Design and Analysis of Studies of the Health Effects of Ozone

Alvaro Muñoz

Department of Epidemiology, The Johns Hopkins School of Public Health, 624 N. Broadway, Baltimore, MD 21205

The design and analysis of studies that investigate the effect of exposure to ozone on health outcomes need to define carefully the methods for the assessment of the exposure and to determine precisely which is the outcome of biological relevance. The estimation of sample size for longitudinal studies requires the expected rates of change among the exposed and unexposed, the variance of the outcome, and the correlation of measurements taken within an individual. Methods of analysis whose primary interest is in the combination of cross-sectional studies for the determination of the marginal distribution of the outcome are particularly appropriate for biological processes where the effect of exposure is acute. Conditional models are particularly useful for investigating the effect of changes in exposure on changes in outcome at the individual level. In addition, conditional models incorporate a dampening effect of exposure that may provide a reasonable agreement with several biological mechanisms. The identification of susceptible individuals and the description of the behavior of their outcomes over time may be better accomplished by using the within-individual variance as the outcome of interest. Discrepancies of the within- and between-individual regressions may be suggestive of chronic effects, and methodological research in this area is needed. Studies of the health effects of ozone exposure need to address the incorporation of missing data, measurement error, and the combination of complementary studies. — Environ Health Perspect 101(Suppl 4):231–235 (1993).

Key Words: Longitudinal studies, regression methods, missing data, environmental effects

Background

A central objective of studies on the health effects of ozone is to determine whether individuals who have been exposed to ozone have adverse health outcomes. The designer of a study needs to consider carefully the specific methods for measuring the exposure to ozone as well as to define precisely the outcome which will be used to identify adverse health effects. It is crucial to measure exposure accurately and to construct summaries that include duration and dose. In addition, it is important to study populations that are exposed to a range of exposures so that informative comparisons among groups can be made. There are three types of outcomes according to the number of possible values of the outcome. The simplest outcome is binary (i.e., yes/no) and typically refers to the presence of disease or a symptom. The occurrence of asthma is an example of a binary outcome that has received considerable attention in respiratory disease epidemiology. The next level of an outcome variable is the case of a categorical outcome including severity of disease. Typically this is the case of an outcome defining stages of disease as severe, moderate, mild, or absent. The third type of outcome is continuous (i.e., any value within the range of biologically possible values). This is the case of commonly used measures of pulmonary function, including the forced volume vital capacity and the forced expiratory volume after the first second in a spirometric maneuver. Studies for epidemiological research are either cross-sectional or longitudinal in nature. Cross-sectional studies involve the assessment of exposure and outcome at a fixed point in time. Longitudinal studies can be viewed as a collection of cross-sectional studies performed in the same group of individuals, thus providing repeated measures of exposure and outcome for each individual at different points in time. The design of cross-sectional studies requires the estimation of the sample size needed to detect the expected health effect with a high probability. The design of a longitudinal study requires the determination of the appropriate sample size as well, but in addition, one needs to specify the frequency of visits and the lag between them. Differences between designs of longitudinal studies can be characterized by three variables, namely, the number of individuals (N), the number of visits (V), which is provided by each individual, and the time lag (T) between the baseline and the last visit (Table 1). Panel studies typically have V large, T small, and N moderate. An example of a panel study is the case of daily follow-up of 100 individuals for a year (i.e., V=365, T=1 year, and N=100).

Longitudinal studies per se typically refer to the case of V small, T moderate or large, and N large. An example of a longitudinal study would be the yearly follow-up of 1000 school children from grade 4 to grade 12 (i.e., V=5, T=8 years, and N=1000). Laboratory experiments on animal models or chamber studies typically have V small or moderate, T small, and N small. An example of a chamber study would be the observation of 30 individuals on a weekly basis for 3 months (i.e., V=12, T=3 months, and N=30).

The analysis of data related to the effects of exposure to ozone on health outcomes requires the use of methods that allow for the incorporation of the simultaneous effect of different exposures or risk factors. The analysis of cross-sectional studies is considerably simpler than that of longitudinal studies. Methods for the analysis of longitudinal data need to incorporate the

| Table 1. Types of longitudinal studies. |
|----------------------------------------|
| Type of study | Number of individuals (N) | Number of visits (V) | Time lag (T) |
|---------------|---------------------------|---------------------|--------------|
| Panel         | Moderate                  | Large               | Small        |
| Longitudinal  | Large                     | Small               | Moderate or large |
| Laboratory    | Small                     | Small or moderate   | Small        |

This manuscript was prepared as part of the Environmental Epidemiology Planning Project of the Health Effects Institute, September 1990 – September 1992. The author thanks the members of the Health Effects Institute Working Group on Tropospheric Ozone, John Tukey, and three anonymous reviewers for helpful comments and suggestions.
intrinsic correlation of the repeated measurements within each individual. Since \( T \) (i.e., \( T/V \)) is typically 6 months or a year, the likelihood of incomplete profiles is high. Profiles are incomplete due to permanent or temporary dropouts. Permanent dropouts are those who have data missing after the last time they were seen in the study, and temporary dropouts are those with intermediate visits missing. The statistical methods for analysis need to include incomplete profiles in an efficient manner.

Longitudinal studies provide data to measure the changes of the outcome at the individual level and to relate those changes to the individual’s exposure over time. Methods of analysis can be classified according to the extent of the parametrization of the functional form of the outcome over time. Fully parametric models use polynomials (of which linear regression is the simplest case) or exponential or logarithmic functions to model the growth curve of individuals over time. Time series analysis models (e.g., autoregressive processes) are not as restrictive, but they assume a specific form of the correlation structure. Fully nonparametric models include smoothing algorithms and graphical summaries of the outcome data over time.

The remainder of this paper is divided into three sections. The first and second sections discuss issues related to studies of acute and chronic effects, respectively. The third section discusses issues common to both studies.

Acute Effects Studies

Although there have been several studies documenting acute (i.e., transient) health effects of exposure to \( O_3 \), there is a need to carry out further studies to determine the full range of acute outcomes.

The outcome of interest could be either continuous (e.g., FEV\(_1\)) or dichotomous (e.g., symptoms), with the main interest being the investigation of the effects of \( O_3 \) concentrations on the levels and variability of the outcomes over time.

Exposure Measurement

The geographic extension of study populations of acute studies is usually of limited range, so that the ambient concentrations of \( O_3 \) to which the population is exposed at a given time, \( t \), is relatively homogeneous. Let \( E_t \) denote the \( O_3 \) concentration at time \( t \) of the geographic area where the study population lives. The major contributor to the differential exposure for different individuals is the pattern of indoor/outdoor and exercise activities. Let \( P_t \) denote the pattern of indoor/outdoor and exercise activities of the \( i^{th} \) individual at time \( t \). It is the combination of \( E_t \) and \( P_t \) that provides the basis for calculating the exposure of the \( i^{th} \) individual at time \( t \) (\( E_{it} \)). Under the current technological limitations for directly measuring \( E_{it} \), studies of acute effects of \( O_3 \) need to devote special care in measuring \( P_t \). Individuals whose \( P_t \) does depend on \( t \) (i.e., pattern of indoor/outdoor and exercise activities is not the same for all \( t \)) provide data on individual changes of exposure even if the regional exposure, \( E_t \), is constant over time. Conversely, even if \( P_t \equiv P_i \) (i.e., individual \( i \) has the same pattern of activities for all \( t \)), the changes in \( E_t \) will result in \( E_{it} \) depending on \( t \). Studies where \( E_t \) and \( P_t \) are constant (i.e., do not depend on \( t \)) for all individuals, \( i \), are of limited utility because they reduce to individuals with patterns of activities that make some of them exposed and others unexposed. Although, in principle, this difference in exposure provides the basis for testing its effect on health outcomes, the confounding between high activity and favorable outcome may intrinsically preclude the detection of the putative effect of \( O_3 \) exposure.

Besides the central issues related to the elements needed to determine exposure at the individual level, epidemiological studies attempting to assess the effect of ozone exposure on health need to collect extensive and detailed data on other variables that could positively or negatively confound or modify the effect of interest.

Outcomes

The outcome (continuous or dichotomous) of acute effects studies is longitudinal in nature. The repeated measurements of the outcomes provide the basis for assessing the changes on an individual over time. Correlation of these changes with the changes in the exposure at the individual level should be a central objective of studies designed to investigate the health effects of ozone exposure.

Design and Analytical Approaches

It is important that, at the planning stage of a longitudinal study, the investigators incorporate the correlation structure of the data for the estimation of the sample size needed to detect the differences of interest. It is not always appropriate to base the calculation of sample size on expected cross-sectional differences.

For the case of a continuous outcome, the simplest model incorporating the correlation of the outcome over time corresponds to the linear model with a random intercept. Specifically, if there are \( n_0 \) unexposed and \( n_1 \) exposed individuals, the outcome \( Y_{it} \) is modeled as \( \alpha + \Delta_0 t + \epsilon_{it} \) for \( 1 \leq i \leq n_0 \)

\[
\alpha + \Delta_1 t + \epsilon_{it}
\]

for \( (n_0 + 1 \leq i \leq n_0 + n_1) \),

where \( \epsilon_{it} \) are normally distributed with mean zero, variance \( \sigma^2 \), and the within-individual correlation of \( \rho \). Thus, the variance of the mean of any indefinitely large number of observations for each individual (i.e., the between-individuals variance) will be \( \rho \sigma^2 \), while the variance of each observation about the population regression line for an individual (i.e., the within-individual variance) will be \( (1 - \rho)\sigma^2 \) and all individuals in a group will have the same slope. The main hypothesis of interest is whether \( \Delta_0 = \Delta_1 \) (e.g., decline of FEV\(_1\) is the same on individuals unexposed and exposed to \( O_3 \)).

Standard procedures of generalized least squares methods show that the individual coefficients \( \Delta_t \) and \( \Delta_0 \) have standard errors

\[
SE(\Delta_t) = SE(\Delta_0) = \sigma(1 - \rho)^{\frac{1}{2}} \left( \sum_{i=0}^{T} (t - \frac{V}{2}) \right)^{\frac{1}{2}}
\]

The asymptotic power to detect a difference \( \Delta_t = \Delta_0 \) at 5% level having available \( n_0 \) unexposed and \( n_1 \) exposed individuals is given by

\[
\Phi \left( (\Delta_t - \Delta_0) \left( n_0 n_1 \sum_{i=0}^{T} \left(t - \frac{V}{2}\right)^2 \right)^{\frac{1}{2}} \right)
\]

\[
\left( \sigma^2 (1 - p(n_0 + n_1)) - 1.96 \right),
\]

so the power is directly related to the magnitude of the difference \( (\Delta_t - \Delta_0) \), the number \( V \) of repeated observations on each individual, and the correlation \( \rho \) between the repeated measurements as well as the sample sizes of both groups. The power is inversely related to the variance \( \sigma^2 \).

Methods to compute the power of studies allowing for the slopes \( \Delta_0 \) and \( \Delta_1 \) to be different for different individuals are needed (1). Similarly, there is need to develop methods to calculate the power of models that directly relate the changes in exposure over time with changes in outcome.

Investigators carrying out longitudinal studies for the investigation of acute effects
of ozone exposure need to use analytical methods that focus on the quantification of the changes in outcome associated with changes in exposure. Under the assumption that the effects are acute and immediately disappear if the exposure is not present, marginal models treating the within-individual correlation as a nuisance are appropriate and particularly attractive since robust methods of estimation have been developed and are readily available (2, 3). If the effects are purely acute, then the comparison of an exposed individual with another, unexposed individual does not need the incorporation of the previous history of the exposure of those individuals. In this case, a cross-sectional design is sufficient, and if longitudinal data are available, the task of the longitudinal analysis consists of combining the different cross-sections or visits into an overall estimate of the effect of the exposure. It is for this reason that the robust methods for combining correlated cross sections are particularly useful.

For many biological processes the effect of an exposure is not purely acute. An exposure needs to be present for a certain amount of time for the outcome of interest to change, and once the exposure is absent it takes a certain amount of time for the outcome to clear the past effects of the exposure. An acute effect can be thought of as one for which current exposure is more important than previous exposures. Statistical procedures allowing past exposure to have a dampening effect (i.e., the further away the exposure, the lower the effect on current outcome) have been proposed and should be explored when analyzing longitudinal data (4). Autoregressive models not only provide an approach to incorporate the correlation structure of the within-individual measurements, they also incorporate the effect of previous exposure on the outcome of interest. In particular, the simplest autoregressive model of the form

\[ Y_{it} = \alpha + \gamma Y_{i,t-1} + \beta E_{it} + \epsilon_{it}, \]  

\[ Y_{it} = \alpha(1 - \gamma)(1 - \gamma^{-1}) \]

\[ + \gamma Y_{it-1} + \beta \sum_{k=0}^{\infty} \gamma^k E_{i,t-k} + \epsilon_{it} \]

with \( \epsilon_{it} \) independent and identically distributed as normal with mean 0 and variance \( \sigma^2 \), yields \( Y_{it} \) in terms of the baseline \( Y_{i0} \) and the exposure history \( E_{it} \) for \( 1 \leq i \leq t \) as

\[ \gamma < \gamma < 1, \text{ and hence the coefficient of exposures at times prior to the current value decrease as the time lag increases. This may not necessarily be the case for certain outcome-exposure associations, and alternative analytical approaches should be employed. On the other hand, it offers a simple way to model a mechanism that is expected in many biological processes.} \]

The above models are useful for the analysis of continuous outcome (e.g., forced expiratory volume after one second). If the outcome is binary (e.g., presence of asthma), appropriate methods using logistic regression models should be used (5). In this case, one models the log of the odds of having an asthma attack in the next examination on the basis of the presence of asthma in the current examination and the current environmental exposures. Another important aspect of acute studies of the health effects of ozone is the identification of susceptible subgroups. A susceptible individual is the one who changes more relative to another under the same change in exposure to ozone. In other words, a susceptible individual may tend to have higher within variance. Experimentation with high within variance as a selector of sensitivity may be well justified, but we cannot be certain that it will work. Longitudinal data provide the basis for the estimation of the within-individual variance, and analytical methods using the within variance as the outcome should be explored.

**Chronic Effects Studies**

Only a few epidemiological studies have been done on the chronic effects of ozone exposure. More studies are needed, and they should be designed to distinguish between acute/transient effects and those that have a long term effect on premature aging of the lung, symptoms, and mortality.

**Exposure**

In contrast to the acute effects studies, the studies to investigate chronic effects should pay special attention to cumulative exposures. Methods to summarize the duration and intensity of long-term exposure have been extensively studied in occupational epidemiology (6) and should be useful in studies designed to investigate the chronic health effects of ozone exposure.

Given the relatively homogeneous ozone concentration on a limited geographic area, it is important that locations with different histories of ozone concentrations be studied and compared. Obviously, valid inferences require the use of populations comparable to each other except for the exposure to ozone.

*Studies should be appropriately designed to provide comparable populations.*

It is difficult to reconstruct the exposure \( E_{it} \) of the \( i \)th individual at the time \( t \) in the past. It requires that the regional ozone concentration \( E_t \) at time \( t \) in the past be available and that the individual provide an accurate history of his or her pattern of activities \( P_{it} \) at time \( t \) in the past. In the absence of these data, it would be very useful to use markers \( M_t \) of the cumulative exposure to ozone on the \( i \)th individual. The availability of these markers may provide the basis of simple cross-sectional comparisons.

**Outcomes**

Studies investigating the chronic effects of ozone exposure should compare cross-sectional means and patterns of change of the outcome over time. The simplest measure of change over time is the difference between two time points \( Y_{it} - Y_{i,t-1} \). Recent work by Ware et al. (7) discusses appropriate methods for statistical analysis and use change as the outcome variable. Alternative methods (4) focus on the difference between the observed outcome at a given time \( t \) and the expected value of the outcome given the past values of the outcome

\[ (Y_{it} - E [Y_{it} | Y_{i,t-1}] ) \]

Both methods attempt to identify the variables that explain the variability of the changes over time. Of primary importance is the test of the effect of the exposure to ozone after adjusting for confounders and other known explanatory variables.

An alternative outcome of interest may be the within variance of the outcome on individual \( i \). It may be that the chronic effect of ozone is to make the outcome very susceptible (or volatile) to a given exposure. This approach is close to the challenging experiments (e.g., histamine) done in the area of respiratory disease epidemiology. In this case, the response to a challenge is used as an outcome as opposed to the usual setting of investigating its effect on pulmonary function. It could very well be that individuals with a long-term exposure to ozone are more reactive to a challenging exercise.

**Design and Analytical Approaches**

A central purpose of a study design to investigate the chronic effects of exposure to ozone is to obtain populations that are comparable to each other except for the exposure to ozone. One approach is to design a study where all individuals have been under the same regional exposure (i.e., close geographic location) but whose patterns of
indoor/outdoor and exercise activities are diverse. An advantage of this design is that
the regional nature of the study population makes the individuals comparable in several
respects, including coupling with exposures to other pollutants. The main difficulty is
that very active individuals (i.e., more likely to be exposed to ozone) may have associated
a favorable outcome. Another approach is to select different locations with different
histories of ozone concentrations and compare the outcome of groups of individuals
from the different locations. The main difficulty here is that ozone elevations are typi-
cally coupled with other environmental exposures, making the effect nonidentifiable.
Furthermore, individuals in different geographic locations intrinsically may have dif-
ferent patterns of activities, making it difficult to distinguish the independent
effects of ozone exposure.

A design that may offer some advantage could involve the comparison of groups of
individuals in different locations with different histories of ozone elevations but
could match the individuals in different locations by their pattern of indoor/out-
door and exercise activities. The objective would be that individuals who have the
same pattern of activities but are subject to different ozone concentrations may provide
different outcomes.

A very important aspect of the analysis of data from studies investigating the long-term
effects of ozone exposure is the identification of patterns of exposure with outcome. The
main components for the determination of pattern of exposures are the duration and
intensity of the exposure. Much effort has been dedicated to advantages and disadvan-
tages of different summary measures, including the maximum concentration, the integral
of the concentrations over time, the weighted mixture of different concentrations at dif-
ferent times, etc. Attention should also be given to patterns incorporating transitions of
past exposures (8).

Analytical methods for longitudinal data have been the subject of active statistical
research in the last decade. The emphasis has been on how to incorporate the correla-
tion structure of the repeated observations on a given individual. Robust methods, ran-
dom effects, and autoregressive models correspond to different ways of handling the
within-individual correlation of the outcomes taken at different visits. An important
result is that when modeling the cross-sectional (e.g., marginal) distribution of the
outcome, the estimation of the regression coefficient should not be affected by the
method of handling the correlation of the
within-individual outcome measures. Discrepancies between estimates of the
regression coefficients when using different methods for the incorporation of the correla-
tion of the outcome over time could be used as a diagnostic regression measure of incon-
sistencies of the between- and within-individual regressions. If the effect of exposure
using individual regressions (i.e., changes in exposure to changes in outcome) are of lesser
magnitude than the effect of exposure using the between-individual regressions, a chronic
(i.e., long-term) effect may be suggested. Therefore, a diagnostic regression procedure
may be useful for the identification of chronic versus acute effects. Refinement and
specificity of the diagnostic tool is needed.

Issues Common to Acute and
Chronic Studies

The general issues of confounding and effect modification are present in both acute
and chronic studies of the health effects of exposure to ozone. Since elevated ozone
concentrations are correlated with elevations of other pollutants and environmental
conditions, it is important to obtain complete information so that controlling for
confounding can be done with the appropriate analytical procedures. The investiga-
tions of interactions between ozone and other pollutants that cause adverse health
effects are of equal importance. It is possible that only situations where critical levels of
other atmospheric pollutants occur are associated with a poor outcome.

Longitudinal studies are bound to have missing data on intermediate visits or on the
last visit for those individuals who drop out of the studies. Analytical methods are
available (9) to incorporate into the analysis visits that are not equidistant due to
gaps caused by intermediate visits missing. The missing information on individuals
who permanently drop out is of greater impact because they directly affect the abil-
ity to assess the long-term effect of the exposure. Investigators in other areas of
research have used multiple imputation (10) for the handling of missing data. Use
and applicability of these methods is an area of important research in the context of
the health effects of ozone exposure. The methods of multiple imputation typically
assume the missing data to be caused by a random mechanism. In studies of health
effects of ozone, the dropouts may be related to disease progression and alterna-
tive methods for the incorporation of informative censoring will need to be developed.

Given the measurement errors to which both the exposure and the outcome are sub-
ject, it is essential that replicate measurements be taken if feasible. Regression meth-
ods incorporating the measurement error have been developed and should be used.
These methods require data on duplicates to estimate the error variance. Failure to cor-
rect for the measurement error may increase the probability of not rejecting the null
hypothesis when the alternative is true.

Both acute and chronic studies need to incorporate data on treatment of chronic
respiratory illnesses (e.g., asthma) and interventions. Studies are needed on the
effect of treatment under different exposures to ozone. As with exposure, the treat-
ment may also be time dependent and, thus, the analytical issues are similar.

During the last decade, methods have been proposed to combine studies to pro-
vide an overview (i.e., metaanalysis) or to combine studies with strengths in comple-
mentary aspects. In particular, the prevalent and incident subcohorts of a longitudinal
study in infectious disease epidemiology describe the mature and early stages of the
natural history. Several approaches (11,12) have been proposed to combine these com-
ponents into a unified data set for the determination of the incubation period of AIDS.
Although of a different nature, acute and chronic effects are complementary. The
outcome of acute studies can be viewed as the exposure of chronic studies. Acute stud-
ies may establish that exposures to ozone are associated with acute/transient changes of
outcome. Chronic studies may establish that individuals whose outcome has a high
within variance are those who will prematurely age with respect to the outcome of
interest. Methods to combine studies of this nature should be the subject of future
research. The issues presented here have a role similar to that of surrogate markers for
the evaluation of effective therapies in infectious disease epidemiology. Bridging the
methodological issues will be an important contribution to scientific research.

In the context of the studies of the health effects of ozone, there are opportuni-
ties to combine laboratory experiments with longitudinal studies. Chamber studies
provide a measure of the changes in outcomes due to a controlled exposure. Using
multivariate methods (e.g., principal components), one can determine the indi-
viduals with the highest variability and enroll them in a follow-up study to assess the
effects of environmental exposure. The analysis of these data will be more informa-
tive if the studies are formally combined into a comprehensive analysis.
REFERENCES

1. Laird NM, Ware, JA. Random effects model for longitudinal data. Biometrics 38:963–974 (1982).
2. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 73:13–22 (1986).
3. Zeger SL, Liang KY. Longitudinal data analysis and discrete and continuous outcomes. Biometrics 42:121–130 (1986).
4. Rosner B, Muñoz A, Tager I, Speizer F, Weiss S. The use of an autoregressive model for the analysis of longitudinal data in epidemiologic studies. Stat Med 4:457–467 (1985).
5. Korn EL, Whittemore AS. Methods for analyzing panel studies of acute health effects of air pollution. Biometrics 35:795–802 (1979).
6. Checkoway H, Pearce NE, Crawford-Brown DJ. Research methods in occupational epidemiology. New York: Oxford University Press, 1989.
7. Ware JH, Douglas WD, Louis TA, Xiping XU, Ferris BG Jr, Speizer FE. Longitudinal and cross-sectional estimates of pulmonary function decline in never-smoking adults. Am J Epidemiol 132:685–700 (1990).
8. Vlahov D, Muñoz A, Anthony JC, Cohn S, Celentano DD, Nelson KE. Association of drug injection patterns with antibody to human immunodeficiency virus Type-1 among intravenous drug users in Baltimore. Am J Epidemiol 132:847–856 (1990).
9. Rosner B, Muñoz A. Autoregressive modelling for the analysis of longitudinal data with unequally spaced examinations. Stat Med 7:59–71 (1988).
10. Little R, Rubin D. Statistical analysis with missing data. New York: John Wiley and Sons, 1987.
11. Muñoz A, Wang MC, Bass S, Taylor JMG, Kingsley LA, Chmiel JS, Polk BF. Acquired Immunodeficiency Syndrome (AIDS)-free time after Human Immunodeficiency Virus type 1 (HIV-1) seroconversion in homosexual men. Am J Epidemiol 130:530–539 (1989).
12. Taylor JMG, Muñoz A, Bass SM, Chmiel JS, Kingsley LA, Saah AJ. Estimating the distribution of times from HIV seroconversion to AIDS using multiple imputation. Stat Med 9:505–514 (1990).