Background and Aims: When executing epidemiological studies relating pollutants to human health (mortality) researchers confront numerous measures of air pollution to consider. Researchers confront both multiple measuring techniques that create series of data for ostensibly the same pollutant measured at the same or neighboring sites (for example, PM2.5, DFPSS and PM2.5 TEOM) and/or measurements of many different, but potentially related pollutants (for example, EC and OC). We discuss steps we have employed, findings of our methods, and suggestions for other researchers.

Methods: We have grouped pollutants using a process that begins with initial analysis of each of the pollutants individually using the same models: 1) average of lag 0 and lag 1 for each pollutant; and 2) pollutant individually as lag 0 through lag 4 or more). Upon review of the results we have clustered the pollutants by the patterns of the estimated coefficients using UNMIX (Lewis, et al., JAWMA 2003), PMR (Ramadan, et al. Chemometrics 2003), knowledge of the source of the pollutants, plus/minus and strength of the coefficients, and Spearman correlation coefficients of the various pollutant series. We also execute two-three pollution models for pollutants that are highly correlated to observe the impact of the multiple pollutants in the same model.

Results: UNMIX and PMF created different factors with only factors created by PMF illustrating significant results. These results were used to identify pollutants considered in subsequent models. In this study, it appeared the significant factors were driven by the chloride ion.

Conclusion: The use of source factors can help identify specific pollutants for further study. This study illustrates the need for researchers to beware of labels of factors and to investigate using multiple methods.