# Effects of Ambient Fine and Coarse Particles On Mortality in Phoenix, Arizona

Merlise A. Clyde     Peter Guttorp     Erin Sullivan

# NRCSE

Technical Report Series

NRCSE-TRS No. 040

February 1, 2000

The **NRCSE** was established in 1996 through a cooperative agreement with the United States Environmental Protection Agency which provides the Center's primary funding.

# EFFECTS OF AMBIENT FINE AND COARSE PARTICLES

# ON MORTALITY IN PHOENIX, ARIZONA[1]

Merlise A. Clyde
Institute of Statistics and Decision Sciences
Duke University

Peter Guttorp
National Research Center for Statistics and the Environment
University of Washington

Erin Sullivan
National Research Center for Statistics and the Environment
University of Washington

**Running Title:** Particulate Pollution and Mortality in Phoenix, AZ

**Key Words:** Bayesian Model Averaging, Bayes Information Criterion, generalized additive models, model uncertainty, PM2.5, PM10

**Corresponding Author:** Merlise Clyde, Institute of Statistics and Decision Sciences, Box 90251, Duke University, Durham, NC 27708-0251.

Phone (919) 681-8440, FAX (919) 684-8594, email: clyde@stat.duke.edu

1This work was supported by NSF grants DMS-97.33013 and by the United States Environmental Protection Agency through agreement CR825173-01-0 to the University of Washington. Although partly funded by the US EPA, this research has as not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

## Abstract

Considerable attention has been given to the health effects of ambient air borne particulate matter as the Environmental Protection Agency (EPA) revises the National Ambient Air Quality Standards (NAAQS). Much of the recent focus has been on the effects of fine particles, with the establishment of additional monitoring platforms to measure both fine and coarse particles for epidemiological studies. Much of the evidence supporting the 1997 standards is based on statistical models using generalized additive models for time series data. Among the statistical concerns raised by National Research Council is the issue of whether observed statistical associations are a result of multiple testing and selection effects due to model selection. We propose a method based on Bayesian Model Averaging to estimate relative risks that incorporates both uncertainty due to estimation but also uncertainty due to model choice. This incorporates uncertainty regarding the choice of confounding variables, choice of pollutant (fine and coarse particles), and which lags of all variables should be included in the model. This approach is illustrated using fine and coarse particulate matter data from the National Exposure Research Laboratory research monitoring platform in Phoenix, AZ for the time period of May 1995 to June 1998. We consider elderly non-accidental mortality for three regions of increasing size in the Phoenix metropolitan area, as well as using accidental mortality for all ages as a control population. We find a weak, but suggestive, particulate matter effect on elderly mortality only for the geographic region where fine particles are expected to be spatially homogeneous; posterior probability of a PM effect is 0.91 and 90% probability interval for the relative risk (RR) with a 1 IQR increase in both fine and coarse particulate matter levels is $(1 < RR \leq 1.04)$. While previous scientific information would suggest that fine particles should have a larger effect than coarse particles on health outcomes, we find instead that the effect of coarse particles on mortality is stronger, with very little support for models that include fine particles, but not coarse particles.

## Introduction

Although new standards for particulate matter (PM) were proposed by the Environmental Protection Agency (EPA) in 1997, there remains a high degree of uncertainty surrounding the effects of ambient fine and coarse particles on human health. Epidemiological models based on daily PM measurements and other confounding variables are often employed to estimate health effects. Because model selection is often used to eliminate unimportant confounding variables and to choose appropriate lags for these models, statistical significance may be overstated (Hodges 1987; Draper 1995), and observed associations may be the result of multiple hypotheses tested in the course of variable selection, an issue raised by the 1998 National Research Council report on "Research Priorities for Airborne Particulate Mater". Different model selection strategies have, in some cases, led to very different models and conclusions for the same set of data (Schwartz 1993; Clyde 2000). Rather than selecting a single model, we approach the problem using Bayesian Model Averaging (BMA), a technique that accounts for model uncertainty by averaging over a wide class of models supported by the data.

It has been hypothesized that fine particles are more damaging to human health than large particles. We evaluate the effect of particle size on elderly mortality using data collected at the National Exposure Research Laboratory's research monitors in Phoenix, AZ from May 1995 to June 1998. We apply Bayesian Model Averaging to incorporate uncertainty about which pollutants (fine versus coarse PM), confounding variables, and transformations and lags of these variables should be included in the model. We construct a posterior distribution for the relative risk based on a simultaneous IQR (interquartile range) increase in both fine and coarse PM variables adjusted for model selection effects using BMA.

## Methods

*Data Sources*

Mortality data were obtained from the Arizona Department of Health Services for 1995-1998 and matched to particulate matter and meteorological data for the same time period. Because of concerns about the effect of spatial heterogeneity and potential bias, we constructed four mortality response variables (Figure 1). Three, ELDERLY MORTALITY METRO AREA, ELDERLY MORTALITY UNIFORM PM2.5 and ELDERLY MORTALITY UNIFORM PM10, are each a daily count of non-accidental deaths for people aged 65 and over in three geographical regions. ELDERLY MORTALITY METRO AREA includes deaths which occurred in the Phoenix metropolitan area using the region defined as Phoenix Division, Arizona by the United States Census Bureau. The second variable, ELDERLY MORTALITY UNIFORM PM2.5, includes deaths in a smaller subset of zip codes which are thought to have spatially similar levels of fine particles or PM2.5 (particulate matter with aerodynamic diameter of 2.5 microns) throughout the region (personal communication Jane Koenig, University of Washington). The third geographical region, is smaller yet and is thought to have fairly homogenous PM10 levels, and therefore spatially homogeneous levels of coarse particles; the corresponding response variable is ELDERLY MORTALITY UNIFORM PM10. Zip codes within Phoenix that define UNIFORM PM10 are 85004, 85006-85009, 85012-85020, 85028-85029, 85031, 85033-85035, 85043, 85051. UNIFORM PM2.5 is defined using the above zip codes for Phoenix plus the following zip codes within Scottsdale, Mesa, Glendale and Tempe: 85201-85205, 85207-85208, 85212-85213, 85234, 85236, 85251, 85253, 85256-85258, 85281-85284, 85296. The fourth response variable, ACCIDENTAL MORTALITY includes all accidental deaths for all age groups which occurred in the Phoenix metropolitan region. As there is no reason to believe that particulate matter should be associated with accidental mortality, this provides a sensitivity check on the methodology. Because early 1995 was a very mild flu year in contrast to the following years, we have chosen to model mortality using a start date of May 1, 1995.

In light of the upcoming review of the National Ambient Air Quality Standards, our analysis focuses on estimating health effects of fine particles PM2.5 and coarse particles (PMC), defined as PM10 - PM2.5. Daily particulate matter readings (Figure 1) from a TEOM monitor (Tapered Element Oscillating Microbalance) were obtained from the EPA's National Exposure Research Laboratory for both PM10 and PM2.5. In Phoenix, particle mass is typically dominated by the coarse fraction. The average of PM2.5 over the period is 13.75 $\mu g/m^3$ (range 0.02 to 40.95 $\mu g/m^3$), while the average for PM10 is 45.42 $\mu g/m^3$ (range 5.19 to 185.66 $\mu g/m^3$). The correlation between fine and coarse particle daily levels is 0.65. Over the 3 year period, roughly 20% of the coarse particle data is missing, while approximately 16% of the PM2.5 observations are missing. Models are based on cases with complete data only.

Additional daily meteorological data were obtained from the U.S. National Climatic Data Center (NCDC) in Ashville, NC, and include minimum and maximum daily temperatures (TMIN and TMAX). Specific humidity (SH) was derived from the NCDC data. To allow for nonlinear effects of temperature and humidity on mortality, we considered squared components of each (TMAXSQ, TMINSQ, and SHSQ). While there are numerous studies of the relationship between mortality and PM, there is no consistent agreement on which lags of PM and confounding variables to include. Many papers consider one lag at a time, or construct 3 day averages, and lags of 3 day averages. We allow for potentially any lag from the present day (lag0) up to a lag of 3 days (lag3) to be included.

*Statistical Model*

For each of the four response variables, we model mortality using Poisson regression with a log link, so that the log of expected mortality follows a linear model. We control for a nonlinear temporal trend in mortality modeled using smoothing splines to represent seasonal variation in mortality due to flu or other unmeasured trends such as population growth over time. We consider 30 knots spaced approximately a month apart to capture variation on a scale of one month or greater. Uncertainty regarding the number of knots was addressed as in Clyde 2000. We also adjust for the potential confounding meteorological variables of minimum and maximum daily temperature, specific humidity, lags of 0, 1, 2, and 3 days in each of these variables, and quadratic terms in each. We consider particle size (fine and coarse) as well as the lag of the effect (0, 1, 2, or 3 days). With the nonlinear BASELINE, there are 29 variables under consideration.

Previous studies have used model selection to determine which variables and lags to include. As variable selection may involve numerous tests of hypotheses, the resulting significance levels may be called into question, and there is the concern, as raised in the 1998 National Research Council report, that the positive associations between health outcomes and particulate matter are a result of multiple testing used in model selection. While results from multiple models may be presented, there has been no consistent method for how to combine multiple inferences for the same set of data. Selecting a single model and making all subsequent inferences based on the selected model often leads to policy decisions that are riskier than one may think (Draper 1995).

Bayesian model averaging (Raftery 1996; Hoeting et al. 1999; Clyde 2000) provides a coherent alternative for combining inferences from different models and addressing model selection in subsequent inferences. Under BMA, each model contributes proportionally based on the support it receives from the data, as measured by the posterior probability of each model. Potential models using the 29 covariates were obtained using the leaps and bounds algorithm to find the best models of each size. To provide a baseline reference analysis, we use the Bayes Information Criterion or Schwarz Criterion (Schwarz 1978) for determining posterior model probabilities, where BIC for model *M* is

$$BIC(M) = \text{deviance}(M) + \dim(M) \log(n)$$

and deviance(*M*) is the deviance statistic for model *M* (-2 log likelihood) and dim(*M*) is the number of variables in model *M*. This imposes a heavy penalty on models that contain a large number of parameters and has been shown to lead to consistent selection of the true model. Assuming that all models are equally likely a priori, the posterior probability of model *M* given the data is

$$Pr(M \mid data) = \frac{\exp\{-BIC(M)\}}{\sum_{m=1}^{K} \exp\{-BIC(m)\}}$$

and the sum in the denominator is over all models. BMA using BIC has lead to improved predictive performance in many situations (Hoeting *et al.* 1999; Clyde 2000) and provides objective probabilities of models.

We consider model uncertainty regarding which meteorological variables should be included, whether there is a coarse or fine particulate matter effect, and which lags of the variables should be included. We calculate the posterior distribution of the relative risk under BMA associated with simultaneous one interquartile range changes in all lags of fine and coarse particulate matter variables included in the models (IQR fine PM = 9.1 $\mu g/m^3$, IQR coarse PM = 17.9 $\mu g/m^3$). The posterior distribution (approximate) for the relative risk given a model which includes any of the PM variables is a log normal distribution, where the log relative risk has a normal distribution centered at the maximum likelihood estimate of the relative risk under that model and with variance derived from the inverse Fisher information matrix for that model (see Clyde 2000 for details). For models that exclude all PM variables (fine, coarse, or any lags), the relative risk is identically 1. The posterior probability that there is no PM effect is obtained by summing the posterior probabilities of all models that exclude PM.

**Results**

Results using model averaging (Table 1) suggest that there is a weak particulate matter effect only for the mortality variable defined over the region with UNIFORM PM2.5. The posterior probability that the relative risk is equal to one is 0.09, with a 0.91 probability that the relative risk is greater than 1. For the other ELDERLY MORTALITY response variables, the data are inclusive about whether or not there is a particulate matter effect (probability of there being an effect or that the relative risk is not one ranges between 0.46 to 0.77). For the ACCIDENTAL MORTALITY response, the data are marginally in favor of there being no PM effect.

Focusing on ELDERLY MORTALITY UNIFORM PM2.5, the probability that the lag 1 coarse PM coefficient is non-zero is 0.87. For all other lags of coarse PM and fine PM, the probability that the coefficient is not equal to zero is less than 0.20, suggesting that the effect is primarily due to coarse particles rather than fine. This is also evident in the model space plot (Figure 2) that illustrates model uncertainty in the top 25 models. The left hand plot in Figure 2 shows an image of the model space for the top 25 models ranked by the log of the posterior model probability (or log Bayes Factor for comparing each model to the worst model). Potential models were obtained using the leaps and bounds algorithm to find the best models of a given size and then ranked based on BIC. Each row corresponds to a model, with the best models at the top. Each column corresponds to a variable, with the last 8 columns corresponding to coarse PM and its lags, followed by fine PM and its lags. Squares that are white indicate that the variable for that column is included in the model for the particular row. Restricting attention to only the top 25 models, only the lag 1 coarse PM variable is consistently included in the top models. The plot on the right hand side of Figure 2 shows 95% probability intervals for relative risks based on each of the top 25 models shown in the model space plot. For the top 25 models, models that include both coarse and fine PM all have probability intervals that include 1, while probability intervals for models with only coarse PM exclude 1. While over all models the data provide weak support for a coarse PM effect (probability = 0.91), the overall probability that there is a fine PM effect is only 0.39, and is inconclusive (marginally in favor of no effect).

All of the top 25 models receive similar support from the data, but lead to potentially different conclusions about the impact of PM on human health. BMA provides a way of combining these inferences to construct an overall estimate that adjusts for model selection. Figure 3 shows the distribution of relative risks based on the simultaneous change in both coarse and fine PM for each of the four response variables. In the distribution for ACCIDENTAL MORTALITY, most of the support is on relative risks equal to 1, however there is some support dispersed over other values. The distributions for the other ELDERLY MORTALITY responses are more concentrated. Again, it is only for the UNIFORM PM2.5 region that there is evidence of weak support for a PM effect. Despite the differences about whether there is a PM effect among the different response variables, the overall posterior mean for the relative risk ranges between 1.01 to 1.02 for the 4 responses (Table 1). However, it is only for the UNIFORM PM2.5 region that the overall 90% probability interval for relative risks does not contain 1.

## Discussion

For Phoenix, Arizona we consider the effect of model uncertainty on relative risk estimates based on simultaneous changes in both fine (PM2.5) and coarse (PM10 - PM2.5) levels using three different geographic regions. We find a weak, but suggestive PM effect only in the region where fine, but not necessarily coarse particles are expected to be spatially homogeneous, with most of the effect being attributed to coarse particles. Results using model averaging with the Bayes Information Criterion for the other response variables suggest that there is not enough evidence in the data to determine conclusively if there is a PM effect.

There are several refinements of the model that could improve the precision. In this analysis, no attempt has been made to impute missing data. Bayesian approaches for missing data would allow us to use all days with complete mortality data and account for uncertainty due to imputing missing observations in a valid way. This would be more efficient than ignoring cases with missing data. For mortality defined over the larger regions, PM measured at the single platform may not be an accurate exposure measure. Additional improvements can be made by expanding the statistical model to allow for spatial variation in PM levels or potential measurement error in the exposure data.

In this analysis we have considered uncertainty in the lag structure of both fine and coarse PM, but otherwise assume that the relationship between PM and mortality follows a simple log-linear relationship as is commonly used. Generalized additive models can be used to model nonlinear PM effects, such as a threshold effect, but complicate relative risk estimates. While exploratory modeling suggests that additional nonlinear functions of fine and coarse PM are not necessary, uncertainty regarding the functional form of the dose response curve can be easily addressed using BMA , if one chooses to expand the class of models under consideration.

In conclusion, the use of BMA in this analysis provides a methodological advance by accounting for both parameter uncertainty and model uncertainty in relative risk estimates and interval estimates, providing more realistic measures of uncertainty. Unlike traditional p-values, the

Bayesian approach provides posterior probabilities of whether there is a PM effect that also incorporate model selection uncertainty. While the best BIC model for the UNIFORM PM2.5 mortality data suggests that there is a statistically significant PM effect, there are other top BIC models that do not suggest this. BMA allows us to combine these inferences and to avoid the problems with false positives that can arise because of multiple hypothesis testing in the model selection framework.

## References


Clyde, M.A. Model uncertainty and health effect studies for particulate matter. *Environmetrics.* 2000; in press.

Draper, D. Assessment and propagation of model uncertainty (with Discussion). *Journal of the Royal Statistical Society, Series B.* 1995; 56:45-98.

Hodges, J.S. Uncertainty, policy analysis, and statistics.*Statistical Science.* 1987; 2:259-291.

Hoeting, J. et al. Bayesian model averaging: A tutorial. *Statistical Science.* 1999; in press.

National Research Council. Research priorities for airborne particulate matter. National Academy Press. Washington, DC; 1998.

Raftery, A.E. Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika.*1996; 83:251-266.

Schwarz, G. Estimating the dimension of a model. *Annals of Statistics.* 1978; 6:461-464.

Schwartz, J. Air Pollution and daily mortality in Birmingham, Alabama. *American Journal of Epidemiology.* 1993; 137:1136-1147.

**Table 1.** Summaries from Bayesian model averaging for the four analyses. For Uniform PM2.5, $1^+$ is rounded down so that 1 is actually not contained in the interval.

| ***Response Variable*** | ***Probability relative risk is 1 given data*** | ***Posterior mean relative risk*** | ***90% Probability Interval*** |
|---|---|---|---|
| Elderly Mortality Uniform PM10 | 0.54 | 1.01 | [1, 1.03) |
| Elderly Mortality Uniform PM2.5 | 0.09 | 1.02 | ($1^+$, 1.04) |
| Elderly Mortality Metro Area | 0.33 | 1.01 | [1, 1.03) |
| Accidental Mortality All Ages Metro Area | 0.51 | 1.02 | [1, 1.08) |

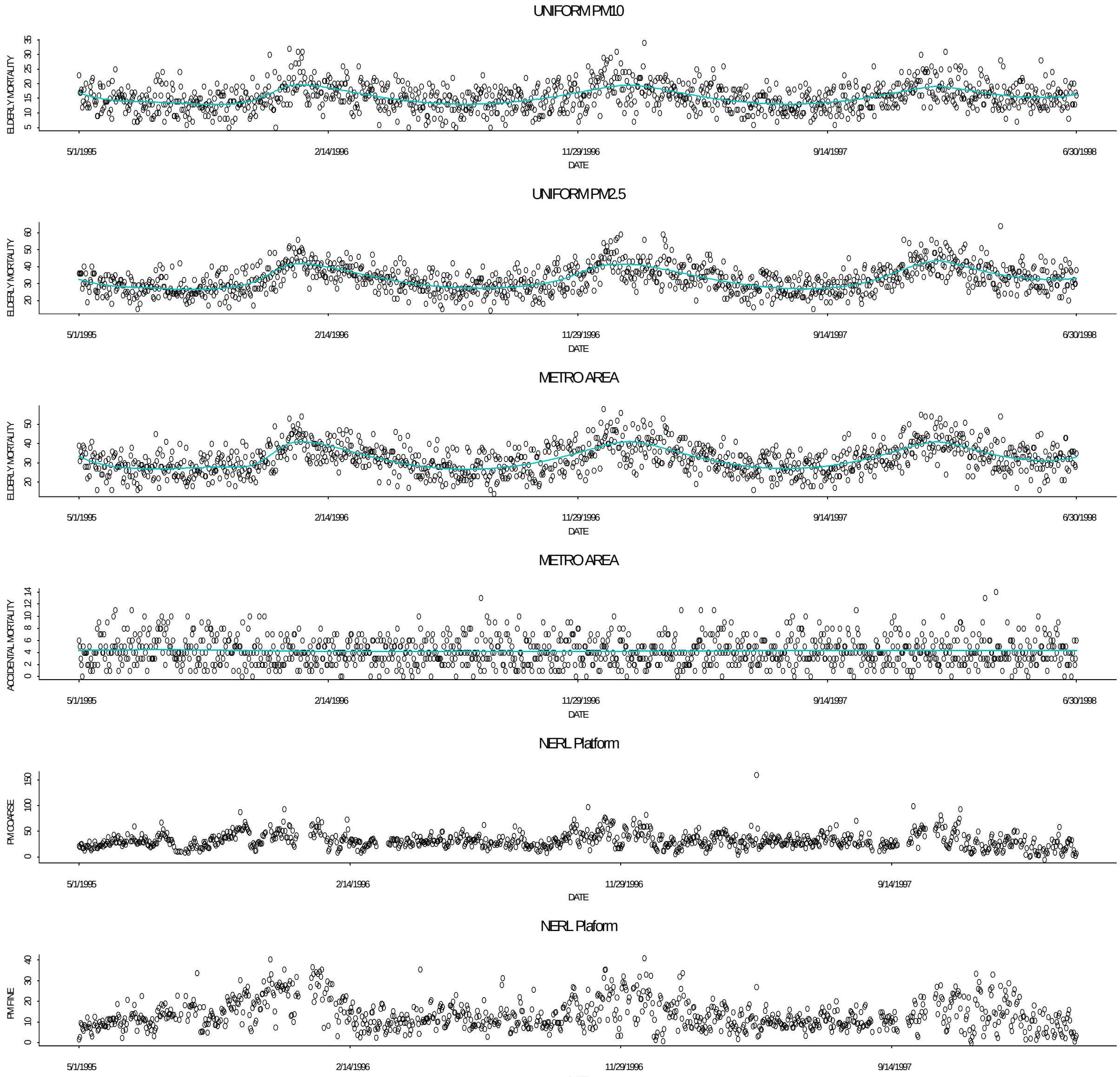


**Figure 1.** Time series of the four mortality responses, PM Coarse, and PM Fine for Phoenix, AZ from May 1, 1995 to June 30, 1998.

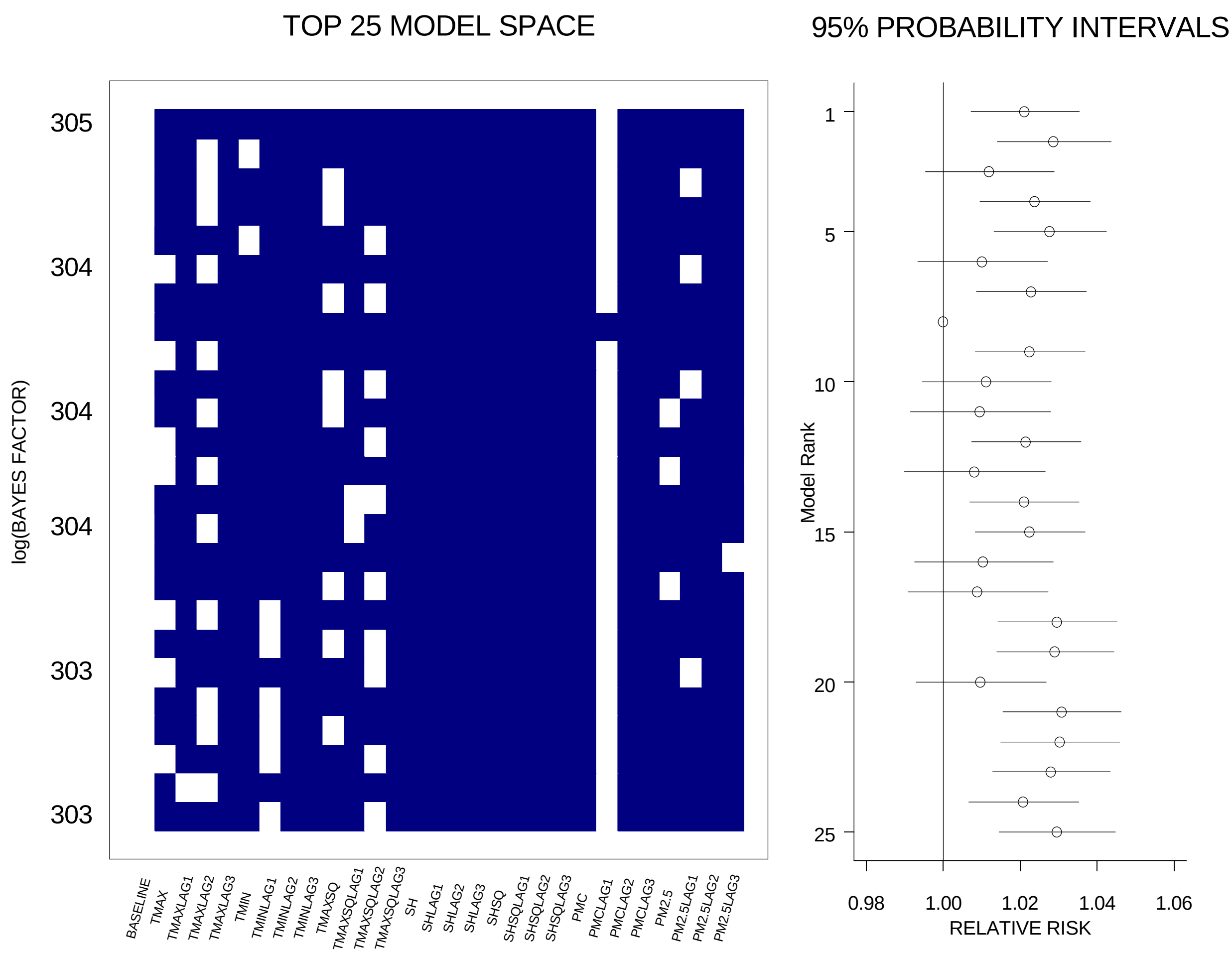


**Figure 2.** The top 25 models ranked by posterior model probabilities and associated 95% probability intervals for the relative risk under each model using elderly mortality for the region with uniform PM2.5. Rows in the model space correspond to models and columns correspond to variables, with white squares indicating that the variable for that column is included in the model for that row. The y-axis for the model space plot is the log(Bayes Factor) for comparing that model to the lowest probability model and is proportional to - BIC. Points in the probability intervals are the maximum likelihood estimates of relative risk under that model.

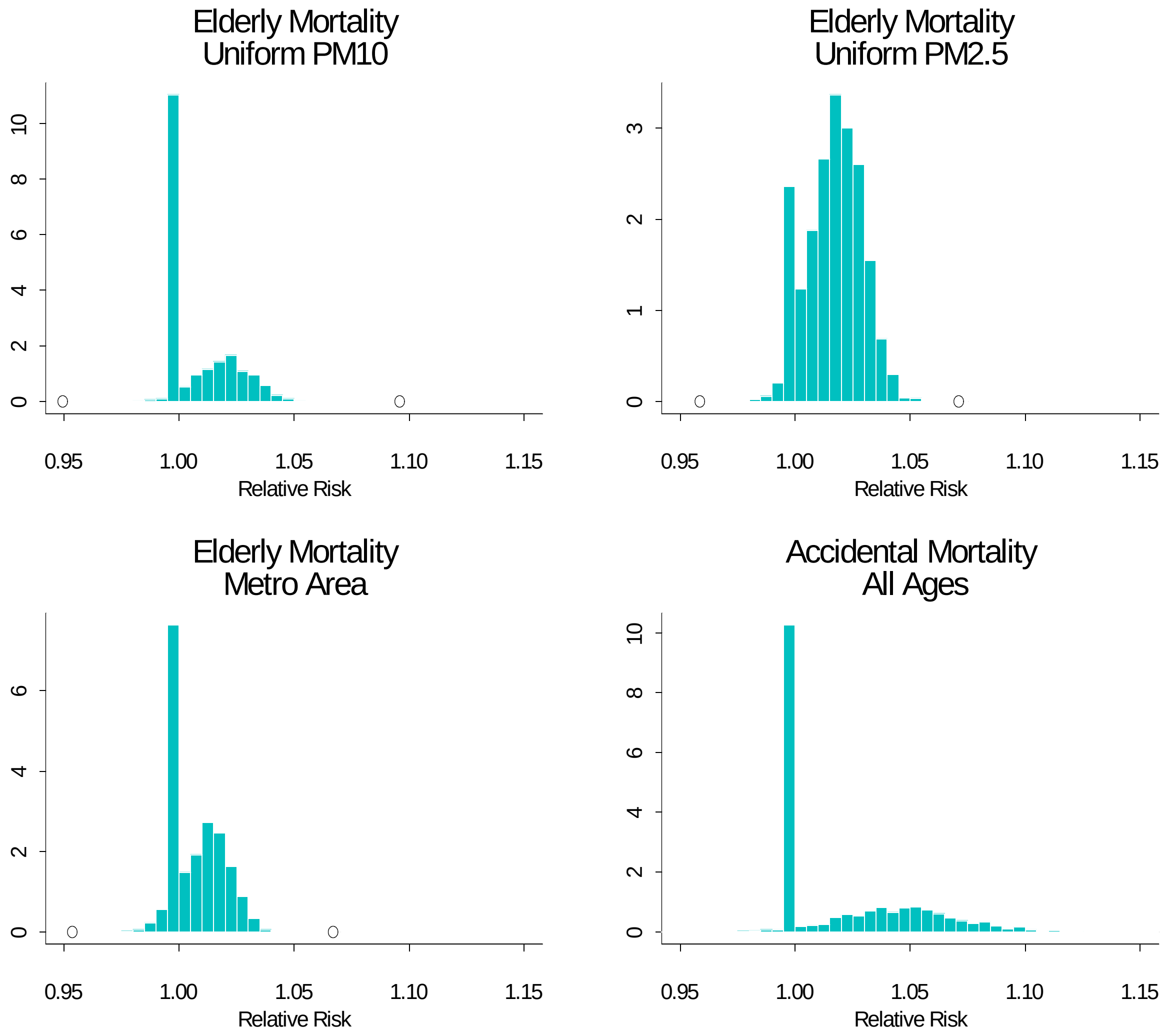


**Figure 3.** Distribution of relative risk incorporating both estimation uncertainty and model uncertainty using elderly non-accidental mortality for the three different regions and accidental mortality for all ages for the metro area. The spikes at one correspond to models that do not include any fine or coarse particulate matter variables or lags. The points indicate the range of the distribution.