A Review of the Statistical Methodologies Applied in the Time-Series Epidemiologic Studies of Ambient Carbon Monoxide and Health

Paul Switzer, Ph.D.
Department of Statistics
Stanford University
Stanford, CA 94305-4065

and

Allen S. Lefohn, Ph.D.
A.S.L. & Associates
111 North Last Chance Gulch
Suite 4A
Helena, MT 59601

April 23, 1999

Introduction

The number of CO-related deaths and acute poisonings is a serious public health concern. The epidemiological data in the draft Air Quality Criteria for Carbon Monoxide Document (U.S. EPA, 1999) indicate a small but persistent association between ambient CO levels and hospital admissions. However, EPA (1999) notes that there remain important uncertainties in the current epidemiologic database for ambient CO and other air pollutants.

There is an increasing realization that health effects observed in association with any single pollutant may actually be mediated by multiple components of the complex ambient mix of air pollutants and physical stressors (e.g., temperature, humidity, etc.). In addition, observed effects of CO and other pollutants appear to be subtle in relation to effects of other risk factors, such as smoking, socioeconomic status, and demographic factors. This lends uncertainty to the potential public health consequences of further reductions of ambient CO or any other single ambient air pollutant at present-day U.S. CO levels.

Moolgavkar and Luebeck (1996) critically reviewed epidemiologic studies of particulate air pollution and mortality in U.S. cities with respect to important methodological issues. The authors noted that many of the studies suffered from serious deficiencies in their control of the confounding effects of other pollutants. Several investigators have noted a high degree of correlation among both air pollutant and weather variables (Saldiva et al., 1994; Saldiva et al., 1995; Schwartz and Morris, 1995; Pereira et al., 1998). Moolgavkar and Luebeck (1996) concluded that the small risks reported to be associated with the particulate component of air pollution could easily be attributed to residual confounding by copollutants. In addition, the authors noted that most of the studies had not considered modification of air pollution effects by seasonal factors. Although Moolgavkar and Luebeck (1996) focused on reviewing epidemiologic
studies of particulate air pollution and mortality, their concerns are relevant for other pollutants such as carbon monoxide, sulfur dioxide, and ozone because of the similar methodologies employed.

In all of the epidemiologic studies, population exposures are assumed to be adequately represented by measurements at ambient monitoring stations. However, this assumption is not necessarily correct. Because of the presence of indoor sources, such as gas appliances, indoor CO concentrations often exceed the outdoor levels (U.S. EPA, 1991). When there are no indoor sources associated with personal exposure, fixed-site ambient monitors do not necessarily represent the outdoor CO concentrations in a given urban area because of spatial and temporal variations (U.S. EPA, 1999). The lack of spatial homogeneity does not lead to the correct temporal pattern. Carbon monoxide is not necessarily homogeneously distributed in space (i.e., not all monitors in a region have high values on the same day). Although some epidemiological investigators have averaged the concentrations measured at several nearby monitors, this may not be appropriate. Episodic concentrations are important for assessing human health effects and averaging results in “smoothing” the information, which will result in weakening the linkage between exposure and effects (Lefohn, 1997).

Because of the above considerations, the link between ambient concentration and personal exposure is another concern. Morris and Naumova (1998) point out that levels measured at ambient monitors poorly represent individual exposures. The authors note that in some cases individual exposures can be far higher than those measured at ambient monitors. It is clear that studies are needed to explore the relationship between the daily variability of personal exposure to CO and the daily variability of the CO index used in the epidemiologic time series analyses to more accurately assess what risk ambient CO poses to public health.

In an attempt to better understand the limitations of the methodologies used in the epidemiological studies, we examined the 26 peer-reviewed papers listed in Appendix A. EPA has indicated that these are among the most important publications associated with the epidemiological chapter in the CO Criteria Document. Our specific focus is on the ten papers listed in Appendix B that we believe illustrate our most important methodological concerns. All the papers report measures of relative risk derived from fitted statistical models by comparing the model-predicted daily mortality/morbidity for a high versus a low pollutant level.

Conclusions and Recommendations

Based on our review of the 26 epidemiological studies, our main conclusions and recommendations are stated in the seven points below:

1. To deal with the effect of confounding variables, analysis should be done in two stages. In the first stage the effects of confounders are modeled. Low frequency mortality/hospitalization (MH) variations need to be filtered out and adjustments should then be made of the filtered MH series for daily weather variables and daily co-pollutant concentrations. All adjustments of the MH series should then be fixed. In the second stage the effects of daily CO on the fixed adjusted daily MH series are assessed. In fact, it
appears that most studies seem to simultaneously estimate or re-estimate models with both confounders and CO variables. In a few papers there is an implied two-stage approach but it is not clear how it was carried out and, even in these cases, the approach did not extend to other-pollutant confounders.

2. There would still remain the issue of whether the CO effects are additive, i.e., should the CO effect be modeled to be the same at all levels of the confounding variables? For example, is it reasonable for CO effects on MH to be the same at both low temperatures and high temperatures? With rare exceptions, e.g., Morris and Naumova (1998) these papers treat CO as an additive effect with the implications that this carries. More insight can be gained by examining the joint effects of temperature and CO, say, via a nonparametric response surface. In that case it would not be necessary to do the prior temperature adjustment to the MH series. A less desirable alternative that addresses the same issue is to do separate analyses for different seasons or broad temperature ranges.

3. With few exceptions, the CO response has been modeled in a proportional way, say via a linear term in Poisson regression. The implication is that a 1-ppm CO increase has the same effect at any CO base level. For these rich data sets, it should be possible to explore the dose-response function nonparametrically. This analysis becomes especially important if one ties the modeling to a regulatory purpose or one wishes to compare results from different studies or geographic regions. Dose-response analysis would also replace the reporting of effects in terms of the confusing interquartile CO increase, because the interquartile CO is many times different for different locations.

4. Since daily weather has a demonstrated effect in all studies, it is important that careful attention be paid to weather modeling. The common practice of choosing a single temperature variable, even if it is the best of several, is not adequate. Kelsall et al. (1997) demonstrates that same-day temperature affects MH in a very different way from the average temperature of the preceding three days, both of which have explanatory power.

5. Model selection issues could have important effects on computed significance levels and confidence coefficients since, in all these studies, models are selected which maximize the apparent effect of CO. It is difficult to do a mathematical correction, but alternatives such as data-splitting, cross-validation, and data resampling provide ways of more honestly assessing the uncertainty of effect sizes – provided that the complete modeling procedures are replicated. It is important that the investigators provide more details of the statistical/mathematical protocols utilized in their analyses so that modeling assumptions are clearly defined.

6. The modeling efforts do not consider the clear possibility that the effects of CO may not be incremental MH but rather MH cases brought forward in time. If CO hastens the onset of MH, is it a small time difference? Several papers insert comments at the end to this effect. Serious thought should be given to explicitly modeling time-advance pollutant effects instead of, or in conjunction with, incremental effects. The distinction is of paramount importance for regulatory purposes.
7. It may appropriate to identify alternative exposure indices that accumulate indoor and outdoor personal exposures in the analyses. The CO index in these studies is typically taken to be an average of daily maximum values over several ambient monitoring sites. The relation to personal exposure is known to be not very strong. Studies with demographic breakdowns suggest the possibility of personal heterogeneity of exposure. In addition, there may be a potential for confounding with non-ambient exposure (e.g., home space heaters are correlated with ambient concentrations).

Capsule Critiques of Selected Papers

Table 1 lists the 10 papers that were selected for a focused review. Six of the papers deal with daily hospital admissions and carbon monoxide and the other four are associated with daily mortality and carbon monoxide. It was not possible to extract all the important methodological details from the published papers, particularly with regard to computations of confidence intervals. All of these papers recognize the confounding effect of time trends in mortality/hospitalization (MH), separate from the daily fluctuations. All recognize the confounding effect of weather and the possibility that some pollutants could confound the effects of other pollutants. Potential confounders account for a very large part of mortality variation while the incremental pollutant effects, if any, appear to be in the range of a few percent. This puts an extraordinary burden on the modeling effort and results will be tied to the modeling approach. For the reasons stated in the above conclusions and in the critique of selected studies below, it does not appear that the confounding and modeling issues have been adequately addressed.


The analysis model is not explicitly stated, but it appears that CO effects are assumed to be additive on top of all other corrections (i.e., weather, trend, month, and day of week). This implies that the added burden of CO would be the same in all seasons and temperatures, contrary to other studies that report separate analyses by season and find no CO effect in summer months. Only same day maximum temperature was used in the analysis, which should be regarded as an incomplete temperature adjustment. The multi-pollutant analysis shows CO effects are little changed in contrast to the findings of other studies, possibly because particulates were not included in the pollutant mix. Substantial differences between CO effects are reported for different cities, presumably due to geographic and population differences. No attempt is made to statistically reconcile such differences. Differences in effect estimates suggest that any observed CO effect, for a single city, differentially affects different population groups. Reported CO associations became insignificant, even with a 1-day lag, suggesting an incomplete accounting for temperature effects. A positive feature of this paper is the presentation of dose-response functions not tied to the assumption of proportional response. However, these functions give relative risk (RR) estimates very different from those reported elsewhere in the paper. They are
not reported with confidence bands and have unexplained differences in shape and level among different cities of the study.


Small CO effects [1-%-3% for 1.28ppm increase] are reported for heart failure admissions but not for ischemic heart disease admissions. Confounding variables were included simultaneously with pollution variables in the final model, rather than the more conservative approach of modeling residuals after first correcting for confounding variables. Effects are assumed to be proportional to CO in the fitted model. The possibility of differential CO effects at different temperatures is considered only via a season interaction term. Because the paper takes the trouble to nonparametrically map the joint humidity-temperature response, it would seem to be natural to do the same for the more important joint CO-temperature response, rather than rely on the additivity hypothesis. The reported low correlation between temperature and weather variables is not explained in terms of mediating variables such as season, and is contrary to correlations reported in other studies. Dose-response functions are not estimated nonparametrically; the underlying proportionality hypothesis was not tested. There is no discussion of adjustment of significance levels for model selection (e.g., selection of lags or averaging times). The interquartile range of CO values in this study is narrow – 1.58 to 2.86 ppm, which puts much burden on the model and makes difficult any extrapolation to CO levels near air quality standards.


The general structure of the analysis model is explicitly described. Effects of multiple pollutants, weather, and time-related confounders were assumed to be additive, with the incremental effect of CO being the same at any level of the confounding variables. CO effects were assumed proportional to concentration. Pollution variables enter into the model on the same footing as confounding variables, despite the implication in the paper that the confounding variables were given precedence. The joint effect of temperature and CO was not estimated nonparametrically (see the comment above on Schwartz & Morris, 1995). The interquartile range of CO values in this study is also narrow – 2.34 to 4.00 ppm.


This study compares eight urban U.S. counties. Weather data were included only for the same day of hospital admission, leaving the possibility of an incomplete correction. Interactions were not included in the model, thereby forcing pollutant effects to be the same at all temperature levels, in distinction to reports of other studies. The use of raw correlations between CO effect estimates and weather variables in a complex model is not warranted as an indicator of potential
confounding. Statements that small correlations between CO and weather variables in some counties make confounding unlikely are not warranted because these are unadjusted correlations (e.g., season). The study correctly points out that estimates of CO effects on hospital admissions cannot be used to estimate a reduced hospital burden under lowered CO scenarios because the study design cannot address the question of hospital admissions “brought forward a few days”.


Unlike other studies, this one allows for the possibility of differing CO effects in different temperature ranges, with significant association reported only at lower temperatures. Only same-day temperature data were used to correct the reported association, leaving out the possibility of multiple-day temperature effects. Pollutant effects disappeared even for one-day lags, creating some doubt about the robustness of the conclusions. Potential confounding variables and CO were entered together into the model instead of the more conservative modeling of residuals. The estimates of the combined effects of CO and temperature show a complex pattern, suggesting that simplified additive models obscure important information. This appears to be the only study to attempt to examine joint effects. The nonparametric estimates of dose-response, reported separately by temperature range, indicate strongly a threshold at about 2 ppm CO. However, it is not clear that temperature adjustments were made within temperature ranges, which could have a major effect.


The interquartile CO range for this study was 1.28 to 2.20 ppm. Single time lags for pollutants (3 days for CO) were chosen based on the strongest apparent effect with little discussion of significance level adjustment for model selection. CO and particle concentrations were highly correlated and their effect estimates were naturally negatively correlated; no joint dose-response function was provided. Similarly, the effects of weather were taken to be purely additive to those of pollutants. Separate analyses by season showed that asthma admissions increase with CO only in spring and fall, suggesting that there may be other confounding variables not considered which relate to asthma. A personal experience (i.e., A.S. Lefohn) with asthmatic children has shown consistently that children in Montana were most affected by asthmatic attacks during change of season periods (i.e., spring/fall time periods). Personal conversations with physicians in Montana indicate that pollen counts increase dramatically, on an episodic basis, during this period.

Daily CO was found not be a predictor of daily mortality after adjustment for trend, weather, and other pollutants. The methods of analysis are similar to those used in hospital admissions studies. A significant PM$_{10}$ effect for daily cancer mortality is reported, which serves as a cautionary note regarding the need for great care in the handling of confounding variables. Lags were selected for maximum effect without acknowledgement of the implications for statistical significance. This paper points out differences in pollutant effects among different demographic groupings, which raises the issue of the heterogeneity of exposure and the relationship of exposure to pollutant measures used in epidemiologic studies.


Daily weather data were considered either on the mortality day or lagged up to two days – but not in combination. Mortality data were considered as one- to three-day averages, whichever produced the largest pollutant effect. Separate optimization of lags and averaging periods for maximum effect were done separately for each of 11 cities, without regard to physiologic implications or to implications for statistical inference. Little consistency of CO mortality effects was seen across the 11 cities and the CO contribution to mortality, among the pollutant mix, appears inconsequential, as do the contributions of PM$_{10}$ and PM$_{2.5}$. The paper goes on to calculate reductions in death risks from proposed reductions of sulfur in gasoline without recognition of the issue of competing risks and marginally postponed mortality.


Unlike other studies, this one included temperature data both for the mortality day and the preceding three-day average, with the possibility of better accounting for the dominant temperature factor. The two temperature variables gave markedly different dose-response functions, which should serve as a cautionary tale for air-pollution epidemiology. There is also an implication that the fitted daily mortality model, based on time trends and temperature variables, was held fixed, when the pollutant variables were subsequently introduced. This would be appropriate, if it is indeed the case, rather than a model that simultaneously fitted the pollutant variables and the time-temperature confounders. However, the same strategy was not employed to deal with the distinct possibility that some pollutants could confound the effects of other pollutants (i.e., the potential other-pollutant confounders were not fixed in the subsequent analysis). The maximum reported CO effect, 1% mortality increase for 0.8 ppm increase in CO, was clearly at a 3-day lag, while the same-day CO effect was negligible and not statistically significant. This somewhat implausible result should be a caution regarding statistical inference with extensive model selection, when effects are small and there are dominant confounding variables. Because the results are highly model dependent, the authors do caution against using
the results of their analysis for the estimation of the consequences of lowered pollutant concentrations.


This study found no consistent association between daily mortality and CO. It used a “two-stage” analysis that seems to fix the fitted model for time trend and temperature and then enters the pollution variables. This is better than re-estimation of the whole model, where both confounders and pollutant coefficients are simultaneously estimated. All variables were assumed to have proportional effects and their effects were assumed to be additive. Separate season-specific analyses were done.

Discussion

Based on our review we found important shortcoming in the modeling and analysis of the CO/MH daily time series data. We propose the following changes:

- Analysis should be done in two stages.
  
  In the first stage, the effects of confounders are modeled. In the second stage, the effects of daily CO on the fixed adjusted daily MH series are assessed.

- There still remains the issue of whether the CO effects are additive (i.e., should the CO effect be modeled to be the same at all levels of the confounding variables?).
  
  For example, is it reasonable for CO effects on MH to be the same at both low temperatures and high temperatures?

- With few exceptions, the CO response has been modeled in a proportional way via a linear term in Poisson regression. The implication is that a 1-ppm CO increase has the same effect at any CO base level.
  
  For these rich data sets, it should be possible to explore the dose-response function nonparametrically.

- Since daily weather has a demonstrated effect in all studies, it is important that careful attention be paid to weather modeling.
  
  The common practice of selecting a single temperature variable, even if it is the best of several, is not adequate.
• Model selection issues could have important effects on computed significance levels and confidence coefficients since, in all these studies, models are selected which maximize the apparent effect of CO.

It is difficult to do a mathematical correction, but alternatives such as data-splitting, cross-validation, and data resampling provide ways of more honestly assessing the uncertainty of effect sizes – provided that the complete modeling procedures are replicated. It is important that the investigators provide more details of the statistical/mathematical protocols utilized in their analyses so that modeling assumptions are clearly defined.

• The modeling efforts do not consider the clear possibility that the effects of CO may not be incremental MH, but rather briefly time-advanced MH.

The distinction is of paramount importance for regulatory purposes. Serious thought should be given to explicitly modeling time-advanced pollutant effects instead of, or in conjunction, with incremental effects.

• The CO index in these studies is typically taken to be an average of daily maximum values over several ambient monitoring sites.

Studies with demographic breakdowns suggest the possibility of personal heterogeneity of exposure. As indicated in the Carbon Monoxide Criteria Document (EPA, 1999), the relation of ambient monitors to personal exposure is known to be not very strong. Thus, it may be difficult to relate ambient CO monitoring with population level mortality statistics.

References


APPENDIX A . List of Reviewed Papers


APPENDIX B. List of Commented Papers


