS) under Grant ES05711. This study is also funded as a part of a Center Program supported by the NIEHS under Grant ES00260. The authors also thank Dr. Morton Lippmann for his comments on this paper.

References

Akaike H. A new look at statistical model identification. *IEEE Trans Autom Control* 1974: AU-19: 716–722.

Anderson H.R., Ponce de Leon A., Bland J.M., Bowers J.S., and Strachan D.P. Air pollution and daily mortality in London: 1987–1992. *Br Med J* 1996: 312: 665–669.

Borja-Aburto V.H., Loomis D.P., Bangdiwaia S.I., Shy C.M., and Rascon-Pacheco R.A. Ozone, suspended particulates, and daily mortality in Mexico City. *Am J Epidemiol* 1997: 145: 258–268.

Cakmak S., Burnett R., Krewski D. Adjusting for temporal variation in the analysis of parallel time series of health and environmental variables. *J Exposure Anal Environ Epidemiol* Apr–Jun 1998: 8 (2): 129–44.

California Department of Public Health. Clean air for California: initial report of the Air Pollution Study Project. State of California, Department of Public Health, San Francisco, CA, 1955.

Caroll R.J., Ruppert D., and Stefanski L.A. Measurement error in nonlinear models. Monographs on statistics and Applied Probability 63. Chapman & Hall, London, 1995, p. 22.

Cifuentes L.A., and Lave L. Association of daily mortality and air pollution in Philadelphia, 1983–1988. *J Air Waste Manage Assoc*, 1997. Submitted for publication.

Cleveland W.S. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 1979: 74: 829–836.

Dersimonian R., and Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986: 7: 177–188.

Ellis F.P. Mortality from heat illness and heat-aggravated illness in the United States. *Environ Res* 1972: 5: 1–58.

Hastie T., and Tibshirani R. Generalized additive models. Chapman & Hall, London, UK, 1990.

Hoek G., Schwartz J.D., Groot B., and Eilers P. Effects of ambient particulate matter and ozone on daily mortality in Rotterdam, The Netherlands. *Arch Environ Health* 1997: 52: 455–463.

Ito K., Kinney P., and Thurston G.D. Variations in PM10 concentrations within two metropolitan areas and their implications to health effects analyses. *Inhalation Toxicol* 1995: 7: 735–745.

Ito K., Thurston G., Nadas A., and Lippman M. Monitor-to-monitor temporal correlation of air pollution and weather variables in the North-Central U.S. *J Exposure Anal and Environ Epidemiol* 2001: 11 (1): 21–32.

Kelsall J.E., Samet J.M., Zeger J.E., and Xu J. Air pollution and mortality in Philadelphia 1974–1988. *Am J Epidemiol* 1997: 146: 750–762.

Kinney P.L., and Ozkaynak H. Associations of daily mortality and air pollution in Los Angeles County. *Environ Res* 1991: 54: 99–120.

Kinney P.L., and Ozkaynak H. Associations between ozone and daily mortality in Los Angeles and New York City. *Am Rev Respir Dis* 1992: 145: A95 (Abstract).

Kinney P.L., Ito K., and Thurston G.D. A sensitivity analysis of mortality/PM-10 associations in Los Angeles. *Inhalation Toxicol* 1995: 7: 59–69.

Lipfert F.W. Air pollution and community health: a critical review and data sourcebook. Van Nostrand Reinhold, New York, NY, 1994.

Mage D.T., and Buckley T.J. The relationship between personal exposures and ambient concentrations of particulate matter. Paper No. 95-MP18.01. Proceedings of the 88th Annual Meeting of the Air and Waste Management Association, Pittsburgh, PA, 1995.

Mage D., Wilson W., Hasselblad V., and Grant L. Assessment of human exposure to ambient particulate matter. *J Air Waste Manage Assoc* 1999: 49: 1280–1291.

Moolgavkar S.H., Luebeck E.G., Hall T.A., and Anderson E.L. Air pollution and daily mortality in Philadelphia. *Epidemiology* 1995: 6 (5): 476–484.

Ozkaynak H., Xue J., Severance P., Burnett R., and Raizenne M. Associations between daily mortality, ozone and particulate air pollution in Toronto, Canada. Paper presented at the Colloquium on Particulate Air Pollution, Irvine, CA, January 24–25, 1995.

Ostro B.D., Sanchez J.M., Aranda C., and Eskeland G.S. Air pollution and mortality: results from a study of Santiago, Chile. *J Exposure Anal Environ Epidemiol* 1996: 6: 97–114.

Pope C.A. III, and Kalkstein L.S. Synoptic weather modeling and estimates of the exposure–response relationship between daily mortality and particulate air pollution. *Environ Health Perspect* Apr 1996: 104 (4): 414–420.

Samet J.M., Zeger S.L., Kelsall J.E., Xu J., and Kalkstein L.S. Particulate air pollution and daily mortality: analysis of the effects of weather and multiple air pollutants. The Phase I.B Report of the Particle Epidemiology Evaluation Project. Health Effects Institute, Cambridge, MA, March 1997.

Samet J., Zeger S., Kelsall J., Xu J., and Kalkstein L. Does weather confound or modify the association of particulate air pollution with mortality? An analysis of the Philadelphia data, 1973–1980. *Environ Res* 1998 Apr: 77 (1): 9–19.

Schwartz J. Health effects of air pollution from traffic: ozone and particulate matter. In: Fletcher T., and McMichael A.J. (Eds.), Health at the Crossroads: Transport Policy and Urban Health. John Wiley & Sons Ltd. NY, NY. 1997.

Simpson R.W., Williams G., Petroeschevsky A., Morgan G., and Rutherford S. The association between outdoor air pollution and daily mortality in Brisbane. *Aust Arch Environ Health* 1997: 52: 442–454.

Snedecor G.W., and Cochran W.G. Statistical Methods. 7th edn. Iowa State University Press, Ames, IA, 1980, pp. 343, 353.

Thurston G.D., and Kinney P.L. Air pollution epidemiology: considerations in time-series modelling. *Inhalation Toxicol* 1995: 7: 71–83.

Touloumi G., Katsouyanni K., Zmirou D., Schwartz J., Spix C., de Leon A.P., Tobias A., Quennel P., Rabczenko D., Bacharova L., Bisanti L., Vonk J.M., and Ponka A. Short term effects of ambient oxidants exposure on mortality: a combined analysis within the APHEA project. *Am J Epidemiol* 1997: 146 (2): 177–185.

U.S. Bureau of the Census. County and City Data Book, 1983. U.S. Department of Commerce. U.S. Government Printing Office, Washington, DC, 1983.

U.S. Environmental Protection Agency (U.S. EPA). Air quality criteria for ozone and other photochemical oxidants. EPA Report No. EPA-600/P-93/004aF. Environmental Criteria and Assessment Office, Research Triangle Park, NC, 1996, pp. 7–94.

U.S. Environmental Protection Agency (U.S. EPA). Regulatory impact analysis for the particulate matter and ozone National Ambient Air Quality Standards and proposed regional haze rule. Appendix J. Table 3. Office of Air Quality Planning and Standards, Research Triangle Park, NC, 1997.

Verhoeff A.P., Hoek G., Schwartz J., and Wijnen J.H. Air pollution and daily mortality in Amsterdam. *Epidemiology* 1996: 7: 225–230.

Wechsler C.J., Shields H.C., and Naik D.V. Indoor ozone exposures. *J Air Pollut Control Assoc* 1989: 39: 1562.