HEI HEALTH EFFECTS INSTITUTE

Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies

Part X: Robust Composite Scores Based on Median Polish Analysis

Paul J. Catalano, John Rogus, and Louise M. Ryan
Harvard School of Public Health and Dana-Farber Cancer Institute, Boston, MA

Includes the Commentary of the Institute's Health Review Committee

Research Report Number 65
January 1995

HEI HEALTH EFFECTS INSTITUTE

The Health Effects Institute, established in 1980, is an independent and unbiased source of information on the health effects of motor vehicle emissions. HEI studies all major pollutants, including regulated pollutants (such as carbon monoxide, ozone, nitrogen dioxide, and particulate materials) and unregulated pollutants (such as diesel engine exhaust, methanol, and aldehydes). To date, HEI has supported more than 120 projects at institutions in North America and Europe.

Typically, HEI receives half its funds from the U.S. Environmental Protection Agency and half from 28 manufacturers and marketers of motor vehicles and engines in the United States. Occasionally, revenues from other public or private organizations either support special projects or provide resources for a portion of an HEI study. For this study, the Institute acknowledges the cooperation and support of the National Toxicology Program (NTP), which consists of four charter agencies of the U.S. Department of Health and Human Services. The NTP sponsored the inhalation component of this project as part of its studies on the toxicologic and carcinogenic effects of ozone. However, in all cases HEI exercises complete autonomy in setting its research priorities and in disbursing its funds. An independent Board of Directors governs the Institute. The Research Committee and the Review Committee serve complementary scientific purposes and draw distinguished scientists as members. The results of HEI-funded studies are made available as Research Reports, which contain both the Investigators' Report and the Review Committee's evaluation of the work's scientific and regulatory relevance.

HEI Statement

Synopsis of Research Report Number 65 Part X

An Innovative Approach to Analyzing Multiple Experimental Outcomes: A Case Study of Rats Exposed to Ozone

BACKGROUND

One major component of urban smog is ozone, a highly reactive gas that forms when emissions from mobile and industrial sources react chemically in the presence of sunlight. Because ozone can damage cells, prolonged or repeated exposures may be a risk factor for lung cancer. Therefore, the National Toxicology Program (NTP) conducted a bioassay of prolonged exposure to evaluate ozone's carcinogenicity in rodents.

Another concern is that prolonged ozone exposure could cause noncancerous lung diseases such as fibrosis and emphysema. The NTP's bioassay project presented a unique opportunity for a collaboration between the HEI and the NTP. The HEI funded eight studies to investigate whether rats exposed to 0, 0.12, 0.5, or 1.0 parts per million (ppm) ozone, for 6 hrs/day, 5 days/week, for 20 months, developed alterations in their respiratory systems that were characteristic of chronic respiratory diseases. This report by Dr. Catalano and colleagues describes the procedures that were used to analyze, interpret, and correlate the data from all of these studies.

During the course of the laboratory studies and after they were completed, the investigators and biostatisticians grouped multiple endpoints from individual studies to examine the effects of ozone exposure for three respiratory disease surrogates: centriacinar fibrosis (defined as fibrosis at the junction between the conducting airways and the alveolar gas exchange area), airway disease (defined as conditions that contribute to a decreased ability of the conducting airways to transport air to and from the alveoli), and chronic rhinitis (defined as inflammation and irritation of the nasal passages). Chronic rhinitis was evaluated through endpoints from a single study, whereas centriacinar fibrosis was assessed via endpoints from four studies, and airway disease from endpoints from five studies.

APPROACH

Drs. Paul Catalano, Louise Ryan, and colleagues used a statistical method called median polish analysis to evaluate jointly the multiple endpoints selected for each disease surrogate. The statisticians chose median polish analysis because each endpoint was not measured on all rats and this technique accommodates missing data points. Using the median polish technique, they first developed composite scores that summarized the selected endpoints for each individual animal. Then they used these scores to analyze ozone concentration–response trends, and differences between rats exposed to ozone and control rats.

RESULTS AND IMPLICATIONS

The clearest changes induced by ozone were seen in the evaluation of chronic rhinitis. A statistically significant trend was noted toward increased responses with increasing ozone concentrations, and the differences between measurements in control rats and rats exposed to 0.5 or 1.0 ppm ozone also were statistically significant. Although marginally significant or significant trends toward increased responses were seen for centriacinar fibrosis and airway disease, the differences between control rats and rats exposed to ozone were not statistically significant. Furthermore, the results for centriacinar fibrosis depended on which endpoints were included in the analysis.

Median polish analysis has not been applied before to the type of data analysis required by the NTP/HEI Collaborative Ozone Project; therefore, it must be considered an experimental method. Problems associated with the use of median polish analysis include the need to select, a priori, the appropriate experimental variables to include in the analysis, and the expected direction (increase or decrease) of their change. These problems were reflected in the results and are discussed in the HEI Health Review Committee's Commentary. Nevertheless, the use of median polish analysis was an innovative approach that was carefully applied by Dr. Catalano and colleagues. They obtained results for some disease surrogates that appear to have biologic meaning and identified significant differences across ozone exposure concentrations. Thus, the method may find wider use in similar analyses of toxicologic data.

This Statement, prepared by the Health Effects Institute (HEI) and approved by its Board of Directors, is a summary of a research project sponsored by HEI from November of 1990 through May of 1994 . This study was conducted by Dr. Paul Catalano and associates. The following Research Report contains an Introduction to the NTP/HEI Collaborative Ozone Project, the detailed Investigators' Report, and a Commentary on the study prepared by the Institute's Health Review Committee.

Library of Congress Catalog Number for the HEI Research Report Series: WA 754 R432.
The paper in this publication meets the minimum standard requirements of the ANSI Standard Z39.48-1984 (Permanence of Paper) effective with Report Number 21, December 1988, and with Report Numbers 25, 26, 32, 51, and 65 Part IV excepted. Reports 1 through 20, 25, 26, 32, 51, and 65 Part IV are printed on acid-free coated paper.

TABLE OF CONTENTS

Research Report Number 65

Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies Part X: Robust Composite Scores Based on Median Polish Analysis

Paul J. Catalano, John Rogus, and Louise M. Ryan

INTRODUCTION

The National Toxicology Program and Health Effects Institute Collaborative Ozone Project

The NTP/HEI Collaborative Ozone Project was a four-year project that was organized to evaluate the effects of prolonged ozone exposure on lung injury in animals. The ozone exposures were conducted by the National Toxicology Program (NTP) at Battelle Pacific Northwest Laboratories. Eight groups of investigators addressed the pathologic and physiologic consequences of prolonged ozone exposure, supported by the Health Effects Institute (HEI). A full description of the NTP/HEI Collaborative Ozone Project and the exposure protocol can be found in the Introduction and Supplement to Research Report Number 65 Part I. This information also will be published in Part VI of Research Report Number 65.

Briefly, in 1987, the Health Effects Institute entered into a partnership with the National Toxicology Program to evaluate the effects of chronic ozone exposure in rats. The NTP, consisting of four agencies of the U.S. Department of Health and Human Services, coordinates the nation's testing of potentially toxic and hazardous chemicals. The Health Effects Institute, an independent research organization supported by both government and industry, provides unbiased information on the health effects of motor vehicle emissions.

Because of the widespread exposure to ozone and concerns about its potential health effects, HEI and the California Department of Health and Human Services nominated ozone for carcinogenicity and toxicity testing by the NTP. The NTP, recognizing that cancer was only one of the chronic diseases of concern, included additional animals for HEI-supported studies of the pathologic and physiologic consequences of prolonged ozone exposures. The HEI animals were housed in cages that would otherwise have been empty. By developing a partnership, the HEI and NTP were able to leverage their funds to develop a comprehensive research program that extended beyond carcinogenicity endpoints; the HEI-sponsored research focused on the relation between long-term ozone exposure and the pathogenesis of chronic lung diseases, such as asthma, emphysema, and fibrosis. The Health Effects Institute would not have been able to undertake such an expensive project, which requires special facilities and trained personnel, without the NTP's support of the inhalation component and the cooperation of the NTP's contractor, Battelle Pacific Northwest Laboratories.

For the HEI component of the Project, eight studies were selected for funding from proposals submitted in response to the Request for Applications (RFA) 90-1, Health Effects of Chronic Ozone Inhalation: Collaborative National Toxicology Program–Health Effects Institute Studies, Part A: Respiratory Function Studies, and Part B: Structural, Biochemical, and Other Alterations. Because of the complexity of a project with many investigators and many endpoints, the HEI Health Research Committee also funded a Biostatistical Advisory Group to provide assistance with experimental design, animal allocation, and data analyses. Figure 1 presents a diagram of the studies in the NTP/HEI Collaborative Ozone Project and their relations to each other. They include those studies that were part of the NTP bioassay, the eight HEI-funded studies, and the biostatistical study. In addition, HEI engaged Battelle Pacific Northwest Laboratories to provide support services for the HEI-sponsored investigators.

Starting at six to seven weeks of age, male and female F344/N rats were exposed to 0, 0.12, 0.5, or 1.0 parts per million (ppm) ozone, six hours per day, five days per week. These concentrations were selected to include the maximum concentration the animals would tolerate (1.0 ppm), the current National Ambient Air Quality Standard (NAAQS) for ozone (0.12 ppm), and an intermediate concentration. The NTP's carcinogenicity bioassay consisted of a two-year study and a lifetime study in rats and mice, and a study of male rats exposed to 0.5 ppm ozone and two levels of a human pulmonary carcinogen, 4-(*N*-methyl-*N*-nitrosamino)-1-(3-pyridyl)-1-butanone (NNK). The design of the HEI studies was directed, to some extent, by the constraints of the NTP protocol. These included ozone exposure concentrations that were set by the NTP, a limit on the sample size (164 rats) to the number of available exposure chambers, and quarantine restrictions that did not allow reentry of animals into the exposure chambers once they had been removed, thus eliminating the possibility of conducting serial tests.

The Biostatistical Advisory Group developed a sample allocation scheme that allowed several researchers to obtain measurements on tissue samples from the same subset of study animals, providing the maximum overlap of animals and tissues among the eight studies while ensuring balance with respect to dose, gender, and time of death. When the ozone exposure of the HEI animals ended (at 20

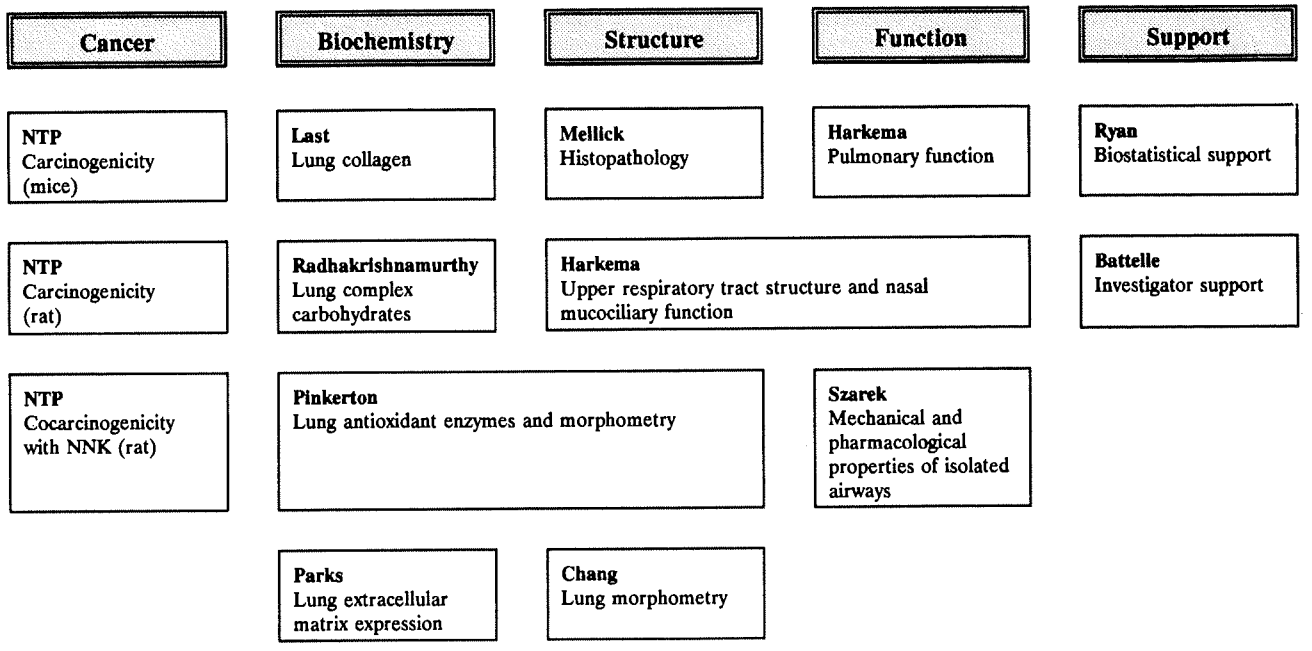

Figure 1. The NTP/HEI Collaborative Ozone Project: individual studies.

months), several investigators traveled to Battelle Pacific Northwest Laboratories to conduct their assays or to obtain samples on site. Battelle personnel prepared the tissues for off-site investigators and shipped them directly to their laboratories.

Because the studies varied in duration from six months to two years, HEI is publishing the reports for each individual study after the Institute's review process for each study is complete. Each Investigator's Report and a forthcoming Integrative Summary Report will be Parts of Report Number 65 of the HEI Research Report series.

The present report by Dr. Paul Catalano and colleagues, Part X, introduces the statistical methods used by the Biostatistical Advisory Group. The report continues by assessing and quantifying the multiple endpoints analyzed by the individual investigators in the NTP/HEI Collaborative Ozone Project, and explores the relation between the different types of effects caused by long-term ozone exposure on lung function and on the nasal mucociliary apparatus (Harkema), airway reactivity (Szarek), structural (Pinkerton, Chang), or biochemical (Last, Radhakrishnamurthy, and Parks) alterations. Many of these analyses are also found in Part XI of Research Report Number 65, which presents an Integrative Summary of the Collaborative Project.

The importance of the collaborative NTP and HEI chronic ozone exposure studies is that they provide an unparalleled opportunity to examine the effects of prolonged ozone exposure using a variety of scientific approaches. The interaction of a number of methods to analyze the pathologic and physiologic consequences of chronic ozone exposure is one of this project's unique features. The results of these studies will provide new information about the threshold effects of ozone exposure on lung injury and the type and extent of damage in a well-established animal model. These results may be helpful for evaluating current standards of ozone exposure as they apply to human health and for designing future animal and human studies.

Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies

Part X: Robust Composite Scores Based on Median Polish Analysis

Paul J. Catalano, John Rogus, and Louise M. Ryan

ABSTRACT

This report describes some of the statistical methods used to analyze data from the National Toxicology Program/Health Effects Institute Collaborative Ozone Project. The purpose of the collaborative study was to assess the health effects of chronic ozone inhalation. Data were obtained from a subset of 164 F344/N rats dedicated to use by the Health Effects Institute from a standard ozone inhalation study conducted by Battelle Pacific Northwest Laboratories for the National Toxicology Program. The study involved eight groups of investigators, each assessing different types of ozone-related health effects. These included studies of respiratory function and of structural, cellular, and biochemical changes in the lungs and airways.

Designing and analyzing a study with several groups of investigators raises many statistical challenges. The highest design priority for this study was that each investigation be individually interpretable as an independent study. This meant that each investigator had to receive an adequate number of animals, balanced with respect to level of ozone exposure and other factors such as the gender of the rats and the time they were killed. Another feature of the collaborative study was the opportunity it provided to assess and quantify the effect of ozone exposure on a broad spectrum of endpoints, and to explore the relations between the different types of effect. Maximizing the potential to assess these correlations required that the individual animals studied by the different groups of investigators overlap as much as possible. This aspect of the statistical design required careful consideration of the compatibility between various investigations. Fortunately, the degree of compatibility was substantial. In many cases, for example, it was possible to assess respiratory function in the animals before they were killed, and then to divide the tissue among several different investigators.

This report concentrates on the methods that were specially developed to analyze the data for multiple endpoints collected in the study. Nonstandard techniques were required to accommodate the complex pattern of missing data that was inherent in the study design because no animals were measured by all investigators.

INTRODUCTION

Ozone, a major component of photochemical smog, is one of six criteria air pollutants monitored by the U.S. Environmental Protection Agency. It is estimated that 50 to 70 million people live in areas that are out of compliance with the National Ambient Air Quality Standard for ozone of 0.12 parts per million (ppm)* averaged over one hour, not to be exceeded more than once per year (U.S. Office of Technology Assessment 1988; American Lung Association 1989). Thus, the impact on humans of chronic exposure to ozone is an important research topic. Many studies have investigated short-term exposures of rodents and other species to ozone, but little is known about the effects of chronic ozone exposure (Witschi 1988).

To assess the adverse health effects of long-term exposure to ozone, in 1989 the National Toxicology Program (NTP) began studies of chronic ozone inhalation in rats with planned exposures of either 24 or 30 months. To gain maximum information from these investigations, the NTP established a unique collaborative arrangement with the Health Effects Institute (HEI). This collaboration provided an opportunity for several researchers to obtain measurements and tissue samples from a subset of the study animals. The arrangement allowed each group of investigators

* A list of abbreviations appears at the end of the Investigators' Report.

This Investigators' Report is one section of Part X of Health Effects Institute Research Report Number 65, which also includes an Introduction to the NTP/HEI Collaborative Ozone Project, a Commentary by the Health Review Committee on the Investigators' Report, and an HEI Statement about the research project. Correspondence concerning the Investigators' Report may be addressed to Dr. Paul J. Catalano, Division of Biostatistics, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115.

This study was supported by HEI funds from the U.S. Environmental Protection Agency and the motor vehicle industry. The inhalation component of this project was sponsored by the National Toxicology Program as part of its studies on the toxicologic and carcinogenic effects of ozone.

Although this document was produced with partial funding by the U.S. Environmental Protection Agency under Assistance Agreement 816285 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and therefore may not necessarily reflect the view of the Agency, and no official endorsement should be inferred. The contents of this document also have not been reviewed by private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views or policies of these parties, and no endorsement by them should be inferred.

to conduct individually interpretable research while providing an opportunity to integrate data from several evaluations of the same animals in order to gain a more comprehensive understanding of the effects of chronic ozone exposure on the respiratory system. Details of this Collaborative Project have been published in other parts of Research Report Number 65 (Harkema and Mauderly 1994; Harkema et al. 1994; Last et al. 1994; Radhakrishnamurthy 1994; Szarek 1994; Boorman et al. 1995; Catalano et al. 1995; Chang et al. 1995; Pinkerton et al.1995).

From a statistical perspective, the design and analysis of the NTP/HEI Collaborative Ozone Project posed some unique challenges. For example, a primary feature of the study was that for each animal measurements were obtained by several different investigators. In some cases, this involved dividing lung tissue between several groups of investigators. In others, it involved in vivo lung function testing before the rats were killed with subsequent dissection and examination by other investigators. Thus, methods for analyzing multiple outcomes were needed. An important secondary feature was that owing to limitations of available tissue, as well as the logistics of supplying several different investigators, no single animal could be examined by all the investigators. The resulting data structure was therefore complex, with multiple endpoints and a complicated, preplanned pattern of missing data.

The study was designed with several analysis goals in mind. First, each distinct endpoint was to be analyzed individually. This meant that the design had to ensure that each study was assigned enough animals to address the questions of interest to that investigation, and that the sample had to be balanced between exposure groups, with respect to gender and any other factors that could affect outcome. Second, an integrative analysis would combine information across endpoints. It was felt that by synthesizing information across endpoints, it would be possible to gain a clearer understanding of the effects of ozone on these related health outcomes. The design aspects of the study are discussed in another HEI report (Boorman et al. 1995). The purpose of this report is to discuss the statistical methods that were specially developed for the integrative analysis.

This paper is organized into several sections. The next section describes our specific aim, which was to develop a statistical framework on which to base the integrative analysis of the NTP/HEI Collaborative Ozone Project. We discuss the conceptual ideas surrounding the integrative analysis and the evolution of these ideas over the course of the study. Then, we provide some general background regarding existing statistical methods for analyzing multiple outcomes and discuss why these methods did not adequately address the needs of our analysis. We follow with details of the method we devised. The concept is based on median polish analysis (for example, see Tukey 1977; Mosteller and Tukey 1977; Hoaglin et al. 1983). We show that the method can be adapted to provide a simple, yet sensitive way to integrate a set of related endpoints. The method is also robust in that it will accommodate a broad variety of situations including nonnormal distributions and missing data.

SPECIFIC AIMS

The NTP/HEI Collaborative Ozone Project design provided a unique opportunity to integrate the results from tests of several different but related endpoints. In the planning phase of the study, there was a conceptual notion of the nature and goals of the integrative analysis. A more detailed picture of its several distinct aspects did not emerge until the first stages of data collection. For example, we had hoped that the overlap of investigators using the same animals would facilitate the correlation of outcomes. However, preliminary analyses made it clear that the signal-to-noise ratio was low enough that any attempts to study correlations between outcomes would have to be largely exploratory. Some attempts were made to explore such relations, and these are described in separate reports, as well as in the overall integrative summary report (Catalano et al. 1995).

The integrative analysis can also be considered a means of protecting the study against Type I error. It is well known that if tests of multiple hypotheses are performed without any adjustment of the resulting p values, then false-positive results become likely. There are essentially two different ways to protect a study against Type I error. One way is through the application of p value correction factors, such as the Bonferroni adjustment and related methods. Another approach is through the use of multivariate analysis methods. Techniques like the Bonferroni adjustment require that if K tests are to be performed and the desired Type I error rate for the study is α percent, then each individual test must be conducted at a significance level well below α percent. There are several related ways to perform such corrections. The Bonferroni adjustment is probably the most extreme, requiring that each of the K individual tests be conducted at significance levels of α/K. The main problem with such adjustments is that they are quite conservative (less likely to detect a real effect), especially when data are correlated, as they are in the NTP/HEI Collaborative Ozone Project, and when a generalized effect on all the endpoints is expected.

It then became our goal to develop a suitable multivariate analysis technique that facilitated a synthesis of the data for several related endpoints, and that helped in making meaningful assessments of the concentration-response patterns. In particular, we were very interested in the fundamental question of whether or not any significant change in health effects was associated with chronic exposure to 0.12 ppm ozone (the current Standard). Owing to the large numbers of endpoints under consideration, we decided to combine them into groups that had some common clinical or health-related interpretation.

CHOICE OF ENDPOINTS

Extensive discussions with the project investigators led to the idea of investigating three major groupings of endpoints related to three different types of respiratory disease: centriacinar fibrosis, airway disease, and chronic rhinitis. To emphasize that the groupings of research endpoints do not fully represent these disease states, we refer to them as disease surrogates. Tables 1, 2, and 3 list the individual endpoints and corresponding sample sizes chosen for the analyses of these three disease surrogates. In this paper, we present summary results from the analyses to illustrate the median polish analysis method. The data will be interpreted and analyzed in the Integrative Summary (Catalano et al. 1995). Detailed discussion concerning the method's statistical properties and sensitivity will concentrate on the centriacinar fibrosis and airway disease surrogates, because they provide interesting case studies. Computer simulations will also be used to evaluate the method.

CENTRIACINAR FIBROSIS

Fibrosis refers to a general condition in which connective tissue becomes thickened and inelastic through the excess accumulation of fibrous material. From a clinical perspective, pulmonary fibrosis would be characterized by alveolitis in its early stages and destruction of alveoli in its later stages. For our purpose, fibrosis does not refer to a clinical condition, but rather to a set of conditions that indicate the process is occurring. The study included measurements of a variety of endpoints related to fibrotic development in the centriacinar region of the lung. We briefly discuss some of the reasons for considering the 10 endpoints which are listed in Table 1 with the available sample sizes within each exposure group.

Because excessive fibrous accumulation in the lungs very nearly defines pulmonary fibrosis, total lung collagen content (measured per lung lobe) is included in the set of fibrosis-related endpoints. Whole lung collagen content is a global measure of fibrous accumulation in the lungs obtained by biochemical analysis of an entire lung lobe. For this analysis, collagen was measured not only in the centriacinar region, but also in the conducting airways.

In addition to this measure, a morphometric assessment of collagen content in the centriacinar region (Chang et al.

Table 1. Endpoints for Analysis of Centriacinar Fibrosis: Sample Sizes by Exposure Group

			Ozone Exposure (ppm)			
Endpoint[a]	Code	Total Rats	0	0.12	0.5	1.0
Total collagen (Last)	COLT	42	12	6	12	12
Collagen volume morphometric (Chang)	COLM	39	11	12	8	8
Fibroblast volume (Chang)	FV	39	11	12	8	8
Interstitial volume (Pinkerton)	IV	39	11	12	8	8
Interstitial macrophage volume (Chang)	IMV	39	11	12	8	8
Glycosaminoglycans: chondroitin 6-sulfate (Radhakrishnamurthy)	C6S	32	11	4	7	10
Glycosaminoglycans: dermatan sulfate (Radhakrishnamurthy)	DS	32	11	4	7	10
Residual volume (Harkema and Mauderly)	RV	61	18	8	18	17
Quasistatic chord compliance (Harkema and Mauderly)	QCC	61	18	8	18	17
Carbon monoxide diffusing capacity (Harkema and Mauderly)	D_{CO}	61	18	8	18	17

[a] The first author of other Research Reports in the NTP/HEI Collaborative Ozone Project series is given in parentheses.

1995) provides a more location-specific measure of collagen. In cases of localized accumulation, this endpoint (if measured at the point of localized damage) is a more sensitive measure of fibrotic response.

Fibrosis is also often characterized by an increase in interstitial volume, possibly because fibrosis often involves quick and complete endothelial regeneration, but prolonged epithelial regeneration accompanied by metaplasia. In this analysis the morphometric assessment of interstitial volume, measured by Pinkerton and associates (1995), was included as an average of the individual cranial and caudal region measurements.

Fibroblasts are major producers of collagen and therefore are associated with aspects of fibrosis. Although there are exceptions to this finding, some studies have shown an association between excessive fibroblasts and the development of fibrosis. Fibroblasts were measured by Chang and associates (1995).

Similarly, previous research showed a connection between macrophage volume and fibrosis. The morphometric measure used in this analysis came from interstitial macrophage measurements also made in the study by Chang and associates (1995).

Glycosaminoglycans (GAGs), complex carbohydrates found in lung connective tissue, are produced by fibroblasts and play an important role in lung connective tissue biochemistry. For our analysis, two major GAG components, chondroitin 6-sulfate (C6S) and dermatan sulfate (DS), as measured by Radhakrishnamurthy (1994), were included.

Three measures of respiratory function, measured by Harkema and Mauderly (1994), were also included in the fibrosis analysis. Residual volume (RV) was chosen as one parameter because it is thought to be a sensitive measure of stiffening at the terminal bronchiole–alveolar duct junction. Fibrosis can be thought to cause airway stiffening deep in the lung, which decreases the volume of gas removed from the lung during the procedure to determine RV. Hence, a decrease in RV can be interpreted as a result of fibrosis. The two other functional parameters, quasistatic chord compliance (QCC) and carbon monoxide diffusing capacity (D_{CO}), are also thought to be somewhat sensitive to lung stiffening and were therefore included in the fibrosis model.

AIRWAY DISEASE

Airway disease refers to a set of conditions that contribute to a decreased ability of the conducting airways to transport air to and from the lungs. Table 2 lists the endpoints that were considered for airway disease. Some of these measurements were taken in the centriacinar region, and others were taken in the conducting airways.

Table 2. Endpoints for Analysis of Airway Disease: Sample Sizes by Exposure Group

Endpoint[a]	Code	Total Rats	Ozone Exposure (ppm)			
			0.0	0.12	0.5	1.0
Mean midexpiratory flow (Harkema and Mauderly)	MMEF	61	18	8	18	17
AB/PAS-positive material Central region stored (Pinkerton)	MUCT	32	8	8	8	8
AB/PAS-positive material Caudal region stored (Pinkerton)	MUCD	32	8	8	8	8
AB/PAS-positive material Cranial region stored (Pinkerton)	MUCR	32	8	8	8	8
Small airway smooth muscle (Szarek)	ASMS	34	9	5	8	12
Large airway smooth muscle (Szarek)	ASML	34	9	5	8	12
Percent bronchiolarization (Chang)	BRON	39	11	12	8	8
Epithelial distance (Pinkerton)	ED	39	11	12	8	8
Shift from ciliated to Clara cells (Chang)	CICL	39	11	12	8	8
Central region percent nonciliated cells (Pinkerton)	NCCT	32	8	8	8	8
Caudal region percent nonciliated cells (Pinkerton)	NCCD	32	8	8	8	8
Cranial region percent nonciliated cells (Pinkerton)	NCCR	32	8	8	8	8
Small airway maximal tension (Szarek)	TENS	34	9	5	8	12
Large airway maximal tension (Szarek)	TENL	34	9	5	8	12
Small airway EC_{50} (Szarek)	ECS	34	9	5	8	12
Large airway EC_{50} (Szarek)	ECL	34	9	5	8	12
Eicosanoid production (Szarek)	PGE_2	34	9	5	8	12
Epithelial volume (Pinkerton)	EV	39	11	12	8	8

[a] The first author of other Research Reports in the NTP/HEI Collaborative Ozone Project series is given in parentheses.

Mean midexpiratory flow (MMEF) is an appropriate measure of airway function because it reflects how well the animal is able to expel air from the lungs. This was one of the suite of respiratory function measurements taken by Harkema and Mauderly (1994).

Measures of stored mucosubstances are natural measures of airway disease because excessive mucus in the lung would be assumed to contribute to decreased airway function. We included three separate measures of stored material that stained positively with Alcian blue/periodic acid–Schiff (AB/PAS) stain in the central, caudal, and cranial regions of the lung. The percentage of bronchiolarization in the proximal alveolar region (Chang et al. 1995) was also included in this analysis.

Smooth muscle thickness in large (intrapulmonary bronchus) and small (branch) airways, as measured by Szarek (1994), was thought to be a relatively sensitive measure relating to airway disease. Several other measures from the Szarek (1994) study relating to airway responsiveness and eicosanoid production, a parameter related to the development of smooth muscle, were taken in distal airways and were included in addition to the airway smooth muscle data. Basal levels of eicosanoid production (mediator prostaglandin E_2 [PGE_2] level) were measured in distal airways. Airway responsiveness was parameterized using the mean maximal airway tension obtained by averaging the concentration-response maximum parameters (T_{max}) obtained from analysis of four contractile agonists. The average estimated concentration of an agonist that elicited a half-maximal response (EC_{50}), taken from the estimated concentration-response curves, was also used as an indicator of airway responsiveness. Separate data were used from both large and small airways.

Several morphometric measures such as epithelial volume (Pinkerton et al. 1995) and airway cell populations (Chang et al. 1995; Pinkerton et al. 1995) were also included in this analysis. The latter measure was parameterized as a ratio of Clara cell to ciliated cell densities (Chang et al. 1995) to reflect the expected shift in these cell types in exposed animals. The percentages of nonciliated cells in the central, cranial, and caudal regions (Pinkerton et al. 1995) were also included as three distinct measures. Because epithelial volume measures (Pinkerton et al. 1995) declined as a function of distance, two measures taken from regression model summaries of the raw data for each animal were used for this parameter in the analysis. The first was the estimated initial epithelial volume (the intercept parameter from the regression model), and the second was the estimated distance at which the epithelial volume tended to asymptote. These two summaries were thought to reflect concisely an animal's epithelial volume changes. For these calculations, summaries for each animal were averaged over the cranial and caudal regions.

CHRONIC RHINITIS

Chronic rhinitis refers to a set of conditions related to inflammation and irritation of the nasal passages. This disease surrogate comprised three primary endpoints, listed in Table 3. All measures were obtained from the study of Harkema and associates (1994).

In the nasal region, increased stored mucus would be considered protective because it would act as a barrier against contaminants reaching more distal airways. However, such increases are likely to lead to chronic irritation and inflammation and, hence, are relevant to the endpoint being considered in this section. Our analysis used stored mucus in the lateral wall because previous studies of ozone exposure had suggested that this would be the most affected location. Similarly, measures of mucous flow rates in the nasal passages (again in the lateral wall area), as well as measures of structural alterations in the epithelial lining of the nasal passages (nasal secretory cells), are relevant for analysis of this disease surrogate.

Table 3. Endpoints for Analysis of Chronic Rhinitis: Sample Sizes by Exposure Group[a]

			Ozone Exposure (ppm)			
Endpoint	Code	Total Rats	0.0	0.12	0.5	1.0
Stored mucus, nasal lateral wall	SMLW	46	15	4	12	15
Nasal mucous flow rate, lateral wall	MFLW	41	13	5	10	13
Nasal secretory cells	NSC	18	5	4	4	5

[a] All measurements were obtained from Harkema et al. 1994.

METHODS AND RESULTS

REVIEW OF METHODS FOR ANALYZING MULTIPLE OUTCOMES

Multivariate analysis refers to a broad class of statistical techniques that can be applied in settings in which more than one outcome variable has been measured for each subject. Multivariate techniques have been widely studied for more than a century, and detailed discussions can be found in any number of excellent books. Although multivariate analysis can have any of several different scientific objectives, we focus here on methods that can be used for data reduction (Johnson and Wichern 1992, p. 2). The problem can be stated as follows. Consider a study involving N individuals, and suppose that K outcomes are observed on each of the individuals. We denote these outcomes by a vector Y_i of the K observations on individual i, say $Y_i1 \ldots Y_iK$. Suppose also that we observe a covariate X_i (such as dose or concentration) for the ith individual and are interested in assessing the relation between this covariate and the outcomes Y_i. A classic approach is to apply a suitable data reduction technique to the multivariate outcome Y_i, obtaining a univariate summary measure (say V_i) that retains much of the important information from Y_i. The reduced variable V_i can then be analyzed as a function of the covariate X_i using standard univariate techniques. Principal components analysis, factor analysis, and canonical correlation analysis are some well-known data reduction techniques (Johnson and Wichern 1992, chapters 8–10; Morrison 1990, chapters 8, 9). A simpler approach is to form an average, possibly weighted, of endpoints of interest.

Unfortunately, several factors limit or complicate the use of multivariate analysis techniques in many applications. First, such techniques are designed to work with normally distributed data. Hence, these methods can and often do fail in settings with extreme or outlying observations. Second, most existing methods do not easily accommodate missing data. For both these reasons, standard multivariate analysis techniques could not be directly applied to the data from the NTP/HEI Collaborative Ozone Project.

ROBUST COMPOSITE SCORES BASED ON MEDIAN POLISH ANALYSIS

If one ignored issues related to missing data and outlying data points, then a simple approach to summarizing multiple outcomes would be to average the responses measured for each animal. A more formal justification for this approach arises from the fact that in many multivariate settings, the first principal component often corresponds closely to the average response (for example, see Morrison 1990, chapter 8). Our goal was to develop an analogous form of this simple data summary that accommodates missing observations and responds robustly to the presence of outliers.

Suppose the data are arranged in an N-by-K table, where the rows represent the N individuals in the study and the columns represent the K endpoints measured for each one. The element Y_{ij} in the matrix will correspond to the value of endpoint j for individual i. In our setting, N is 164, and K is 10 for the centriacinar fibrosis analysis, 18 for the airway disease analysis, and 3 for the chronic rhinitis analysis. To summarize the data in such a table, it is convenient to think of constructing an additive decomposition of Y_{ij} as follows:

$$Y_{ij} = \mu + \alpha_i + \beta j + \varepsilon_{ij}. \quad (1)$$

Such a decomposition models each observed value in terms of a grand mean (overall effect), a component associated with a particular animal (row effect), a component associated with a particular endpoint (column effect), and a (residual). The component of most interest in this setting is the animal (row) effect, which will be used in the subsequent analysis of exposure effects.

The classic way to perform the decomposition (Equation 1) is through a two-way analysis of variance (ANOVA). If the data are complete (that is, every individual is measured for each endpoint), then the row effect is simply the average of the endpoints for a given animal. More specifically,

$$\hat{\alpha}_i = \overline{Y}_i - \overline{Y}; \quad (2)$$

that is, the estimated animal effect is the deviation of the animal's average response from the grand mean.

The ANOVA procedure is known to be optimal when the data are normally distributed. In the absence of normality, the procedure will still be valid so long as the data are homogeneous with respect to their variance. Transformations of the data, such as taking logarithms or transforming to rankits (see Tukey 1977), can be used to reduce heteroscedasticity. An alternative is to use a more robust model. On a technical level, ANOVA minimizes sums of squared deviations from the assumed mean. Hence, extreme deviations (outliers) are amplified by the squaring process. Minimizing the sums of absolute values of deviations provides a process that is less sensitive to large deviations. One disadvantage of this method is its computational complexity.

A robust alternative is median polish analysis, proposed by Tukey (1977) in the context of decomposing a two-way data table (see also Mosteller and Tukey 1977). In terms of results, median polish analysis is similar to (and often the same as) minimizing the sum of absolute deviations (Hoaglin et al. 1983; Fink 1988). However, it is computationally much less

complex. Moreover, median polish analysis may be easier to explain to applied researchers because it is conceptually very similar to ANOVA. Unlike ANOVA, however, median polish analysis imposes no strong parametric assumptions, requires minimal variable preparation (for example, merely standardizing by an estimated SD), and accommodates both extreme values and missing data.

Once the median polish analysis has been applied to a data table, the resulting row effects can be used as composite scores for each animal. To illustrate this approach, consider the 164-by-10 table obtained by cross-classifying all 164 individual animals in the NTP/HEI Collaborative Ozone Project with respect to the 10 endpoints listed in Table 1 for centriacinar fibrosis. As an example, Table 4 lists the resulting data only for five animals, H1, H2, H3, H163, and H164. The abbreviation NA (not available) indicates that the corresponding outcome was not measured for that animal. Of the 10 endpoints, for instance, animal H164 had only total collagen (COLT) measured. The large number of missing values in Table 4 is fairly typical. Table 2 shows, for example, that even the respiratory function endpoint (mean midexpiratory flow, the most commonly assessed endpoint) was measured in only 61 of the 164 animals. The last two rows of Table 4 provide the median and the median absolute deviation (MAD, or median distance around the median) for the data in each column. The fact that some of the variables have a minus sign reflects an priori judgment that the exposure is expected to have a negative (decreasing) effect on those endpoints.

The first step in preparing the raw data for median polish analysis is the standardization step. This involves several things, including centering (subtracting out the median), scaling (dividing by the MAD), and deciding on the appropriate sign (plus or minus) to apply to the transformed data. Operationally, this is accomplished by reversing the signs on those variables that are expected to decrease with increasing ozone exposure (that is, we define composite scores as quantities that are anticipated to increase with exposure). This step is necessary to insure that some measurements do not cancel out the effects of others in the median polish process. For some endpoints, the direction that the exposure's effect will take is well established by theory (for instance, collagen content for analysis of centriacinar fibrosis), while for others (for example, the GAG measurements for analysis of centriacinar fibrosis) there is less a priori information about expected direction.

The latter step is necessary to maximize the expected association between the exposure variable (here, ozone concentration) and the composite scores that will be ultimately extracted. For instance, the discussion on the fibrosis endpoints suggests that lung function measurements such as residual volume (RV), D_{CO}, and QCC will decline with increasing ozone exposure. Hence, after these variables are standardized, the signs are reversed. Thus, the original variables Y_{ij} will all be transformed to

$$Z_{ij} = sgn_j \frac{Y_{ij} - \text{med}_j}{\text{MAD}_j}, \tag{3}$$

where sgn_j, med_j, and MAD_j are the sign, median, and MAD associated with the jth variable. Alternative standardization procedures are possible. For instance, the mean instead of the median can be subtracted, and SD instead of MAD could be used as a scaling factor. Other techniques are also possible, and in practice we have found different standardizing methods to yield closely comparable results.

Table 4. Raw Data for Ten Centriacinar Fibrosis Endpoints[a]

Animal Number	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H1	17.10	NA[b]	NA	NA	NA	NA	NA	−2.70	−1.07	−0.51
H2	18.10	NA	NA	NA	NA	0.11	0.19	−0.70	−0.81	−0.40
H3	17.10	NA	NA	NA	NA	0.10	0.24	−1.30	−1.00	−0.68
H163	18.10	NA	NA	NA	NA	NA	NA	−1.60	−1.07	−0.82
H164	19.10	NA	NA	NA	NA	NA	NA	NA	NA	NA
Median	14.10	0.19	0.15	1.70	0.01	0.24	0.30	−1.60	−0.80	−0.55
MAD	3.00	0.05	0.02	0.28	0.01	0.05	0.04	0.40	0.08	0.11

[a] See Table 1 for explanation of codes in column headings. We analyzed all 164 animals in the NTP/HEI Collaborative Ozone Project; however, data for only five animals are given here as an example of the method.

[b] NA = not available.

Tables 1 and 5 provide a clearer picture of how many fibrosis-related endpoints are measured on various subsets of the animals in the study. For instance, GAG measurements C6S and DS were made on a total of 32 animals while respiratory function measurements were made on 61 animals. The numbers in the various exposure groups are fairly equally distributed, with the exception of the 0.12 ppm ozone level, for which fewer animals were available overall (see the study design paper, Boorman et al. 1995). Table 5 provides information about the number of animals for which there was overlap between various pairs of fibrosis-related endpoints. By design, the morphological measurements had no overlap with biochemical endpoints, and only limited overlap with respiratory function endpoints. The amount of overlap between biochemical and respiratory function endpoints was larger.

Once the variables are standardized, they are ready for median polish analysis. Table 6 provides a step-by-step description of the median polish analysis. In each of the two steps of the median polish, there are four substeps (call these *a*, *b*, *c*, and *d*). The data listed at the very beginning of Table 6 (step 1*a*) are the transformed data. Substep *a* consists of taking the median of each row and placing that median in column 12 of the table. In substep *b*, column 12 is subtracted from the main elements of the table, then the values in column 12 are added to column 13. In substep *c*, medians are taken for each column (including column 13), and the resulting values are placed in row $N+1$ (in our case, row 165, "Work"). At substep *d*, the medians in row $N+1$ are subtracted from the table and added to the values in row $N+2$ (in our case, row 166 "Median"). At step 1 of the polish, the final rows and columns of the table start at zero. These substeps are repeated for step 2 of the polish to illustrate the iterative nature of this process. After the polish is complete, the values in the outermost row and column will be the estimated row (animal) and column (standardized endpoint) effects. Most importantly, the row effects will be the robust composite scores for each animal that will then be used to characterize the concentration response.

INTERPRETATION OF THE SCORES

To evaluate whether the median polish analysis could produce meaningful summaries for individual animals, we conducted simulation studies that generated normally distributed animal effects and compared these quantities with the estimated animal (row) effects from the median polish analysis. We also compared the true animal effects with those obtained from a more naive analysis that simply transformed the outcomes and constructed median summaries for each animal rather than subjecting the entire data array to median polish analysis. Figures 1 through 4 show the results of these simulations for two simulated data sets. Figure 1A shows boxplots* of 165 random values for each of 10 endpoints with varying means and variances. These data were generated by adding normally distributed noise to the endpoint mean and the animal effect for each of the 165 animals. Note that we used 165 animals instead of 164 for computational convenience. The amount of background variability was then varied across endpoints (chosen to mimic actual study data). Each endpoint was standardized by subtracting the mean and dividing by the SD. Figure 1B displays the true animal effects against the values obtained by median polish analysis, and Figure 1C displays the true animal effects against the simple medians. Straight lines in either plot would indicate perfect estimation of the true

* Each boxplot indicates the median (white bar) and the 25th and 75th percentiles (limits of the box) and values occurring up to 1.5 times the length of the box (indicated by the tails extending from the box).

Table 5. Pairwise Common Samples for Centriacinar Fibrosis Endpoints[a]

Endpoint	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
COLT	0	0	0	0	12	12	29	29	29
COLM	–	39	33	39	0	0	7	7	7
FV		–	33	39	0	0	7	7	7
IV			–	33	0	0	4	4	4
IMV				–	0	0	7	7	7
C6S					–	32	19	19	19
DS						–	19	19	19
RV							–	61	61
QCC								–	61

[a] See Table 1 for explanation of codes in column and row headings.

Table 6. Step-by-Step Overview of the Median Polish Analysis[a]

	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}	Work[b]	Median[c]
Step 1a												
H1	1.00	NA	NA	NA	NA	NA	NA	−2.75	−3.37	0.36	−1.20	0.00
H2	1.33	NA	NA	NA	NA	−2.45	−2.33	2.25	−0.12	1.26	0.57	0.00
H3	1.00	NA	NA	NA	NA	−2.64	−1.22	0.75	−2.50	−1.16	−1.19	0.00
H163	1.33	NA	NA	NA	NA	NA	NA	0.00	−3.37	−2.40	−1.20	0.00
H164	1.67	NA	NA	NA	NA	NA	NA	NA	NA	NA	1.67	0.00
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Step 1b												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.02	−2.90	1.68	−0.69	0.69	0.00	0.57
H3	2.19	NA	NA	NA	NA	−1.45	−0.03	1.94	−1.31	0.03	0.00	−1.19
H163	2.54	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.20
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Step 1c												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.02	−2.90	1.68	−0.69	0.69	0.00	0.57
H3	2.19	NA	NA	NA	NA	−1.45	−0.03	1.94	−1.31	0.03	0.00	−1.19
H163	2.54	NA	NA	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	−0.15	0.00	−0.20	0.01	0.06	0.05	0.00	0.00	0.00	0.00	0.00
Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Step 1d												
H1	2.20		NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76		NA	NA	NA	−3.09	−2.95	1.68	−0.69	0.69	0.00	0.57
H3	2.19		NA	NA	NA	−1.51	−0.08	1.94	−1.31	0.03	0.00	−1.19
H163	2.54		NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.20
H164	0.00		NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	−0.15	0.00	−0.20	0.01	0.06	0.05	0.00	0.00	0.00	0.00	0.00
Step 2a												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.09	−2.95	1.68	−0.69	0.69	0.00	0.57
H3	2.19	NA	NA	NA	NA	−1.51	−0.08	1.94	−1.31	0.03	−0.03	−1.19
H163	2.54	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.20
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	−0.15	0.00	−0.20	0.01	0.06	0.05	0.00	0.00	0.00	0.00	0.00

(Table continues next page.)

Table 6. Step-by-Step Overview of the Median Polish Analysis[a] *(continued)*

	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}	Workb	Median[c]
Step 2b												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.09	−2.95	1.68	−0.69	0.69	0.00	0.57
H3	2.22	NA	NA	NA	NA	−1.48	−0.06	1.97	−1.28	0.06	0.00	−1.22
H163	2.54	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.20
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	−0.15	0.00	−0.20	0.01	0.06	0.05	0.00	0.00	0.00	0.00	0.00
Step 2c												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.09	−2.95	1.68	−0.69	0.69	0.00	0.57
H3	2.22	NA	NA	NA	NA	−1.48	−0.06	1.97	−1.28	0.06	0.00	−1.22
H163	2.54	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.20
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.67
Work	0.00	−0.07	0.00	−0.05	−0.01	0.03	0.03	0.00	0.00	0.00	0.00	0.01
Median	0.00	−0.15	0.00	−0.20	0.01	0.06	0.05	0.00	0.00	0.00	0.00	0.00
Step 2d												
H1	2.20	NA	NA	NA	NA	NA	NA	−1.55	−2.18	1.55	0.00	−1.20
H2	0.76	NA	NA	NA	NA	−3.12	−2.98	1.68	−0.69	0.69	0.00	0.56
H3	2.22	NA	NA	NA	NA	−1.51	−0.08	1.97	−1.28	0.06	0.00	−1.22
H163	2.54	NA	NA	NA	NA	NA	NA	1.20	−2.17	−1.20	0.00	−1.21
H164	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.00	1.66
Work	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Median	0.00	−0.22	0.00	−0.25	0.01	0.09	0.08	0.00	0.00	0.00	0.00	0.01

[a] See Table 1 for an explanation of the codes in column headings. We analyzed all 164 animals in the NTP/HEI Collaborative Ozone Project; however, data for only five animals are given here as an example of the method.

[b] Intermediate calculations of row and column medians.

[c] Final row and column medians at each substep.

animal effects. For comparison, Figure 2 is similar to Figure 1 except that the median/MAD standardization was used for each variable rather than mean/SD standardization.

Figures 1 and 2 illustrate that median polish analysis adequately retrieves the animal effects from the data and that the effect is similar using either standardization method (mean/SD or median/MAD). Also, it is clear that, in these simulations, the results of median polish analysis closely approximate the results obtained by using the simple median after standardizing each variable.

To investigate the ability of median polish analysis to retrieve animal effects in the presence of outliers, we modified the simulation to generate background noise from a Cauchy distribution (a distribution with "heavy tails" and prone to extreme values), rather than a normal distribution, and left the animal effects normally distributed. As shown in Figures 3 and 4, this resulted in similar animal effect values, but the individual endpoint distributions often have large outliers. As in the previous simulation, Figure 3 shows the results of using the mean/SD standardization, and Figure 4 incorporates the median/MAD standardization. Notice that again the median polish analysis tends to estimate the animal effects quite well, and in the case of the mean/SD standardization, the results appear less variable than those obtained using the simple animal medians (compare Figure 3B with Figure 3C). These results suggest that the median polish technique is able to estimate a generalized animal effect across a variety of endpoints, some of which may be nonnormal and prone to extreme values. In these initial simulations, the animal effects were randomly generated.

As we illustrate in the next section, when the animal effects are related to covariates (like exposure group or gender), the median polish composites can then be analyzed to investigate their relation to covariates.

RESULTS FROM THE DISEASE SURROGATE ANALYSES

Appendix Tables A.1, A.2, and A.3 give, for centriacinar fibrosis, airway disease, and chronic rhinitis, respectively, the individual animals and endpoints entering each analysis by indicating (with asterisks) which of the 164 total animals were measured. Thus, these tables provide a complete picture of the animal overlap for each set of endpoints. Table 7 summarizes the information from Tables A.1 through A.3 by tabulating the number of endpoints per animal entering each disease surrogate analysis. Table 8 summarizes the data from a different viewpoint by providing the number of studies for which investigators measured each animal out of the total of eight functional, structural, and biochemical individual study components of the NTP/HEI Collaborative Ozone Project.

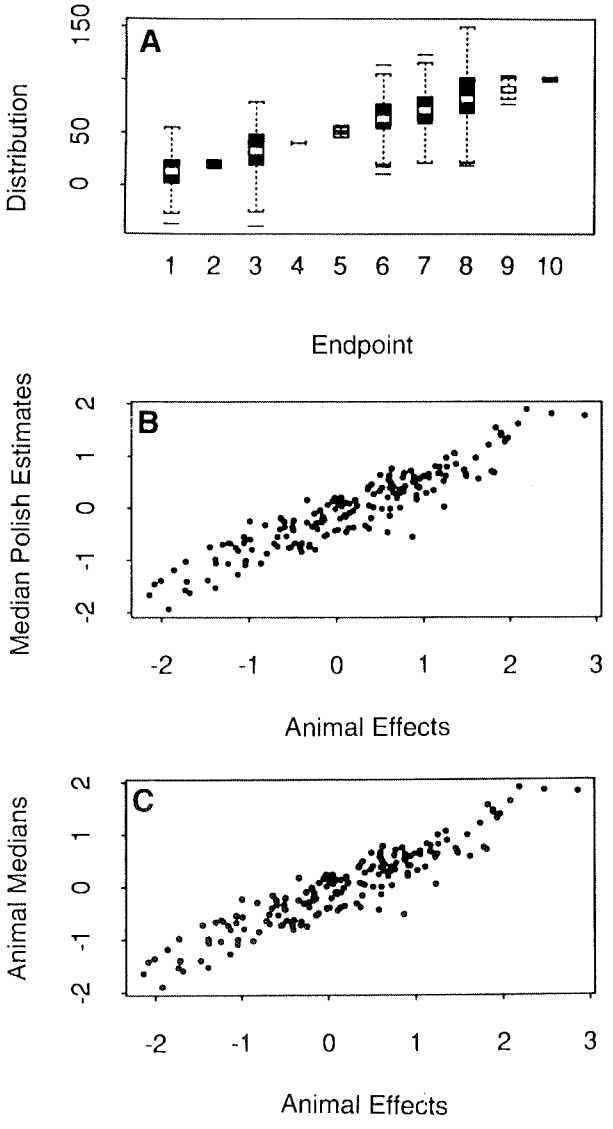

Figure 1. Simulation to evaluate median polish under normal error, using mean/SD standardization: (A) endpoint distributions; (B) median polish estimates versus true animal effects; (C) animal medians versus true animal effects.

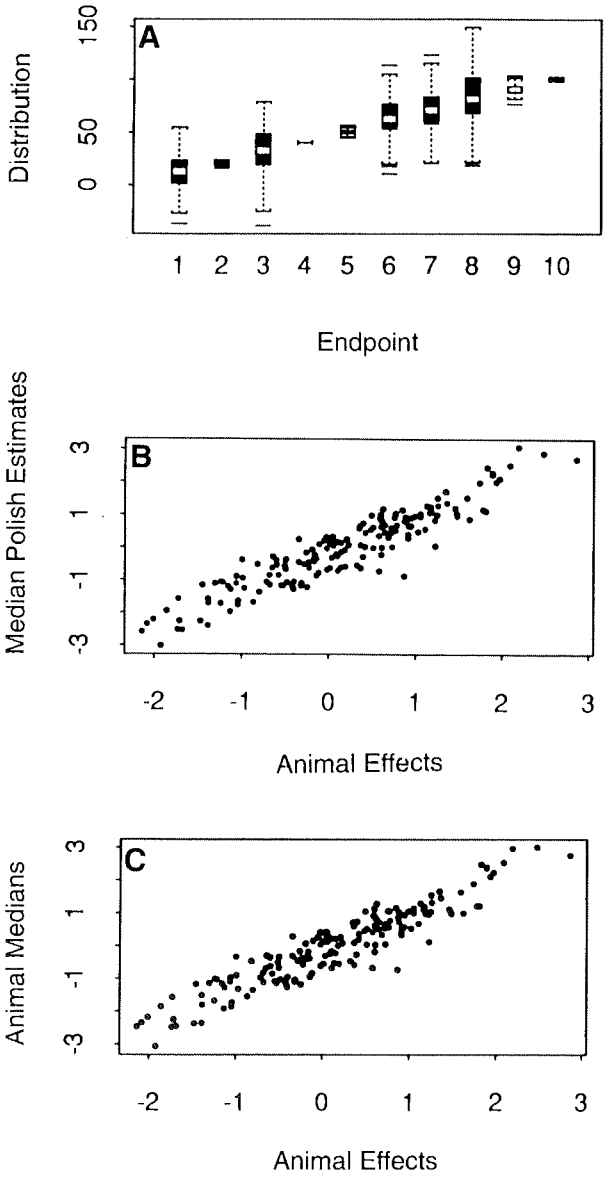

Figure 2. Simulation to evaluate median polish under normal error, using median/MAD standardization: (A) endpoint distributions; (B) median polish estimates versus true animal effects; (C) animal medians versus true animal effects.

Centriacinar Fibrosis

Figures 5 through 8 show the analyses for the centriacinar fibrosis composite score. Figures 5 through 7 plot the results of the median polish analysis for the fibrosis endpoints (centriacinar fibrosis score) against the concentration of ozone exposure for all animals. The p values under each ozone concentration show the results for testing each concentration versus control using a standard t test. None of the three pairwise comparisons is significant. Using linear regression analysis, the concentration response for the fibrosis scores shows a marginally significant trend associated with increasing ozone concentration. In subset analysis by gender, the data for female rats show a clearer trend; in fact, the pairwise comparison for females at 1.0 ppm ozone shows a rather large increase, although not statistically significant ($p = 0.11$), whereas the data for male rats show no evidence of a trend or pairwise differences. This gender pattern was observed in individual study data as well, most notably by Last and associates (1994). Their report indicates that collagen increased only in female animals and only when it was expressed in terms of content (per lung lobe) instead of concentration (per lung lobe weight).

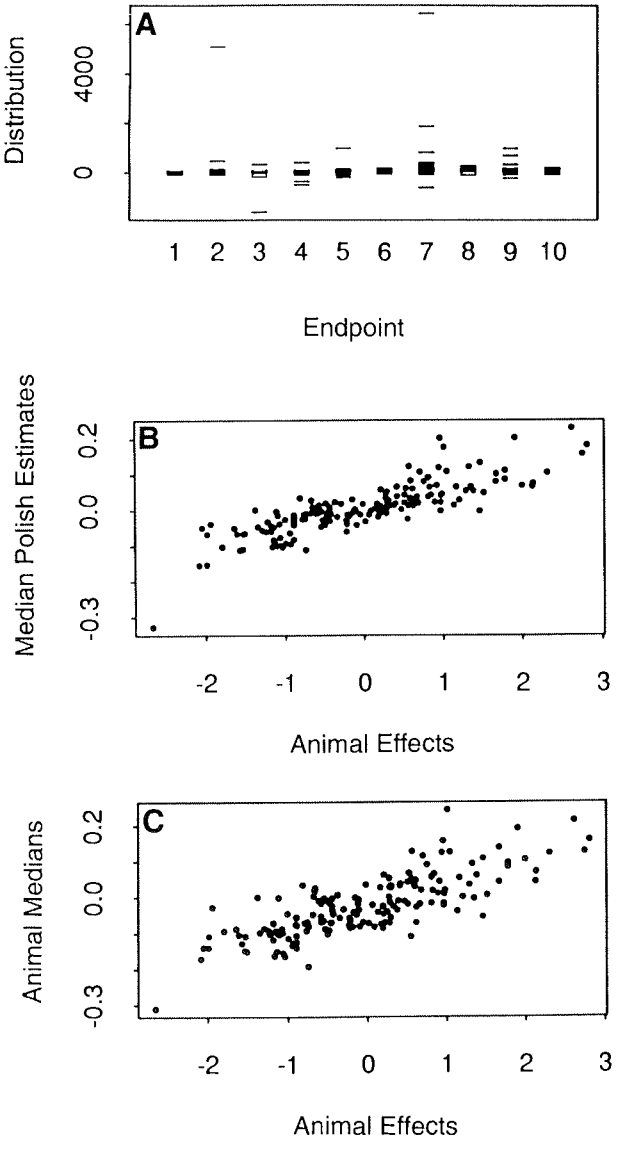

Figure 3. Simulation to evaluate median polish under Cauchy error, using median/SD standardization: (A) endpoint distributions; (B) median polish estimates versus true animal effects; (C) animal medians versus true animal effects.

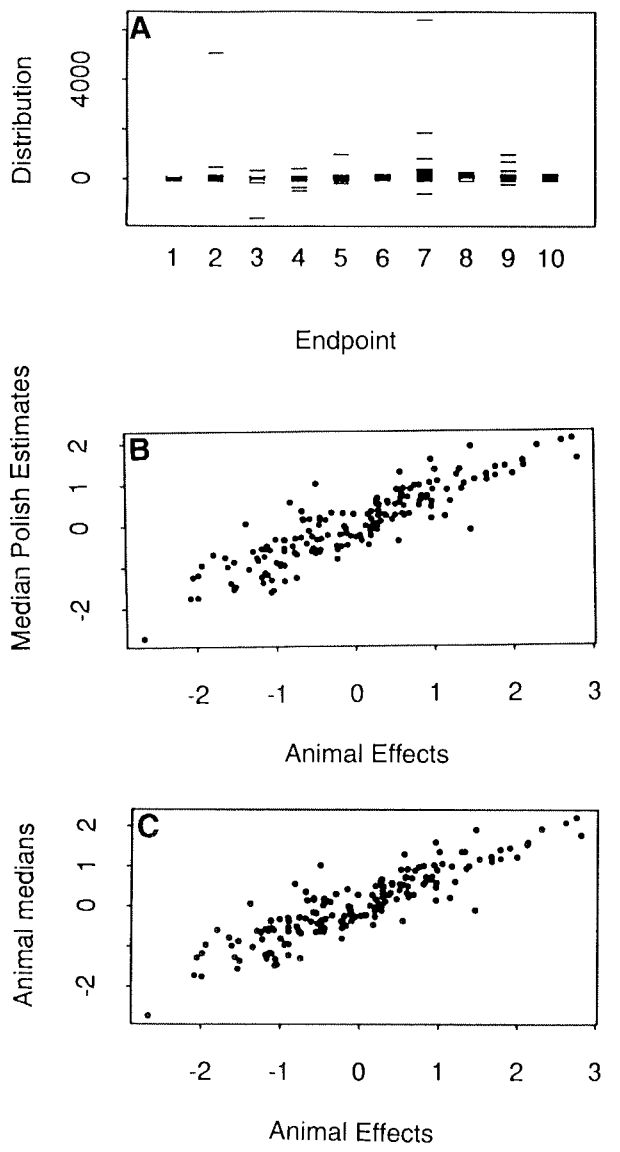

Figure 4. Simulation to evaluate median polish under Cauchy error, using median/MAD standardization: (A) endpoint distributions; (B) median polish estimates versus true animal effects; (C) animal medians versus true animal effects.

To investigate further the effects of different ways of measuring the collagen data on the results of the composite score analysis, we also computed composite scores using the concentration approach to expressing total lung collagen. Table 9 and Figures 9 through 11 show the results of these analyses. It is clear that for the analysis including all animals, the concentration-response trend is attenuated using the collagen concentration approach. Upon further analysis of each gender, it is also clear that changes observed in the females under the collagen content approach are no longer apparent under the collagen concentration approach and that the interpretation of the male data remains the same under either method of standardization. These findings are consistent with those of Last and associates (1994) and illustrate, at least in females, the sensitivity of the results to the collagen metric.

Table 7. Number of Endpoints Measured per Animal in Each Disease Surrogate

	Number of Animals		
Number of Endpoints	Centriacinar Fibrosis	Airway Disease	Chronic Rhinitis
1	13	33	25
2	13	0	40
3	18	6	0
4	49	0	0
5	7	2	0
6	15	0	0
7	4	16	0
8	0	20	0
9	0	1	0
10	0	29	0
Total animals studied	119	107	65
Animals not included[a]	45	57	99
Total animals in Project	164	164	164

[a] This value includes the 20 animals that died and therefore contributed no data for analysis.

Table 8. Number of Individual Studies per Animal for Each Disease Surrogate

	Number of Animals		
Number of Individual Studies	Centriacinar Fibrosis	Airway Disease	Chronic Rhinitis
1	44	47	65
2	59	58	0
3	16	2	0
Total	119	107	65

A priori selection of variables for this analysis and relevant biologic theory assumed that increases in GAGs would be consistent with centriacinar fibrosis. This turned out not to be the case. In fact, exposure resulted in significant decreases in GAGs (see Radhakrishnamurthy 1994). Thus, inclusion of these data in the fibrosis analysis diluted the relation between concentration and response and is partly responsible for the weak concentration-response trends in Figures 5 through 7. Although adjustment for the unexpected direction of the GAG results could have been made in the analysis, it seemed wholly inappropriate to alter theory based on the observed data from a single investigation. Thus, it was decided to allow the GAG data to enter the analysis as originally anticipated, and to report on the ramifications of this choice. Table 10 and Figures 12 through 17 compare the final centriacinar fibrosis analysis with analyses that parameterize the GAG measurements in other ways. The second line of Table 10 and Figures 12 through 14 give the analysis results after reversing the sign (direction) of the expected change in the GAG measurements (to reflect the actual GAG decreases). The third line of Table 10 and corresponding Figures 15 through 17 show results from the centriacinar fibrosis analysis after removing the GAG measurements entirely from the analysis.

It is clear that both of these alternative analyses affect the fibrosis results for the comparisons of exposures to 0.5 and 1.0 ppm ozone (and also the overall concentration-response trend). Although it is tempting to report either of these analyses as the final model (the model excluding the GAG measures may be the more attractive one), the collaborative group agreed that the original intent in creating the composite disease surrogate scores was to bring together parameters from the individual studies and evaluate the data with regard to the most current theory. In this disease surrogate, theoretical considerations indicated that increases in GAGs would be consistent with development of fibrosis. Thus, our strategy was to report the results as originally outlined and discuss any discrepancies between the data and theory, rather than abandon theoretical considerations and report the most statistically significant model.

Because the presence of leukemia may compromise the investigator's ability to detect fibrosis, the analysis was also conducted excluding leukemic animals. In fact, there is some confounding between gender, ozone exposure, and leukemia incidence. Specifically, as shown in Table 11, there was substantially more leukemia among the male

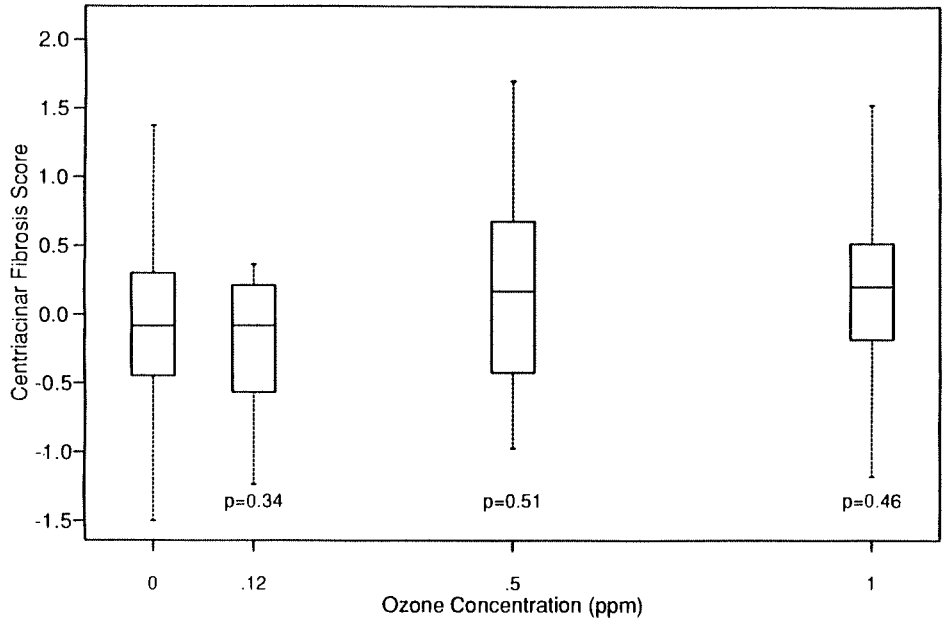

Figure 5. Analysis of centriacinar fibrosis in all animals: Concentration-response trend $p = 0.091$.

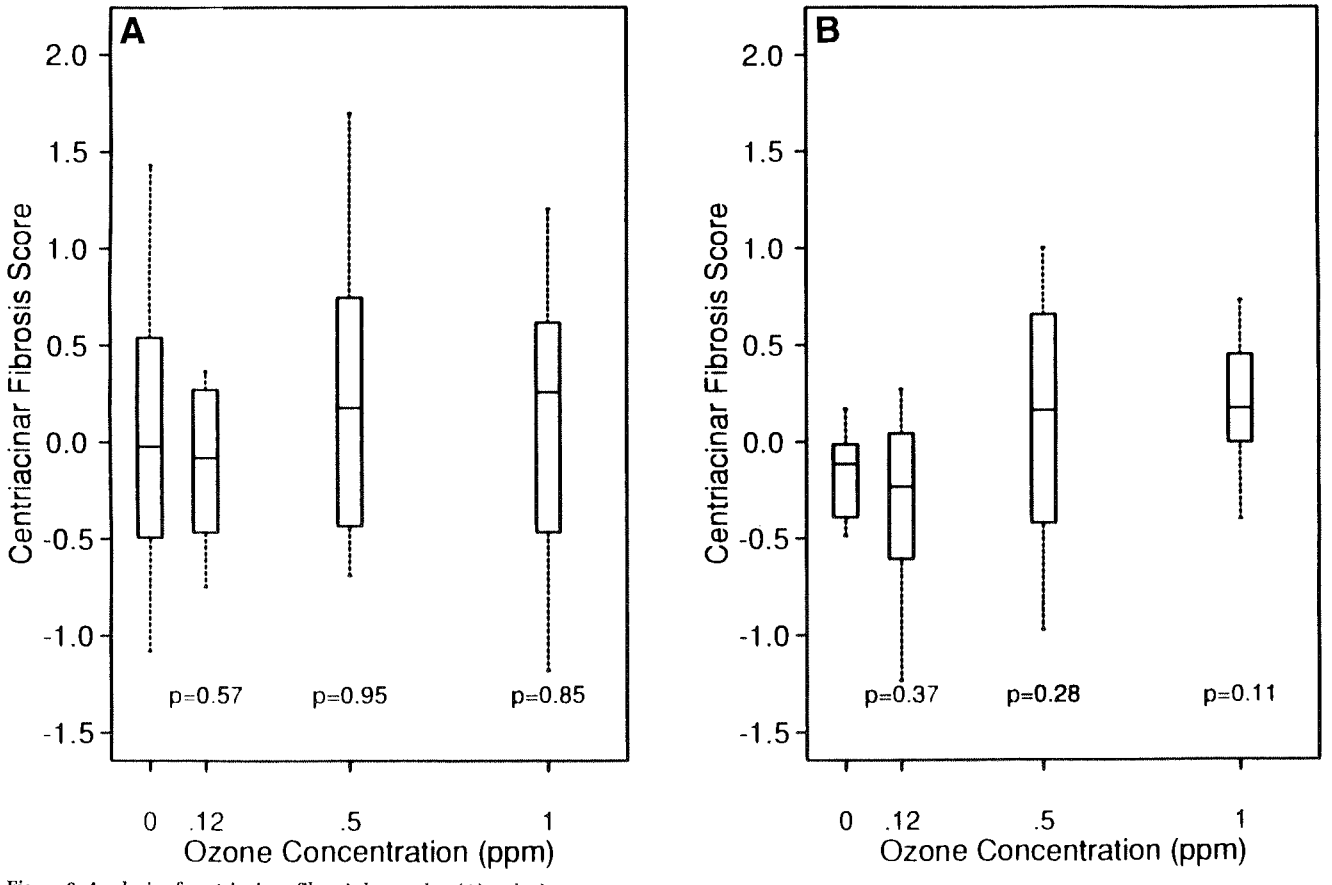

Figure 6. Analysis of centriacinar fibrosis by gender: (A) male data, concentration-response trend $p = 0.97$; (B) female data, concentration-response trend $p = 0.038$.

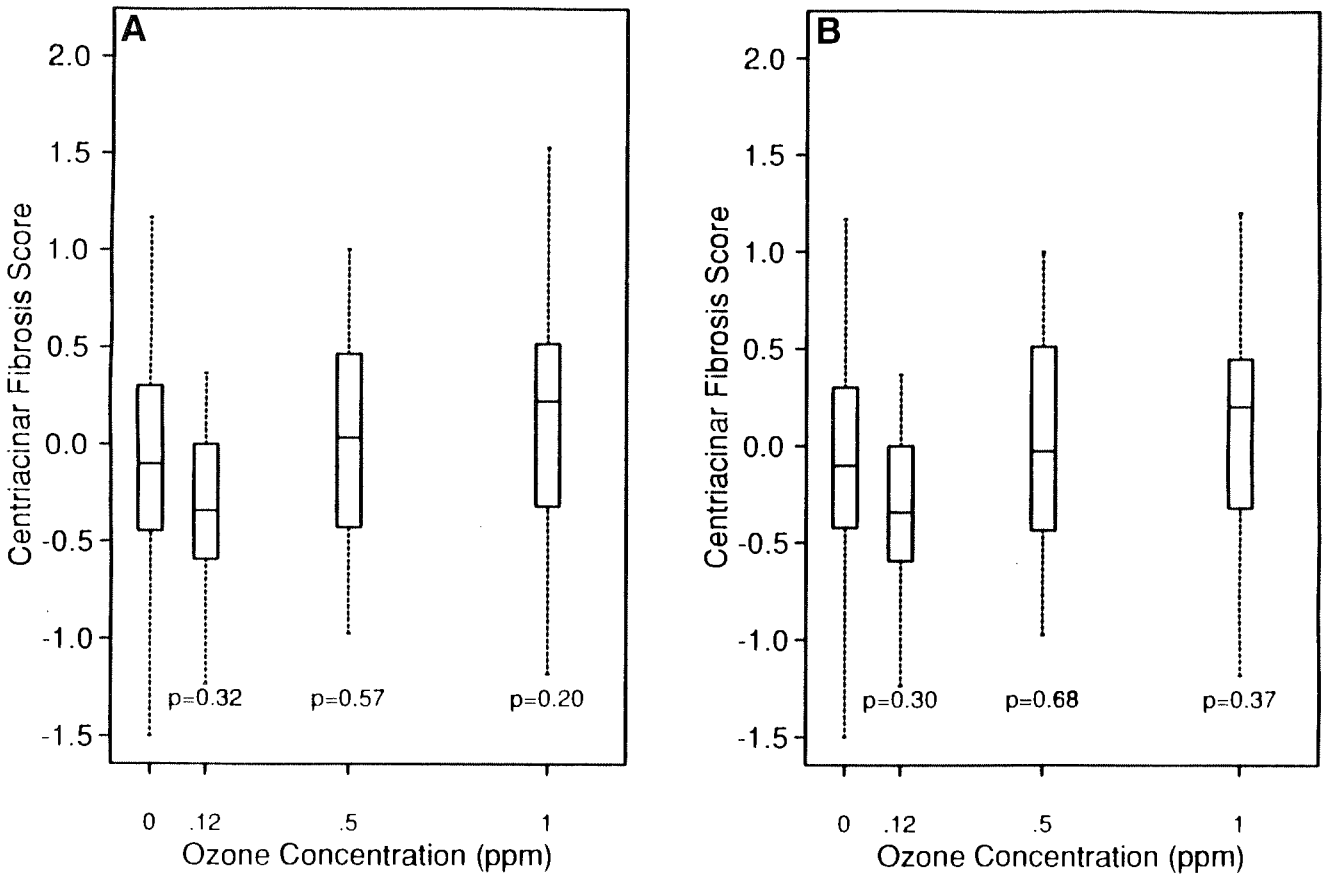

Figure 7. Analysis of centriacinar fibrosis excluding animals with leukemia: (A) excluding animals with advanced leukemia, concentration-response trend $p = 0.056$; (B) excluding animals with any leukemia, concentration-response trend $p = 0.12$.

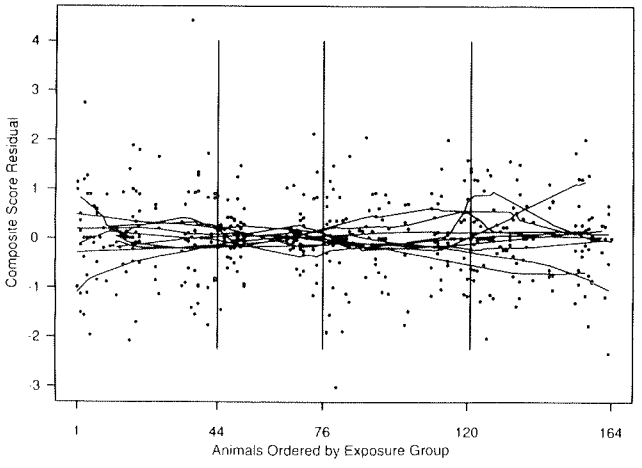

Figure 8. Residuals from the centriacinar fibrosis model. Vertical lines divide the animals by exposure group: from left, 0.0 ppm (control), 0.12 ppm, 0.5 ppm, 1.0 ppm ozone. Numbers on the horizontal axis indicate available animals in each group.

animals, and this trend was especially apparent in controls. When the animals with advanced leukemia are excluded or when animals with any degree of leukemia are excluded, the fibrosis results are not substantially changed (Figure 7).

Airway Disease

Figures 18 through 21 show the analyses for the airway disease composite score. A moderate and statistically significant overall upward trend is apparent for the airway disease analysis, which upon subset analysis is persistent in the males but not in the females. Pairwise comparisons of values for each ozone concentration with control values are not significant, although there is some evidence of a reduction in response at 0.12 ppm ozone, both overall and separately in males and females. In contrast to the results for centriacinar fibrosis, removal of the leukemic animals (Figure 20) increased the strength of the airway disease trend although pairwise comparisons remained nonsignificant.

Table 9. Sensitivity of the Centriacinar Fibrosis Analysis to Collagen Standardization Method

	p Values (Two-Sided)			
Model	Trend	0.12 ppm O_3	0.5 ppm O_3	1.0 ppm O_3
All Animals				
Content method[a]	0.091	0.34	0.51	0.46
Concentration method[b]	0.27	0.87	0.80	0.34
Males Only				
Content method	0.97	0.57	0.95	0.85
Concentration method	0.64	0.96	0.71	0.64
Females Only				
Content method	0.038	0.37	0.28	0.11
Concentration method	0.33	0.69	0.39	0.30

[a] Content method was calculated per lobe.

[b] Concentration method was standardized by lobe weight.

Chronic Rhinitis

Figures 22 through 25 show the analyses for the composite score related to chronic rhinitis. Trends in the association between ozone concentration and response are quite strong for endpoints related to chronic rhinitis. The patterns are solidly maintained in both the male and female subset analyses, and in the animals that are not leukemic. There is clear evidence that the groups exposed to 0.5 and 1.0 ppm ozone are significantly different from the control group. However, there is little apparent difference between the control group and the group exposed to 0.12 ppm ozone, and certainly these comparisons were not statistically significant.

Sensitivity Analysis

The median polish technique also allows for several types of sensitivity analyses. In addition to the more standard analysis of residuals from modeling the composite scores, the table of residuals resulting from the median polish analysis (what remains after removing the animal effects) can be used to evaluate individual animal and endpoint variability. For example, consider Figures 8, 21, and 25, in which the residuals from median polish analysis are plotted for each animal by ozone exposure group. In these figures, vertical lines divide the control group and the groups exposed to 0.12, 0.5, and 1.0 ppm ozone; the numbers on the horizontal axis indicate available animals in each group. The plots contain several points at each distinct animal position along the horizontal axis because each animal contributes as many residuals as there were original observations on that animal. Furthermore, no residuals appear in positions corresponding to animals that did not enter the analysis.

For each of the disease surrogates the residuals all cluster about zero, even across ozone concentrations. This indicates that any information about response due to ozone has been removed from the data (in fact, this information is in the row effect composites that are subsequently analyzed). Exploration of the residuals does, however, allow assessment of individual animals that may be particularly deviant in one or more of their measurements and can provide insight into the possible effect that responses for that animal may have on the overall analysis. For example, in Figure 8 at least one observation is clearly extreme. Further inspection of the data reveals that this animal (H129) is a female control that suffered from advanced leukemia. To investigate residual trends by endpoint, nonparametric smooth curves (Cleveland 1979) are overlaid on the figures for each endpoint entering the analysis. There are no systematic trends in these curves, indicating that information relating to systematic changes in response to ozone exposure have been extracted from the raw data in the process of median polish analysis and moved into the animal effect composite.

Median polish analysis also allows investigation of the sensitivity of the particular choice of responses to the overall conclusions. Consider again the airway disease analysis, which comprises 18 individual endpoints. It is necessary to determine the extent to which the overall conclusions depend on any one endpoint. Obviously, a composite analysis whose results relied too heavily on any one endpoint would not be very convincing. Similarly, results should not change drastically by including a single wrong endpoint. (By "wrong endpoint" we mean one that has no real relevance to the disease surrogate being analyzed.) These phenomena can be investigated by conducting a sensitivity analysis of the technique. Table 12 illustrates such an investigation for airway disease. The first row of Table 12 provides a summary of

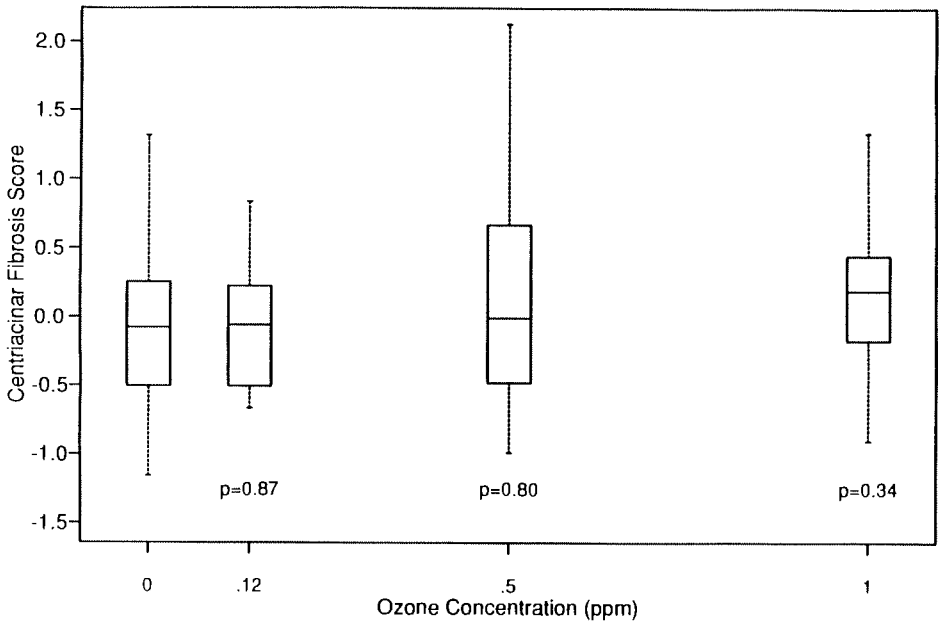

Figure 9. Centriacinar fibrosis analysis of all animals, using collagen concentration: concentration-response trend $p = 0.27$.

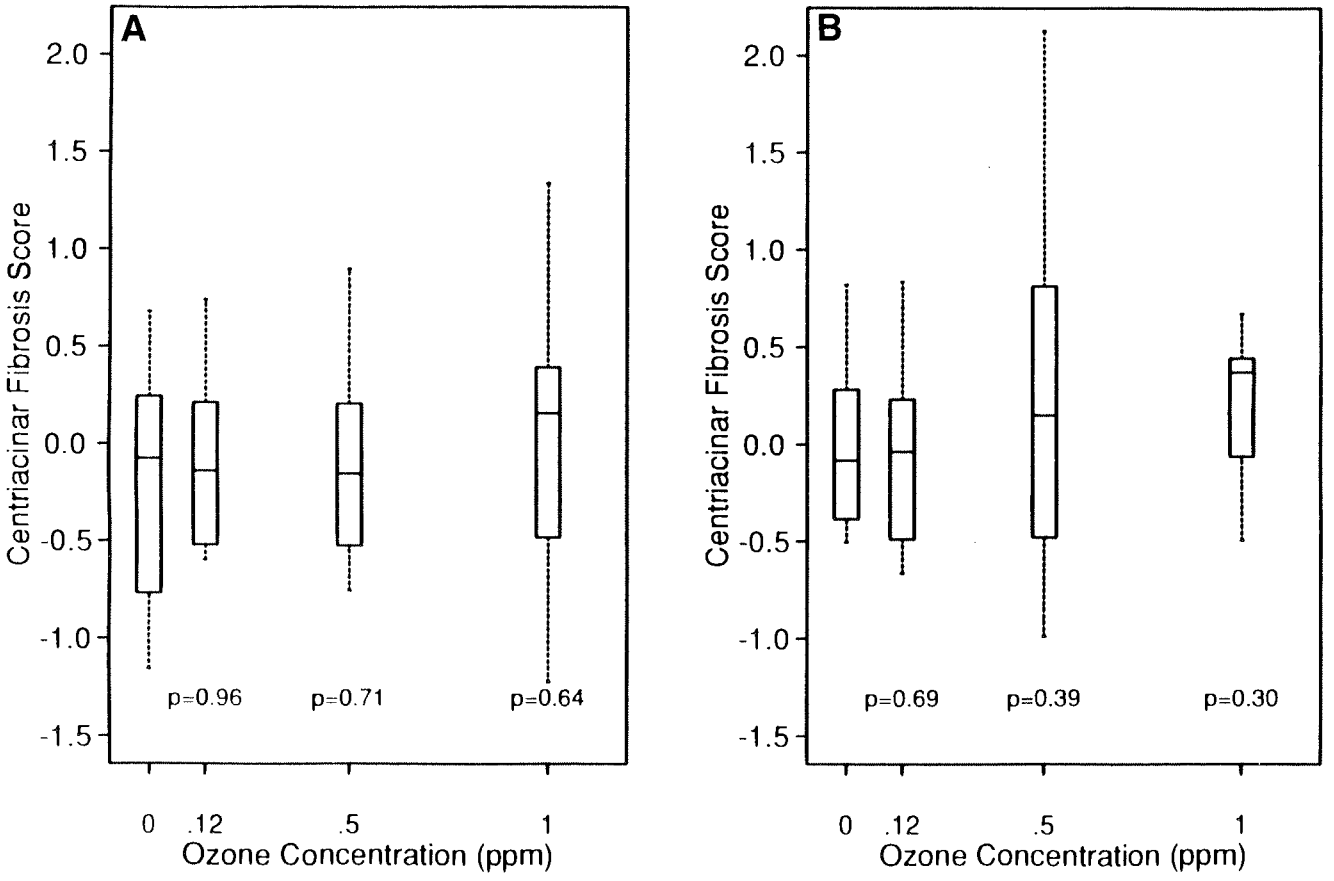

Figure 10. Centriacinar fibrosis analysis by gender, using collagen concentration: (A) male data, concentration-response trend $p = 0.64$; (B) female data, concentration-response trend $p = 0.33$.

the p values for the trend test and the three pairwise tests comparing values for each ozone concentration with control values. Subsequent rows indicate corresponding p values for analyses in which the indicated endpoint is removed from the median polish (that is, analyses in which only subsets of 17 endpoints are analyzed, rather than the full set of 18 endpoints). As is evident from Table 12, the procedure is quite robust. A possible exception lies in excluding MMEF, in which case the pairwise comparison for 1.0 ppm ozone and the concentration-response trend show deviation from the model with all 18 endpoints included. Other results from the analyses using 17-endpoint sets are very close to the results of the analysis of the complete 18-endpoint set.

EVALUATION OF THE METHOD VIA SIMULATIONS

In constructing our robust composite scores, several important issues and questions arise: Does the method estimate effects and produce summaries that are biologically meaningful? How sensitive is the method to the inclusion of endpoints that have no association with ozone exposure? What happens when endpoints have variances of different

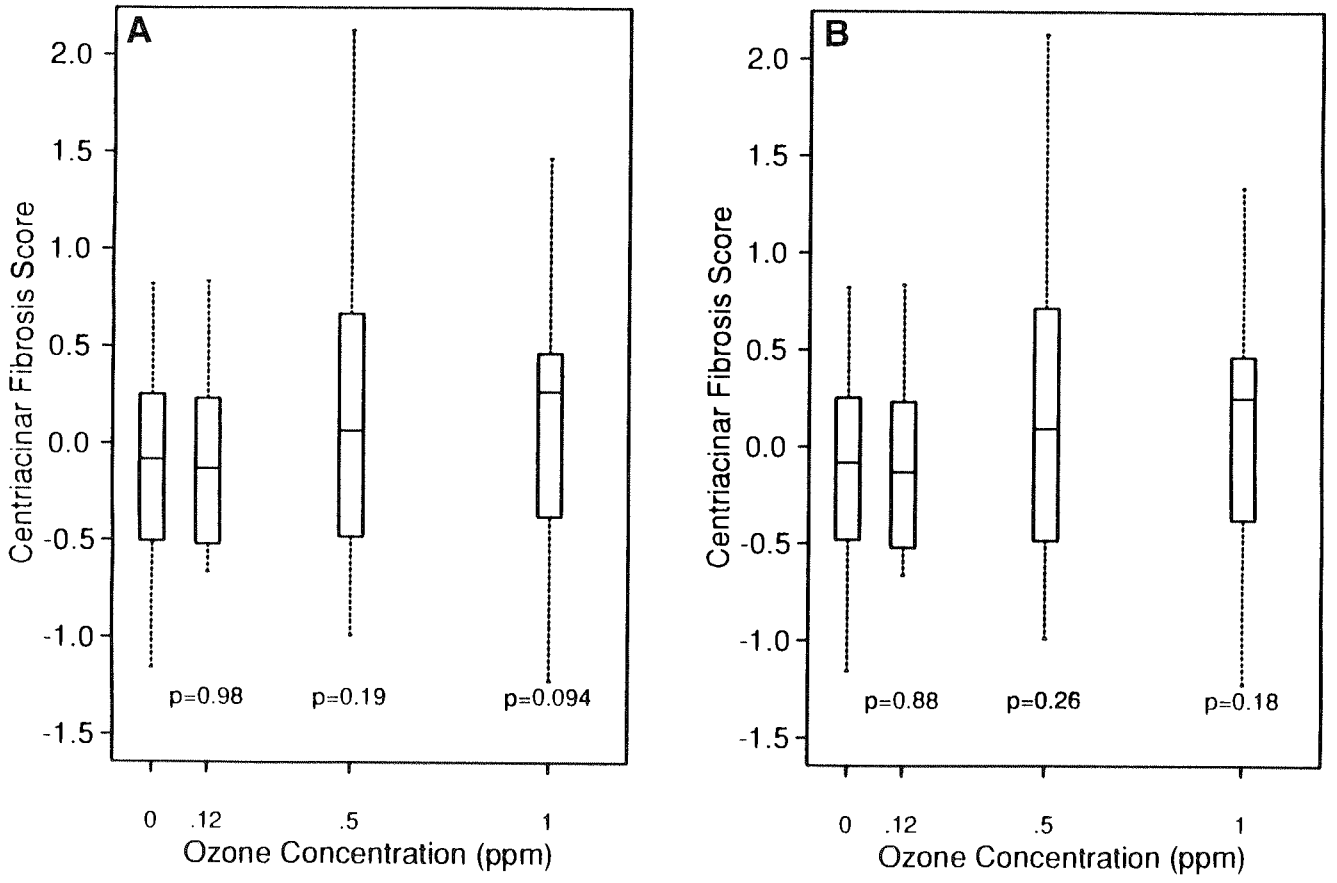

Figure 11. Centriacinar fibrosis analysis excluding animals with leukemia, using collagen concentration: (A) excluding animals with advanced leukemia, concentration-response trend $p = 0.054$; (B) excluding animals with any leukemia, concentration-response trend $p = 0.10$.

Table 10. Sensitivity of the Centriacinar Fibrosis Analysis to Glycosaminoglycan Measurements

	p Values (Two-Sided)			
Centriacinar Fibrosis Model	Trend	0.12 ppm O_3	0.5 ppm O_3	1.0 ppm O_3
Original model	0.091	0.34	0.51	0.46
Model with signs of C6S and DS reversed	0.0011	0.52	0.0050	0.0050
Model with glycosaminoglycans excluded	0.0060	0.64	0.018	0.031

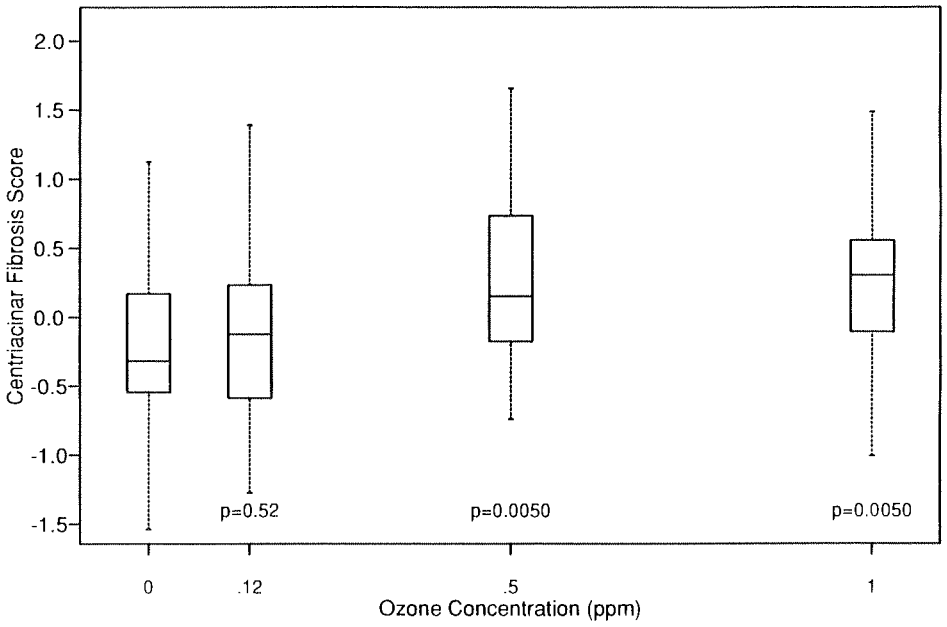

Figure 12. Centriacinar fibrosis analysis of all animals, with direction of GAG effect reversed: concentration-response trend $p = 0.0011$.

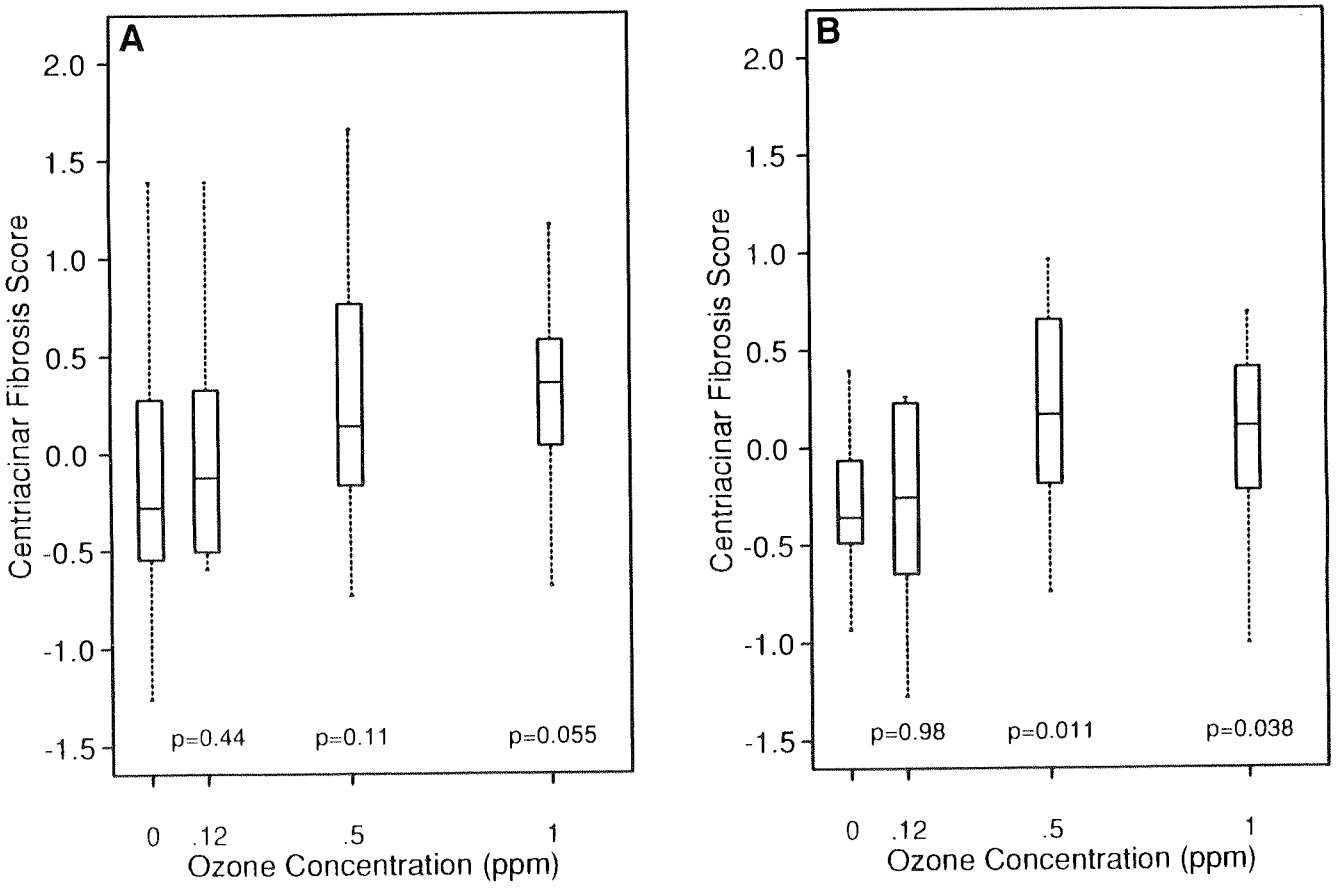

Figure 13. Centriacinar fibrosis analysis by gender, with direction of GAG effect reversed: (A) male data, concentration-response trend $p = 0.034$; (B) female data, concentration response trend $p = 0.013$.

magnitudes or concentration-related patterns? What is the impact of a large number of missing values in the table?

The first question was addressed above in our initial simulation work. To address the other issues we performed several more simulation studies in which data were generated from an additive model of the form given in Equation 1. Each simulated data set had the following characteristics:

- Five doses, 165 animals, approximately 33 animals per dose.
- Five endpoints with several characteristics: two or three of the five endpoints are affected by exposure, with equal within-dose variances; all five endpoints are affected, with equal within-dose variances; two or three of the five endpoints are affected, with unequal within-dose variances; all five endpoints are affected, with unequal within-dose variances.
- Data are complete for each endpoint for each animal, or the pattern of missing data is similar to the missing data structure of the NTP/HEI Collaborative Ozone Project study design.
- Errors are normally distributed, with moderate (0.3) intraanimal correlation.

Figure 26 shows histograms of trend test statistics under 12 different scenarios describing the simulated concentration-response patterns for the endpoints. Under each scenario 100 simulations were conducted. The scenarios were defined as having two, three, or five of the five total endpoints affected, each at four levels for the within-dose variances: variances are all equal and small (row 1), variances are all equal and large (row 2), variances are unequal with smaller variances in the affected endpoints (row 3), and variances are unequal with larger variances in the affected endpoints (row 4). For example, the (3,2) histogram displays a situation in which the three affected endpoints have large variances and the remaining two unaffected endpoints have small variances. The (4,3) plot (in which all endpoints are affected by exposure) has three endpoints with large variances and two endpoints with small variances. In each of these plots, the bars indicate the distribution of the 100 trend test statistics calculated from each set of simulated data.

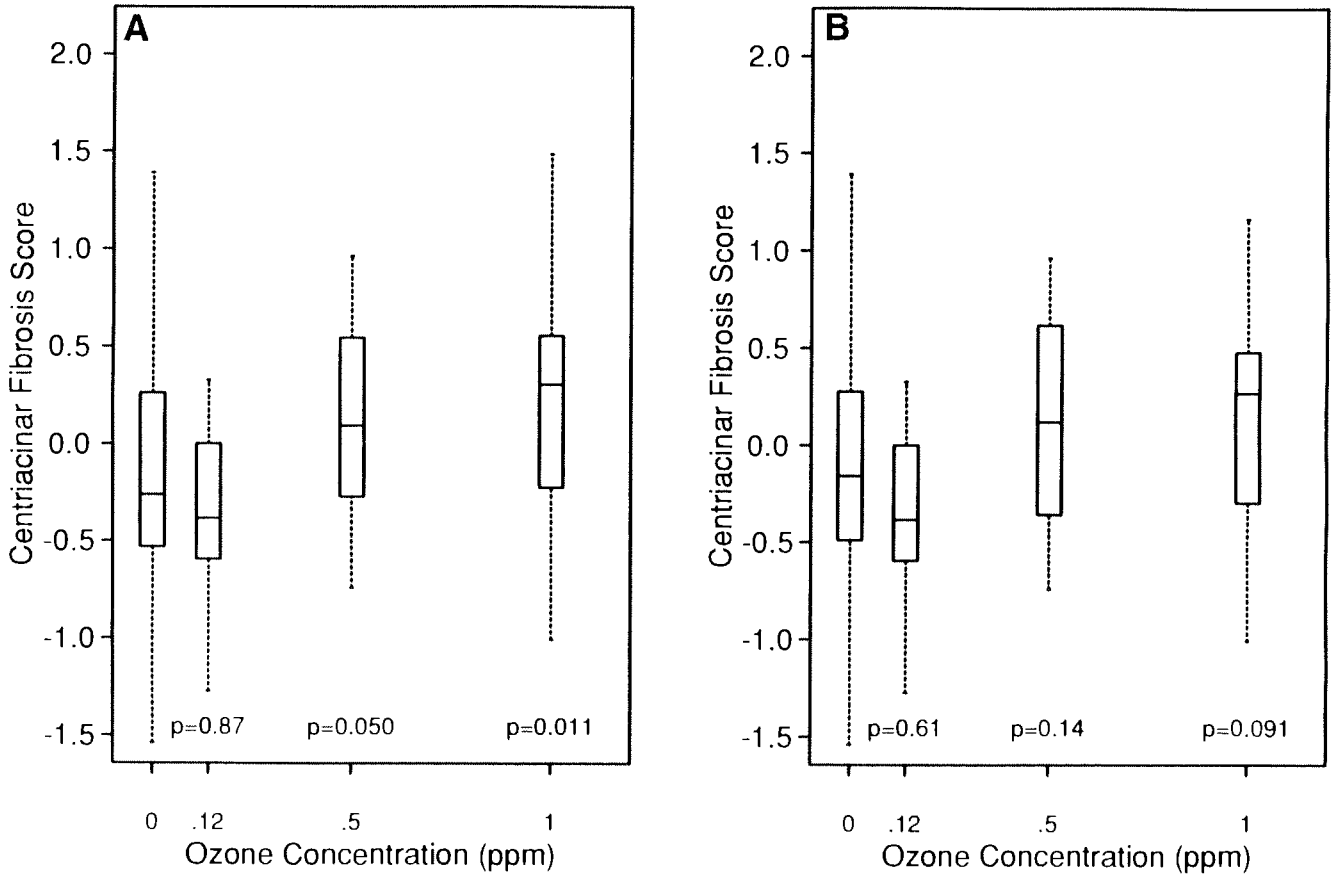

Figure 14. Centriacinar fibrosis analysis excluding animals with leukemia, with direction of GAG effect reversed: (A) excluding animals with advanced leukemia, concentration-response trend $p = 0.0014$; (B) excluding animals with any leukemia, concentration-response trend $p = 0.022$.

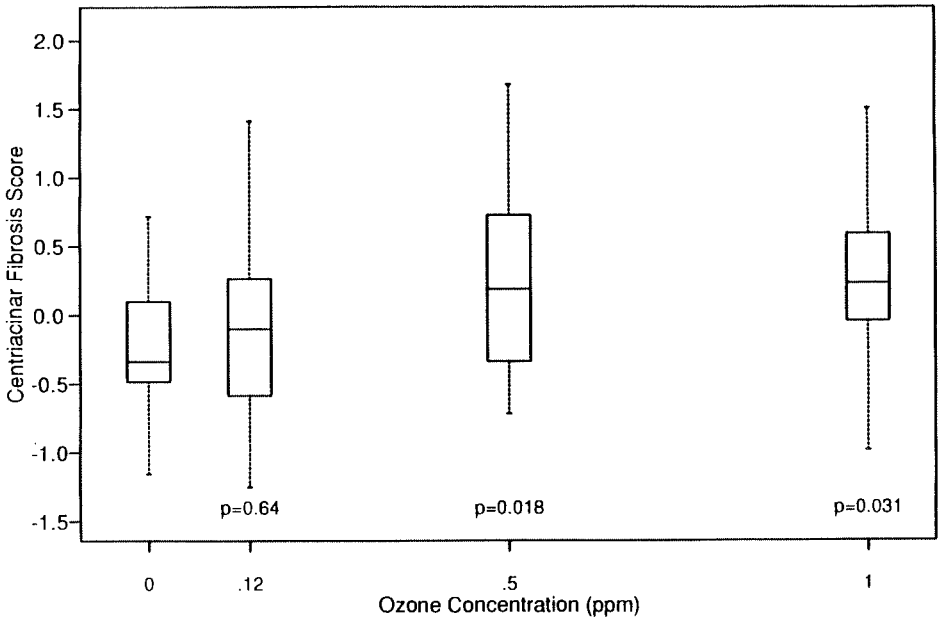

Figure 15. Centriacinar fibrosis analysis of all animals, with GAG measures excluded: concentration-response trend $p = 0.0060$.

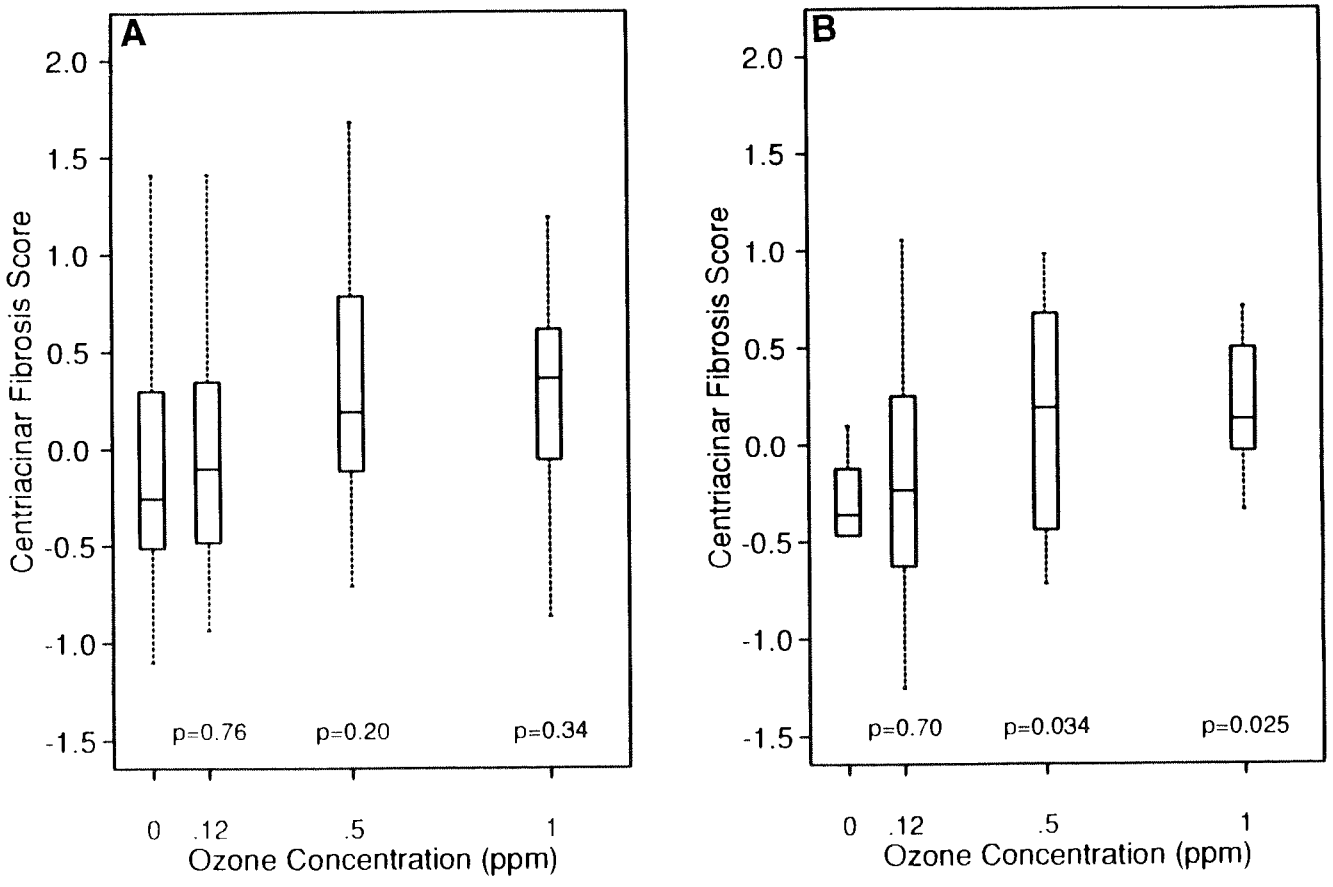

Figure 16. Centriacinar fibrosis analysis by gender, with GAG measures excluded: (A) male data, concentration-response trend $p = 0.30$; (B) female data, concentration response trend $p = 0.0064$.

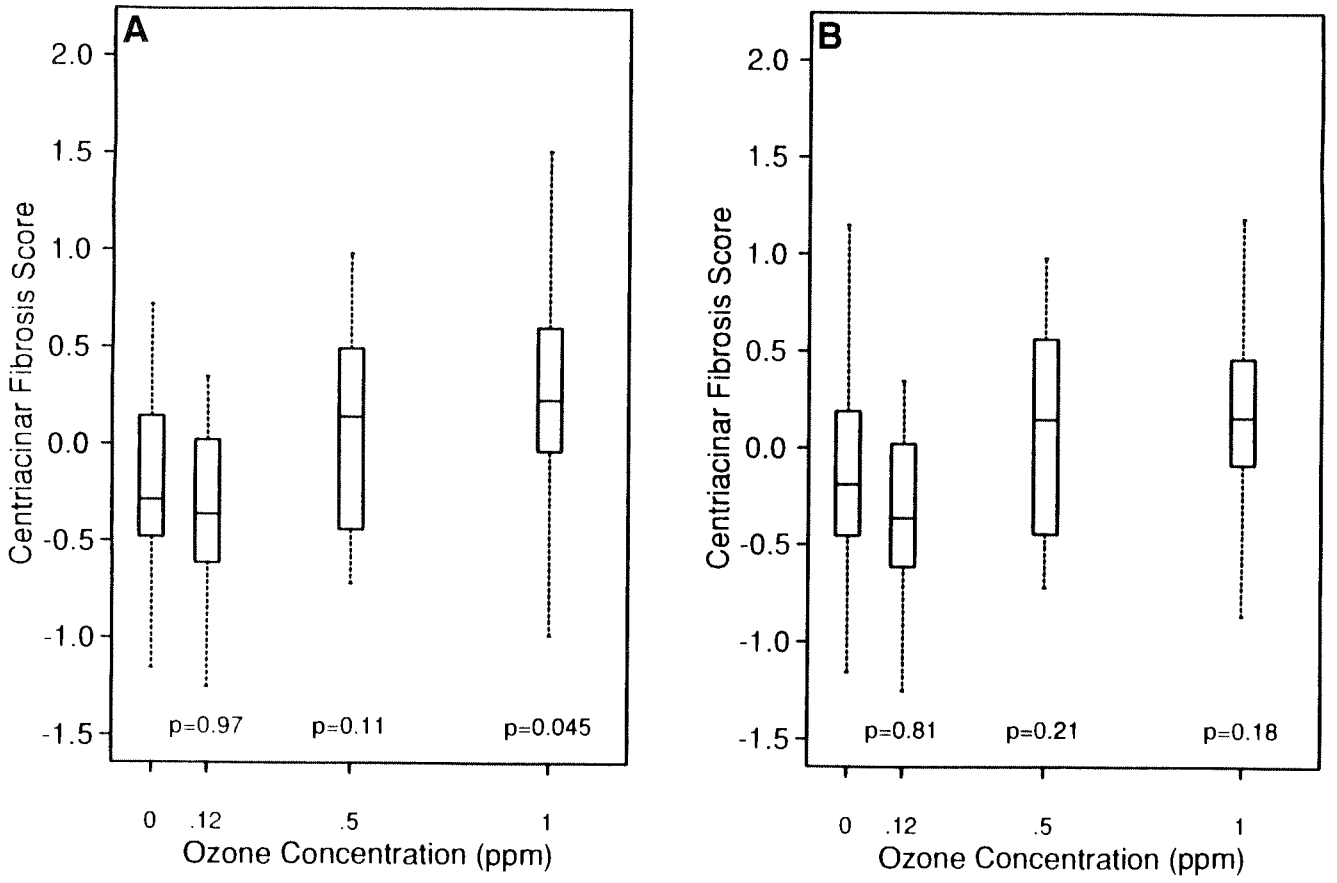

Figure 17. Centriacinar fibrosis analysis excluding animals with leukemia, with GAG measures excluded: (A) excluding animals with advanced leukemia, concentration-response trend $p = 0.013$; (B) excluding animals with any leukemia, concentration-response trend $p = 0.070$.

Table 11. Leukemia Incidence

	All Animals		Control Animals	
Leukemia	Female	Male	Female	Male
Early	2	12	0	4
Advanced	7	16	2	6
Total	9	28	2	10

Figures 27 through 38 correspond to each of the histograms in Figure 26, showing various plots of the simulated data and the settings for the simulation parameters from one run of the simulation. Across the top of Figures 27 through 38, each of the five raw simulated endpoints is plotted against the dose level. Boxplots are used for each dose level to show individual endpoint variability and dose-response trends. Along the bottom are three sets of boxplots that show (left to right) marginal distributions for each endpoint, raw values versus dose (combining endpoints), and the results of performing median polish analysis on the five endpoints and plotting the composite scores (animal effects) versus dose. The test statistics for linear trend from this last plot are reflected in Figure 26 for 100 replications of each scenario. The bottom right of Figures 27 through 38 shows the particular parameters for generating the data. $Mean_i$ refers to the control mean value for endpoint i, $Dose_j$ is the mean level at dose j, $Delta_i$ is an indicator of whether or not endpoint i is affected by the exposure (for example, whether the $Dose_j$ values apply, $Delta_i = 1$; or whether the dose response is flat at $Mean_i$, $Delta_i = 0$). The phi parameters refer to changes in variability across endpoints. Phi refers to the SD in each endpoint as a multiple of the value for sigma (also in the table). Thus, setting phi allows for different variances for each endpoint. The parameter rho is the intraanimal correlation and "Missing = no" refers to the fact that for these simulations, there were no missing data in the table (for example, all measurements on all animals were available).

Figure 39 shows results of the same simulation conducted under scenarios with missing data. The missing data

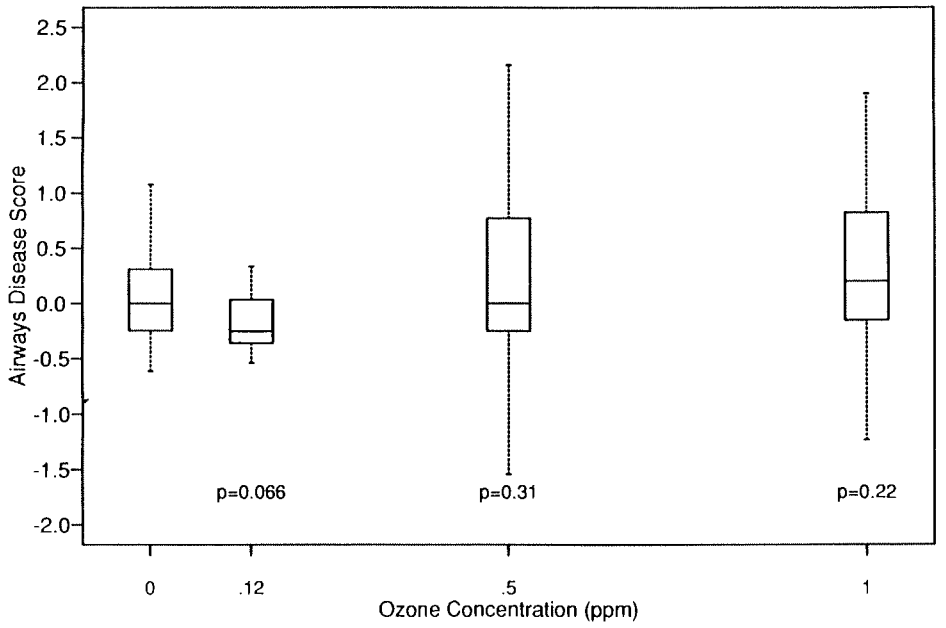

Figure 18. Analysis of airways disease in all animals: concentration-response trend $p = 0.014$.

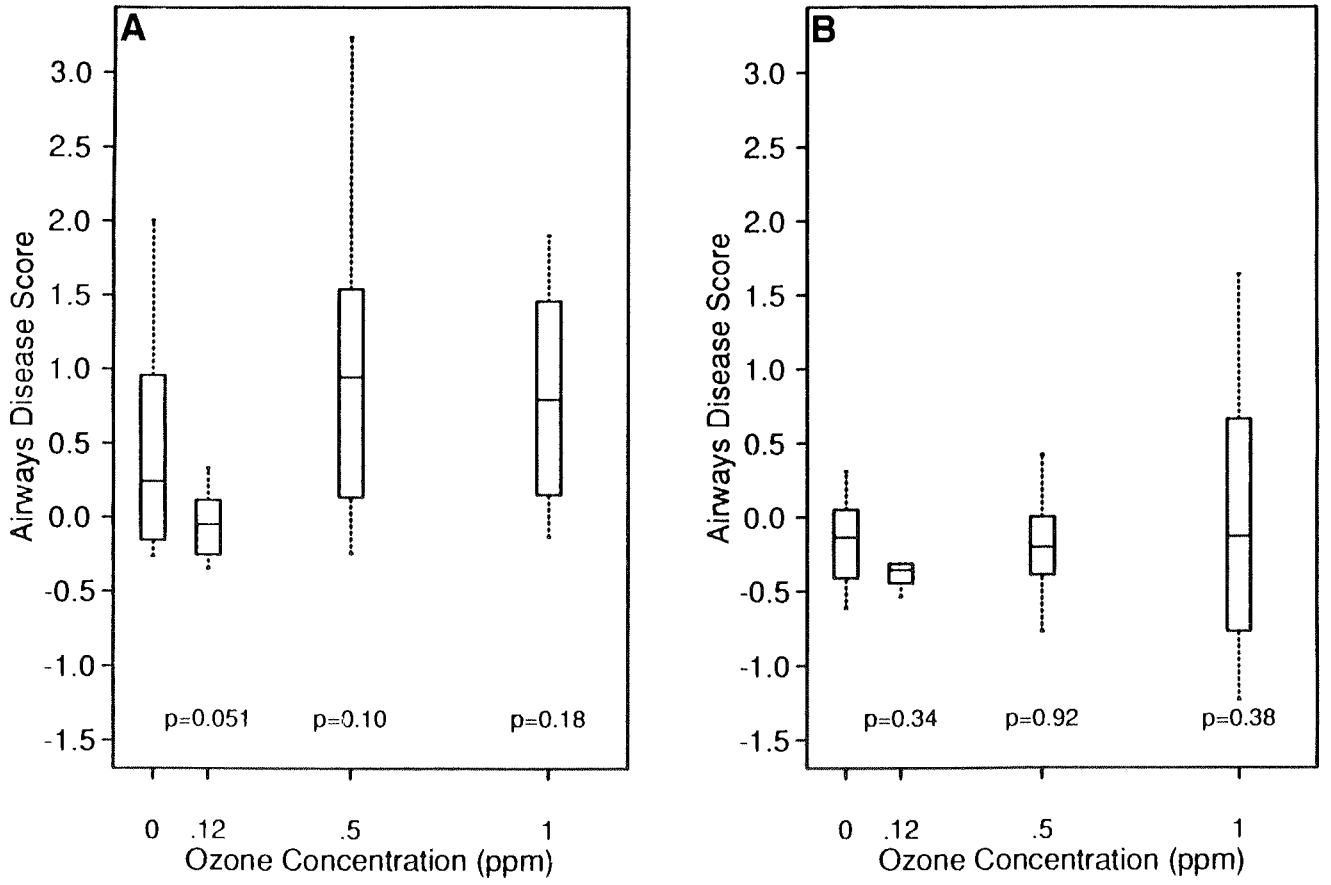

Figure 19. Analysis of airways disease by gender: (A) male data, concentration-response trend $p = 0.033$; (B) female data, concentration response trend $p = 0.16$.

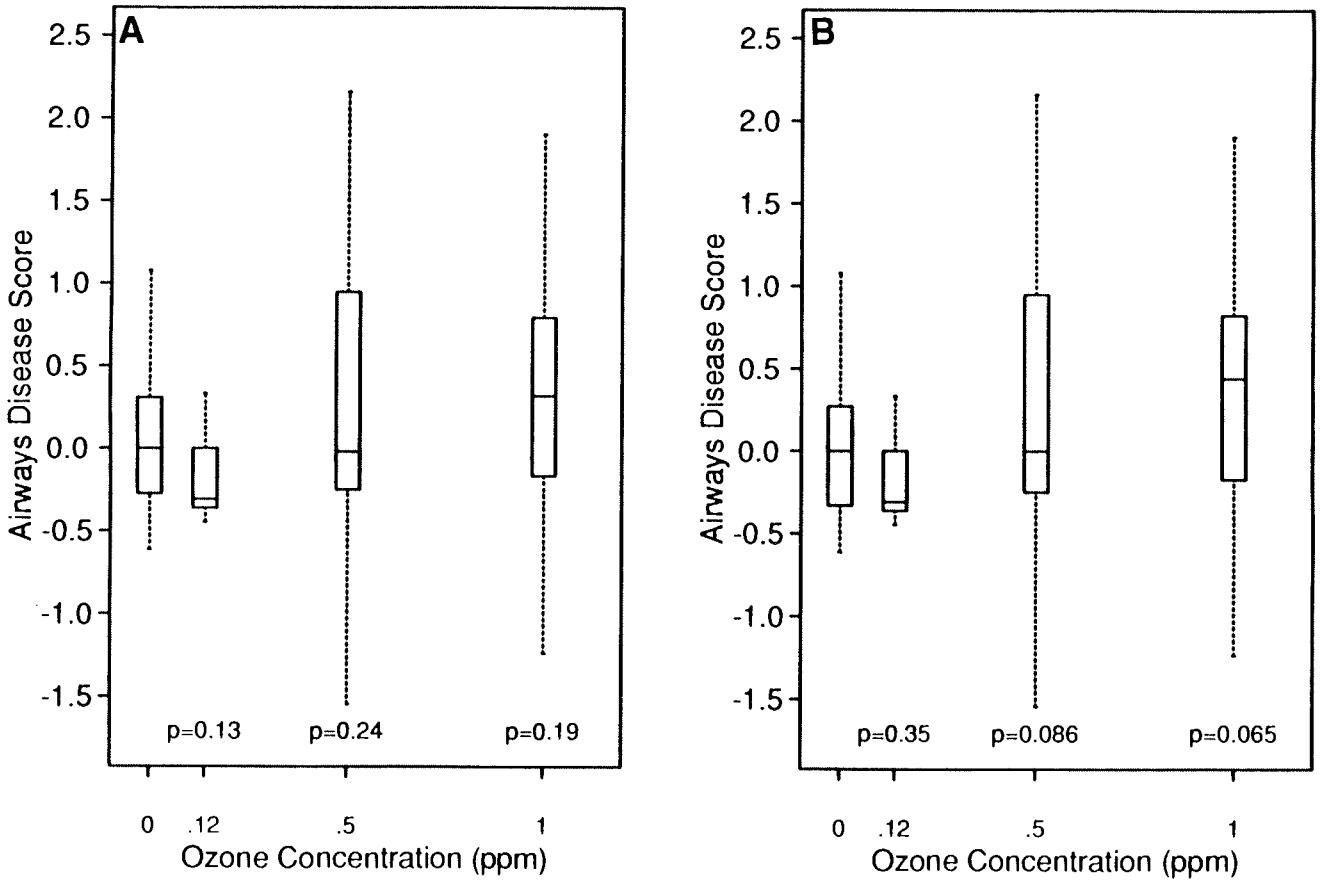

Figure 20. Analysis of airways disease excluding animals with leukemia: (A) excluding animals with advanced leukemia, concentration-response trend $p = 0.016$; (B) excluding animals with any leukemia, concentration-response trend $p = 0.0063$.

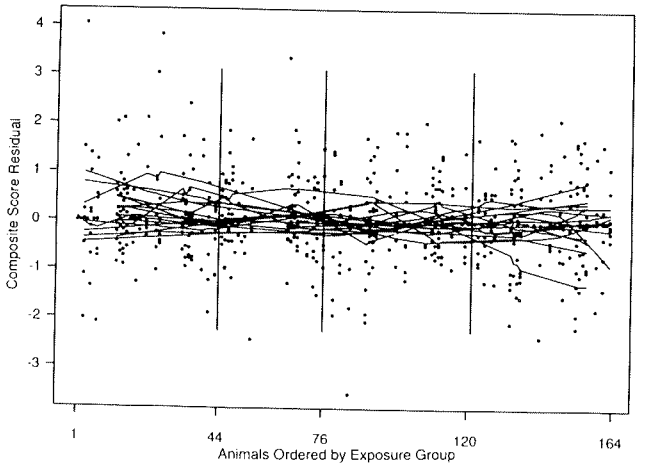

Figure 21. Residuals from the airways disease model. Vertical lines divide the animals by ozone exposure group: from left, 0.0 ppm (control), 0.12 ppm, 0.5 ppm, and 1.0 ppm ozone. Numbers on the horizontal axis indicate available animals in each group.

pattern was chosen to mimic one of the patterns observed in the NTP/HEI Collaborative Ozone Project.

Summary of Simulation Results

In general, the simulations suggest that the median polish approach has considerable power to detect trends, even when the data are quite variable. Also, the method exhibits surprising ability to detect trends even when some individual endpoints are not substantially affected. From the perspective of power, this is desirable. However, there may be potential for confounding if some endpoints, but not all, are affected owing to their relation to some other process. To assess the sensitivity of the method, the impact of deleting one of the endpoints was studied with the rationale that if only one endpoint was driving the whole median polish analysis, it was not providing a true summary. Even when the endpoint with the strongest individual dose-response pattern was removed, the resulting composite scores were not changed dramatically.

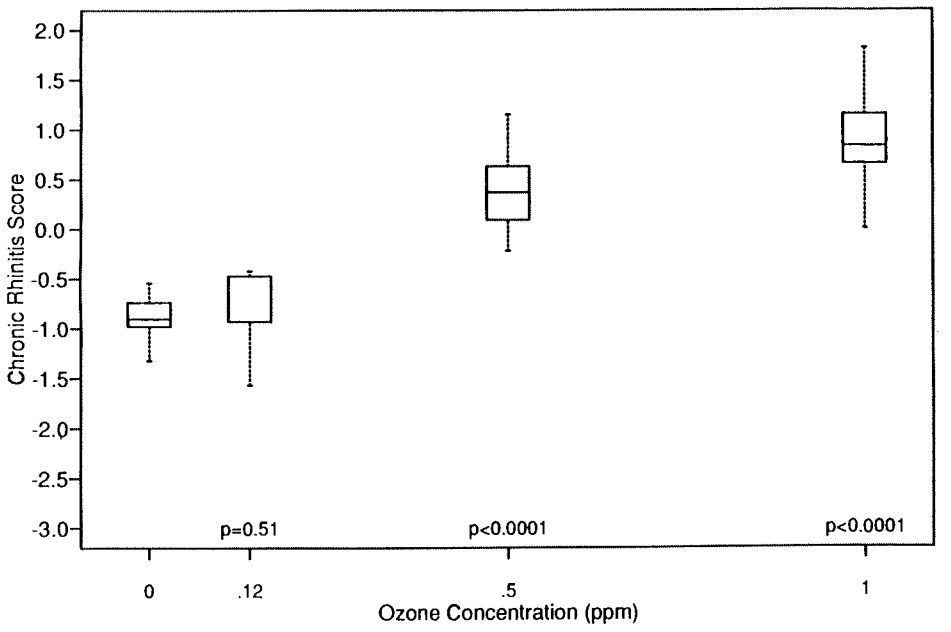

Figure 22. Analysis of chronic rhinitis in all animals: concentration-response trend $p < 0.0001$.

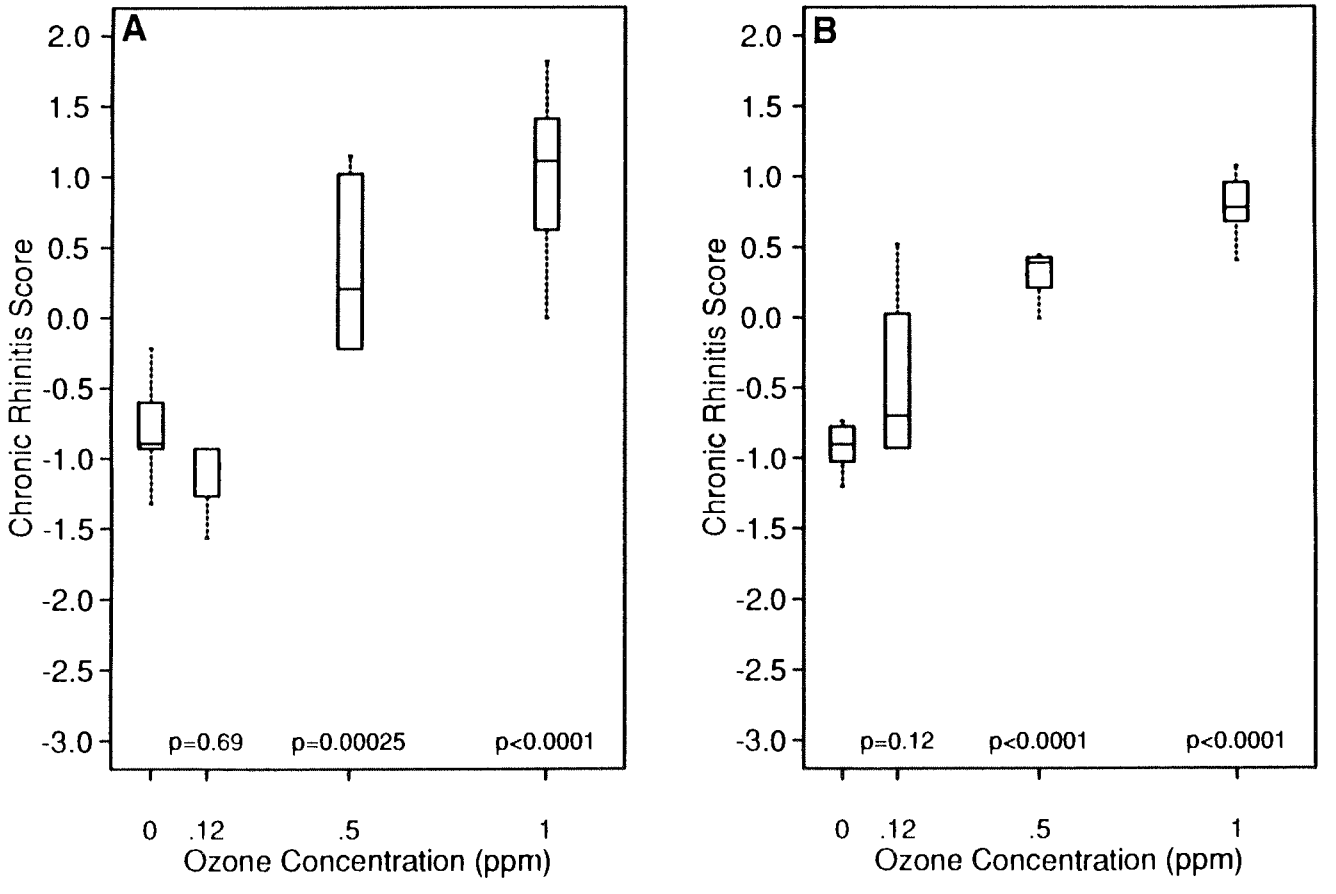

Figure 23. Analysis of chronic rhinitis by gender: (A) male data, concentration-response trend $p < 0.0001$; (B) female data, concentration-response trend $p < 0.0001$.

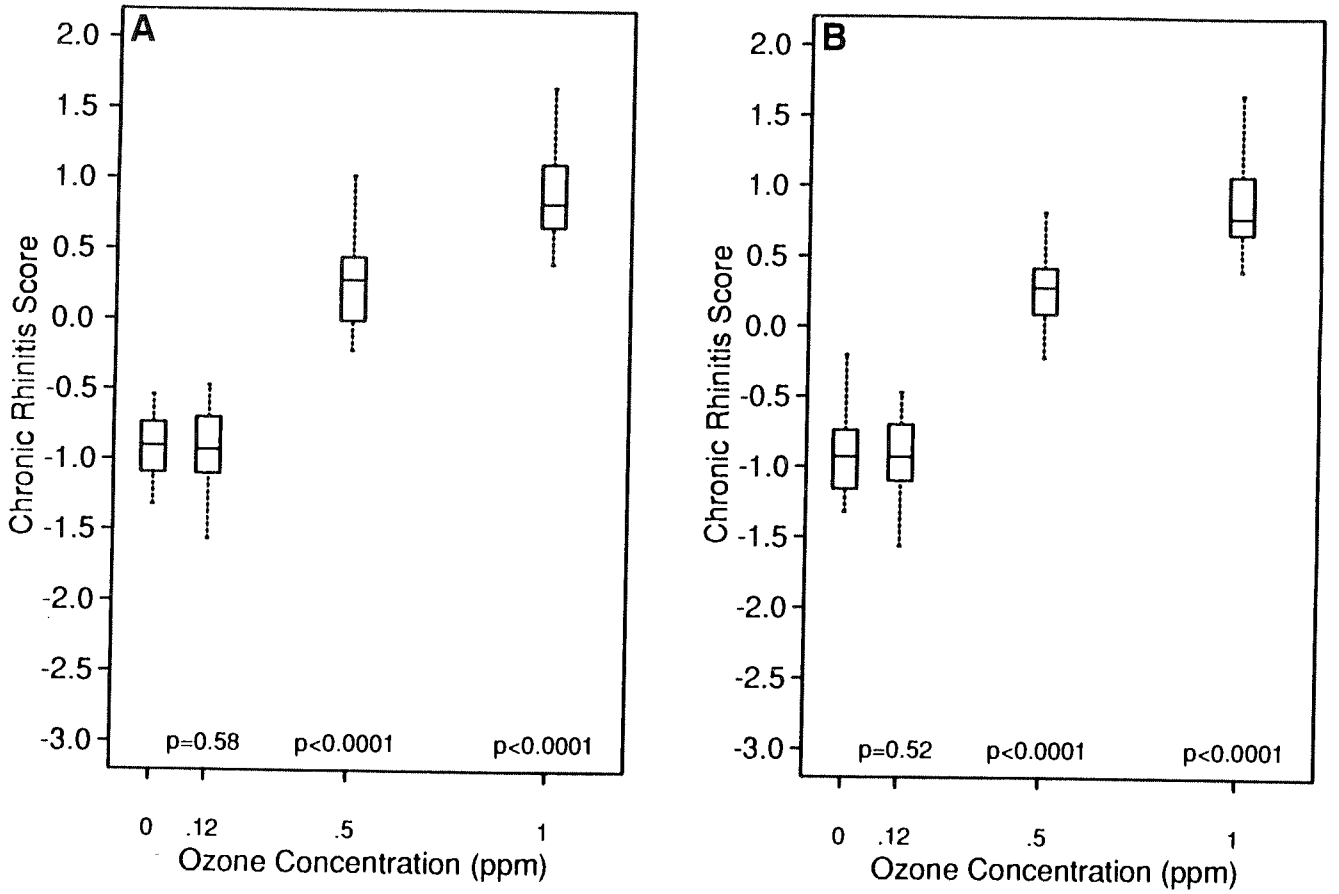

Figure 24. Analysis of chronic rhinitis excluding animals with leukemia: (A) excluding animals with advanced leukemia, concentration-response trend $p < 0.0001$; (B) excluding animals with any leukemia, concentration-response trend $p < 0.0001$.

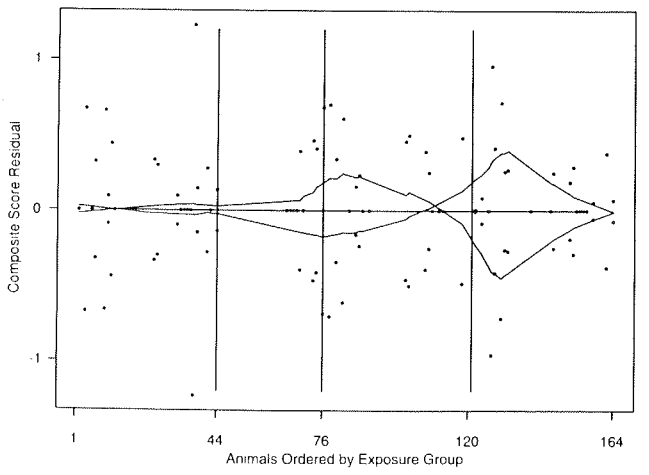

Figure 25. Residuals from the chronic rhinitis model. Vertical lines divide the animals by ozone exposure group: from left, 0.0 ppm (control), 0.12 ppm, 0.5 ppm, and 1.0 ppm ozone. Numbers on the horizontal axis indicate available animals in each group.

As can be seen from Figure 39, results of the simulations incorporating missing data were similar to the complete data case in Figure 26. Of course, there is some reduction in the values of the test statistics in this setting because the sample sizes are reduced with missing data, but in general the simulations show that the method remains quite powerful even when significant portions of the data table are unavailable.

COMPARISON WITH ANOTHER METHOD

To assess the performance of our approach based on median polish analysis, it is useful to compare its results with those obtained using more classic approaches to the analysis of multiple outcomes. As described earlier in this report, methods such as factor analysis or principal components analysis cannot be applied in this setting because of the extensive pattern of missing data. The alternative approach considered in this section is based on generalized estimating equations (GEEs) (Liang and Zeger 1986; Zeger and Liang 1986). This approach uses the ideas of M-estimators (Serfling 1980), by solving equations that involve linear

Table 12. Sensitivity of p Values from the Median Polish Analysis for the Airway Disease Surrogate

	p Values (Two-Sided)			
Model[a]	Trend	0.12 ppm O_3b	0.5 ppm O_3	1.0 ppm O_3
All endpoints	0.014	0.066	0.31	0.22
Mean midexpiratory flow	0.000026	0.20	0.52	0.0021
Central region stored AB/PAS-positive material	0.0078	0.077	0.25	0.15
Caudal region stored AB/PAS-positive material	0.011	0.052	0.33	0.20
Cranial region stored AB/PAS-positive material	0.016	0.090	0.26	0.22
Small airway smooth muscle	0.023	0.084	0.43	0.26
Large airway smooth muscle	0.015	0.059	0.42	0.20
Bronchiolarization	0.013	0.12	0.30	0.19
Epithelial distance	0.022	0.061	0.36	0.27
Shift from ciliated to Clara cells	0.027	0.12	0.33	0.29
Central region percentage nonciliated cells	0.011	0.070	0.28	0.20
Caudal region percentage nonciliated cells	0.014	0.087	0.26	0.20
Cranial region percentage nonciliated cells	0.013	0.069	0.30	0.20
Small airway maximal tension	0.031	0.062	0.40	0.30
Large airway maximal tension	0.019	0.074	0.30	0.25
Small airway EC_{50}	0.0070	0.057	0.26	0.17
Large airway EC_{50}	0.012	0.069	0.27	0.21
Eicosanoid production	0.0085	0.072	0.31	0.16
Epithelial volume	0.012	0.052	0.31	0.20

[a] When an individual endpoint is given, the values are for analyses conducted excluding that endpoint.

[b] The direction of change from control values is opposite to the trend.

combinations of observed values minus expected values. The GEE approach described by Liang and Zeger also allows for repeated measures on the same subjects. Hence, the approach can be applied to model the endpoints listed under the three composite scores in Tables 1, 2, and 3 as a function of ozone exposure. To apply GEEs to the multiple endpoints listed in these tables, we first had to standardize the variables, just as we did for the median polish analysis. The standardization step corresponded to subtracting the mean and dividing by the SD for each endpoint. An appropriate sign was also attached so that all endpoints within each disease surrogate would have the same expected direction with respect to ozone exposure. For instance, total collagen content is expected to increase with ozone exposure, so it has a positive sign for the centriacinar fibrosis analysis; because RV is expected to decrease with increasing ozone exposure, this parameter has a negative sign.

Using the GEE approaches, values corresponding to trend tests applied to the centriacinar fibrosis, airway disease, and chronic rhinitis endpoints were $p = 0.133$, $p = 0.001$, and $p < 0.001$, respectively. These values compare well with the p values obtained from the median polish analysis described previously ($p = 0.091$, $p = 0.014$, and $p < 0.001$ for centriacinar fibrosis, airway disease, and chronic rhinitis, respectively). Thus, the results of our procedure based on median polish analysis provide results qualitatively similar to those based on GEEs (in terms of testing for the effects of ozone concentration).

DISCUSSION AND CONCLUSIONS

Analogously to a standard ANOVA, median polish analysis provides results that decompose the original data table into an overall effect, a column or endpoint effect (one per endpoint, measuring, across all animals, the change of each endpoint from the overall effect), a row or animal effect (one per animal, measuring, across all endpoints, the change of each animal from the overall effect), and a residual effect (one per observation, measuring the unexplainable variation left over from each measurement). For our purposes, interest centers on the row effects, which for each animal are the composite scores across all endpoints. These composite scores measure, with a single-number summary, the

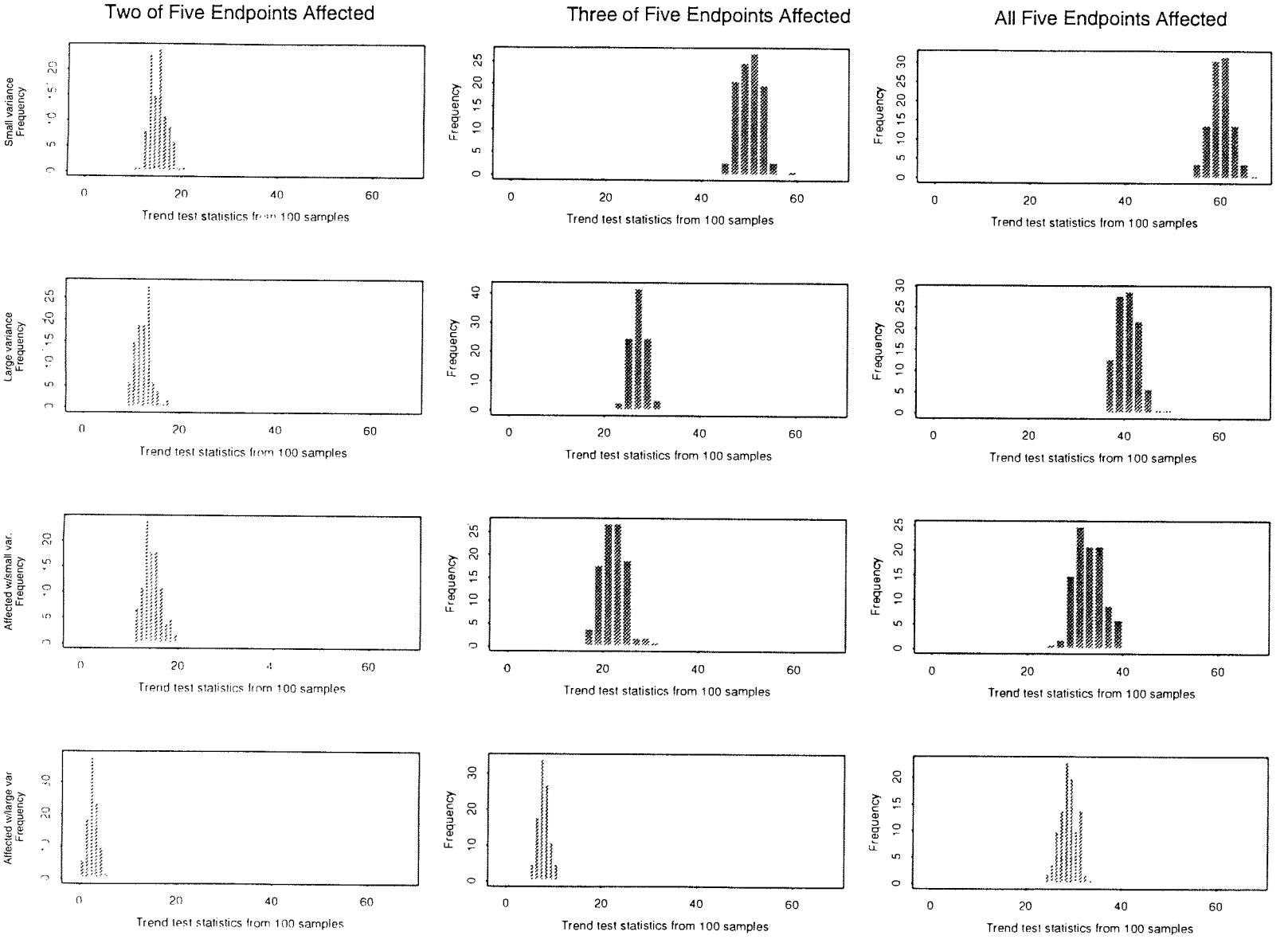

Figure 26. Simulation trend test statistics under 12 scenarios.

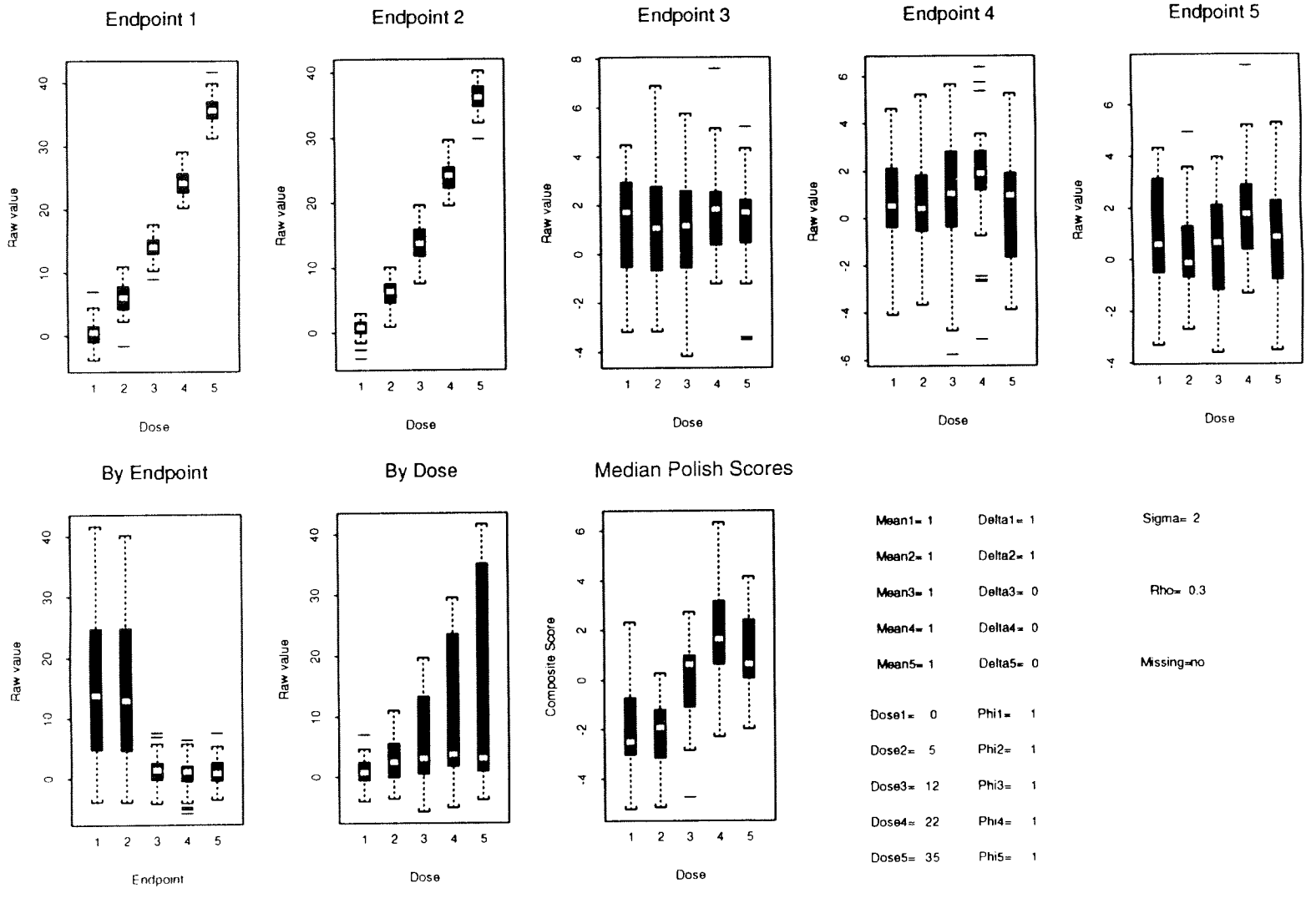

Figure 27. Simulation detail: two of five endpoints are affected with within-dose variances that are equal and small.

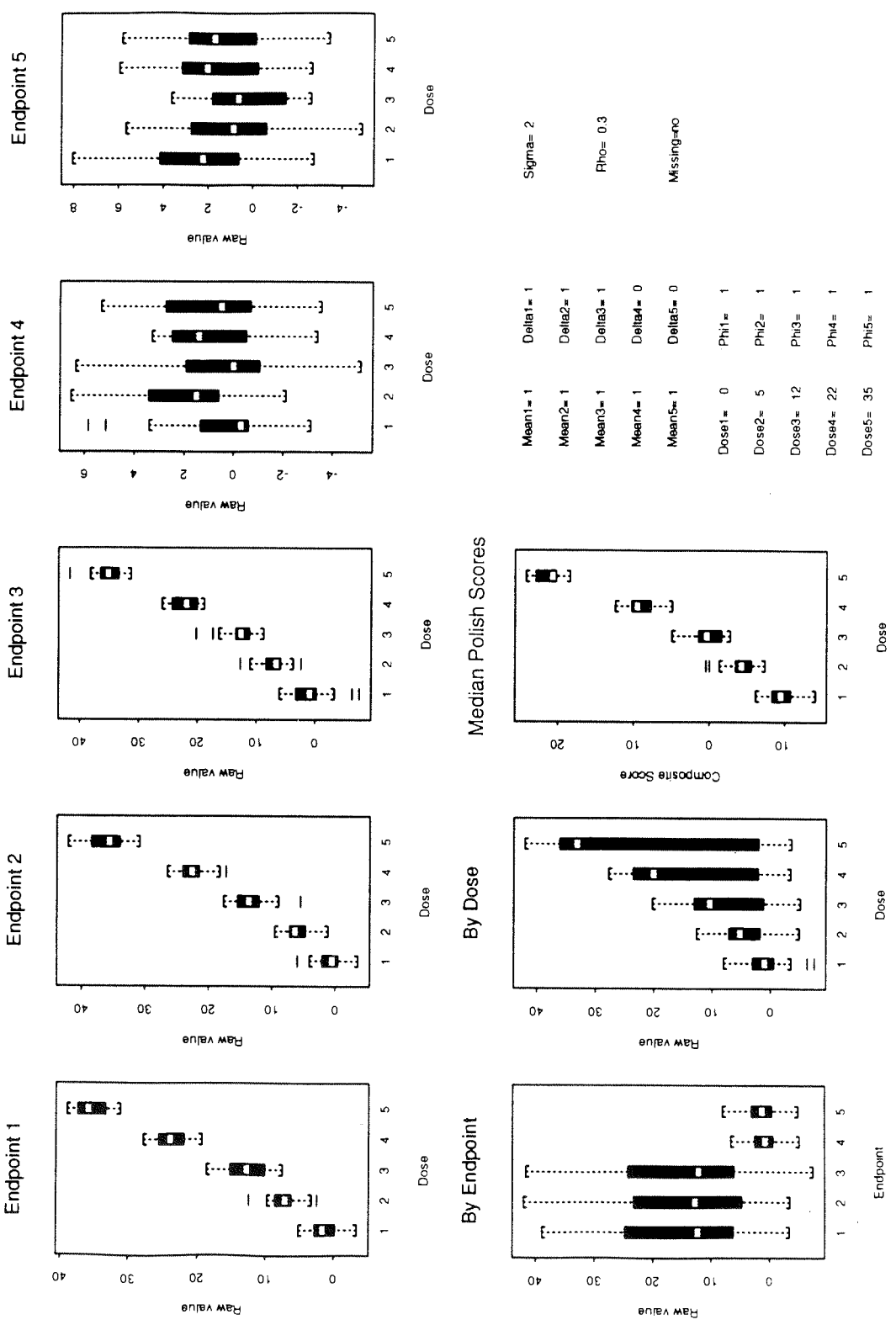

Figure 28. Simulation detail: three of five endpoints are affected with within-dose variances that are equal and small.

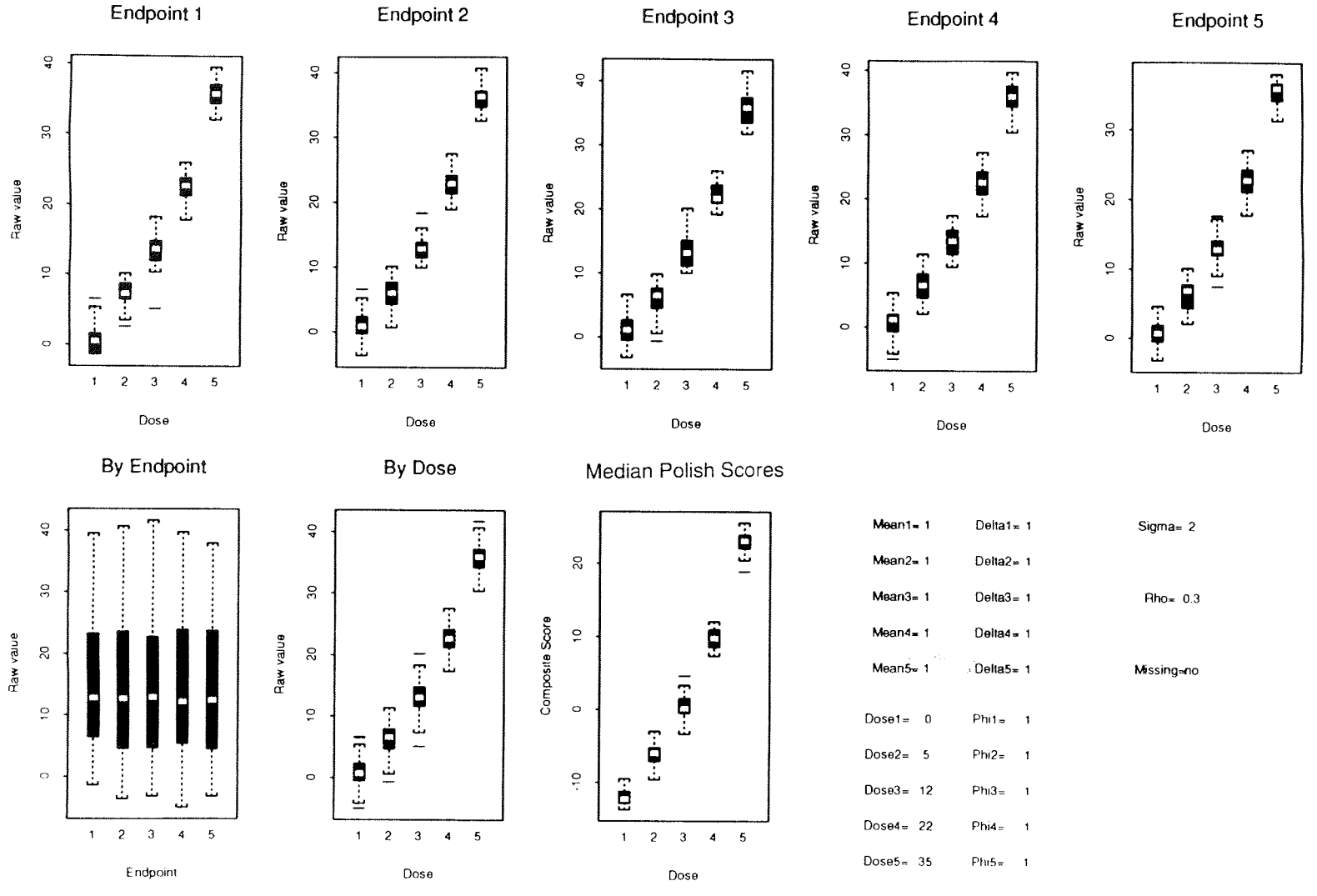

Figure 29. Simulation detail: all five endpoints are affected with within-dose variances that are equal and small.

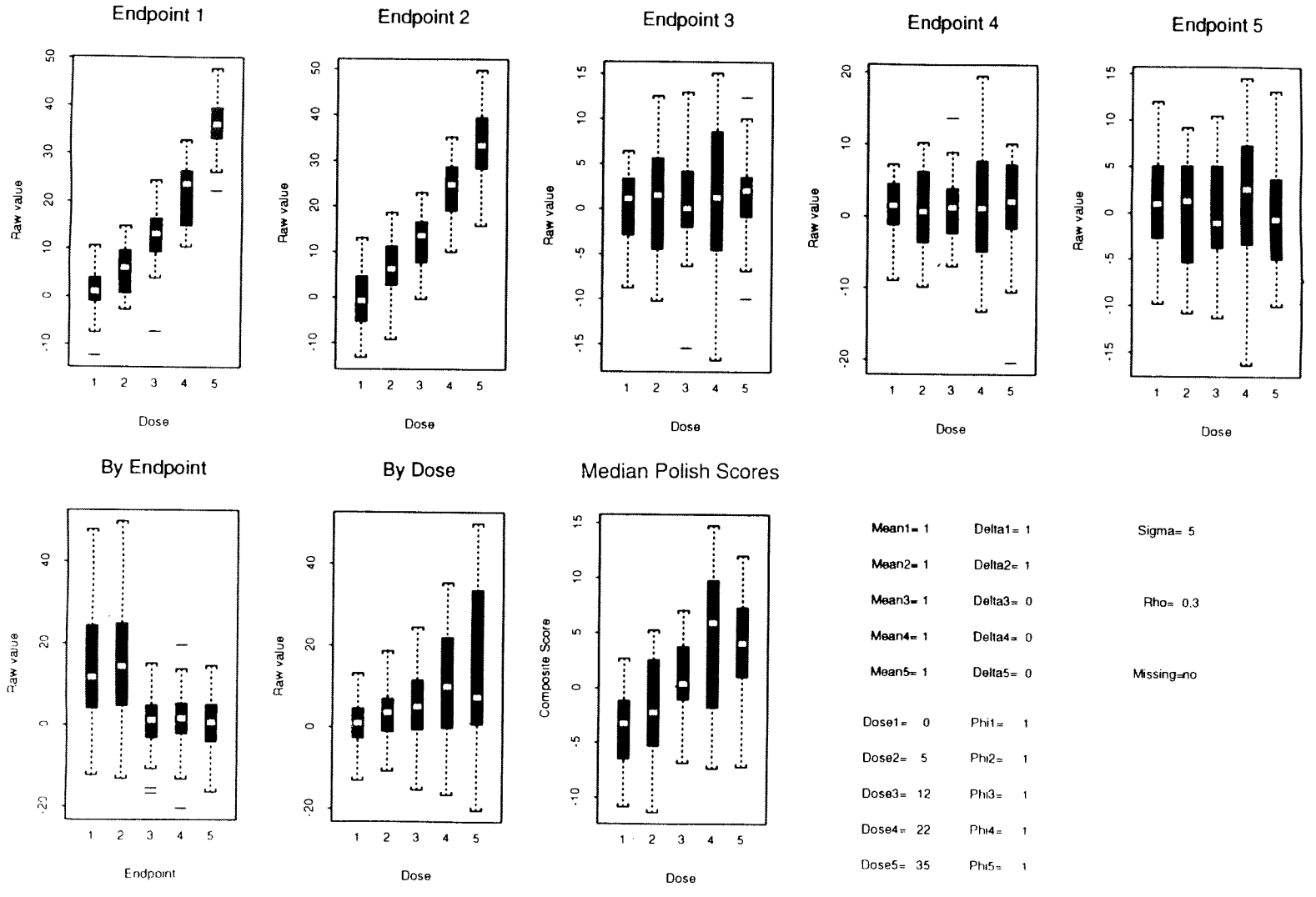

Figure 30. Simulation detail: two of five endpoints are affected with within-dose variances that are equal and large.

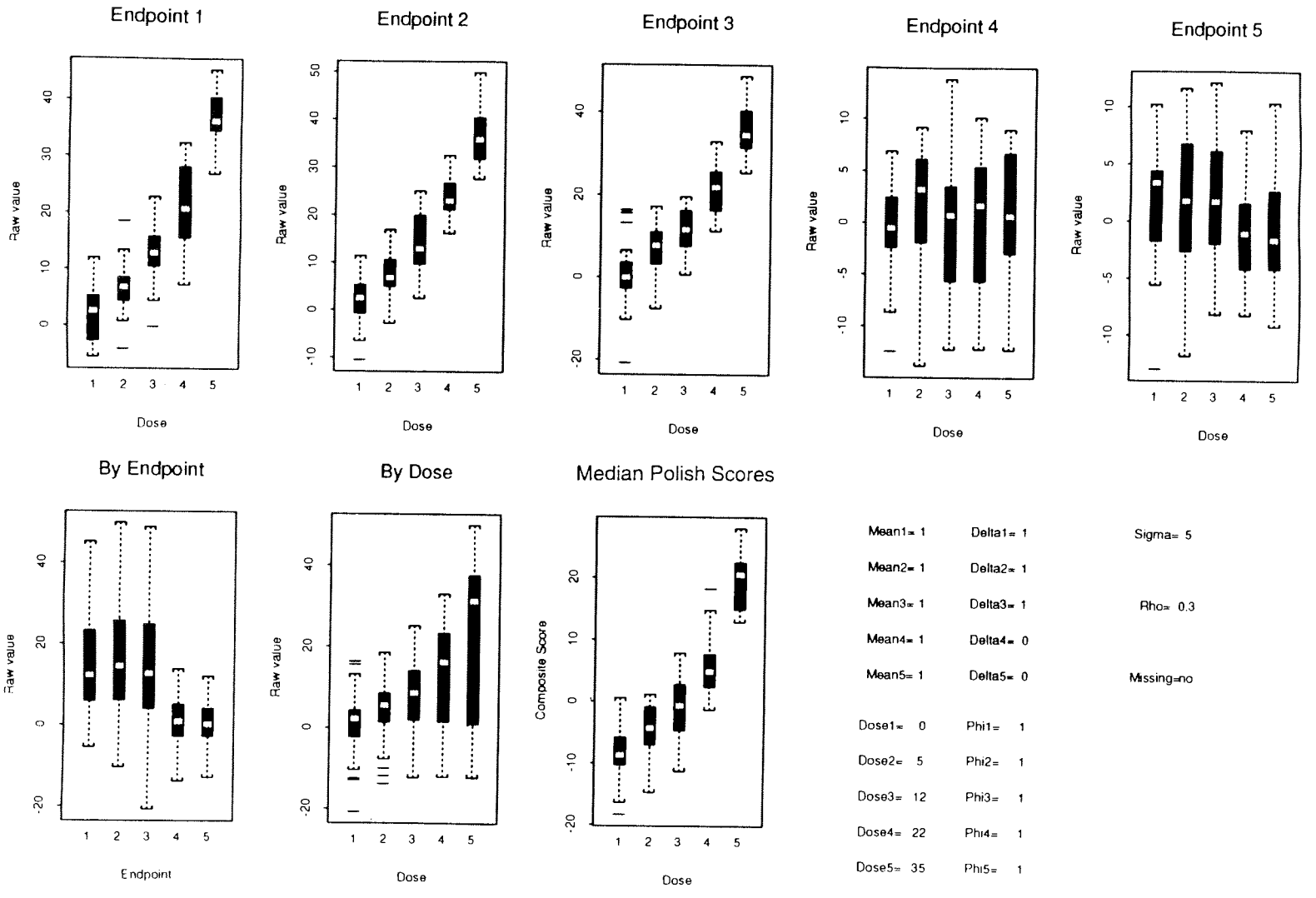

Figure 31. Simulation detail: three of five endpoints are affected with within-dose variances that are equal and large.

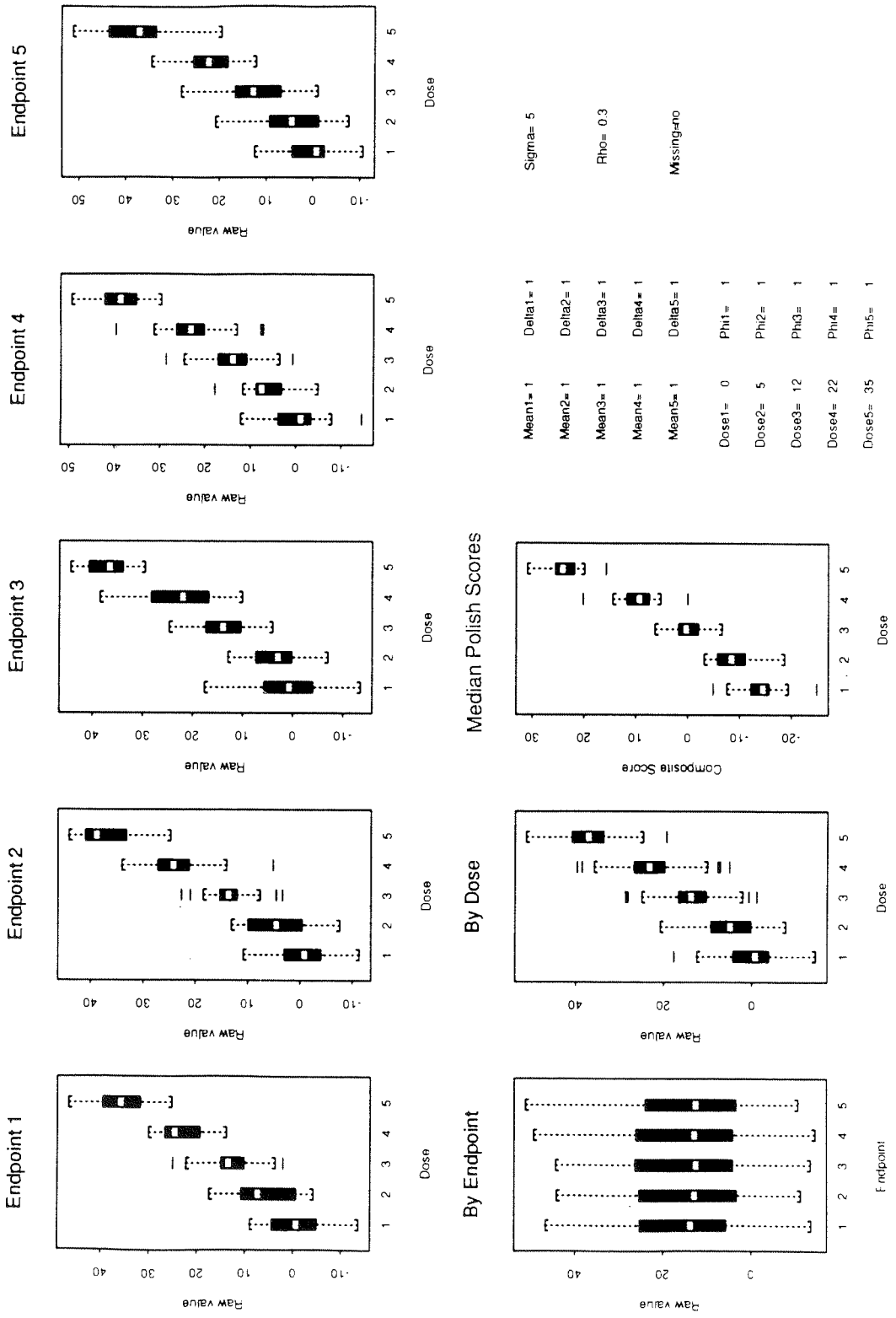

Figure 32. Simulation detail: all five endpoints are affected with within-dose variations that are equal and large.

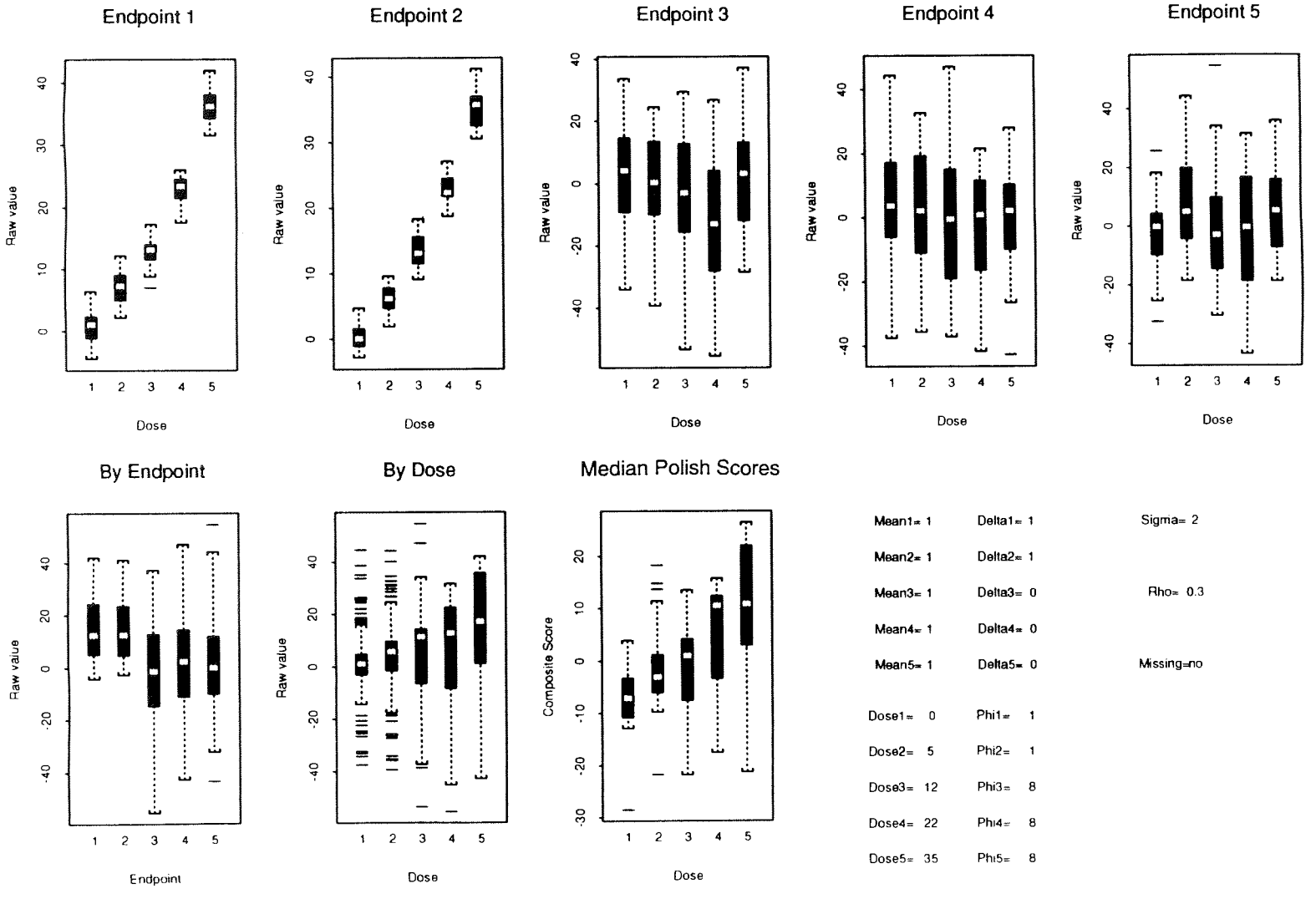

Figure 33. Simulation detail: two of five endpoints are affected within-dose variances that are unequal and smaller in the affected endpoints.

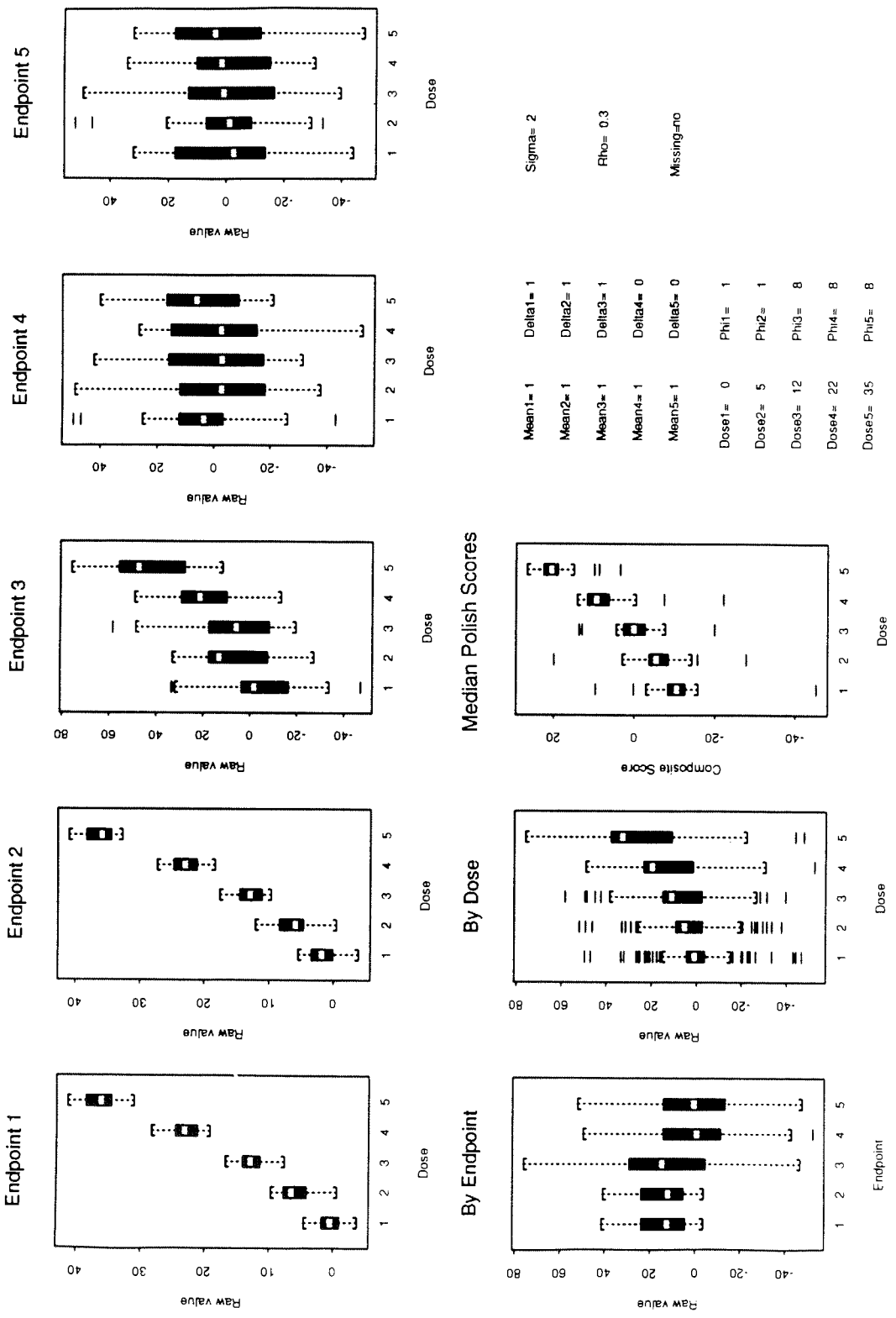

Figure 34. Simulation detail: three of five endpoints are affected with within-dose variances that are unequal and smaller in the affected endpoints.

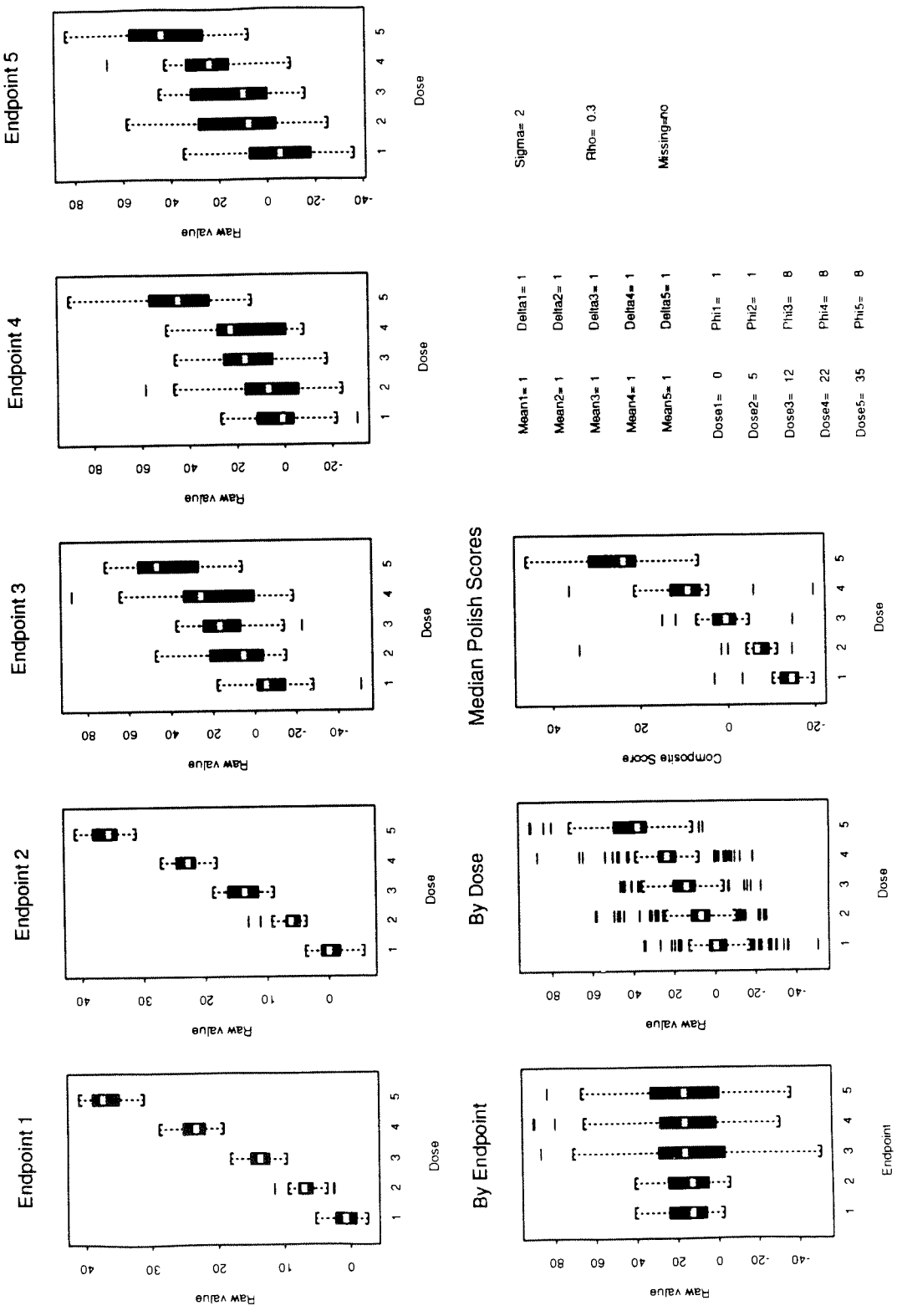

Figure 35. Simulation detail: all five endpoints are affected with within-dose variances that are unequal and smaller in the affected endpoints.

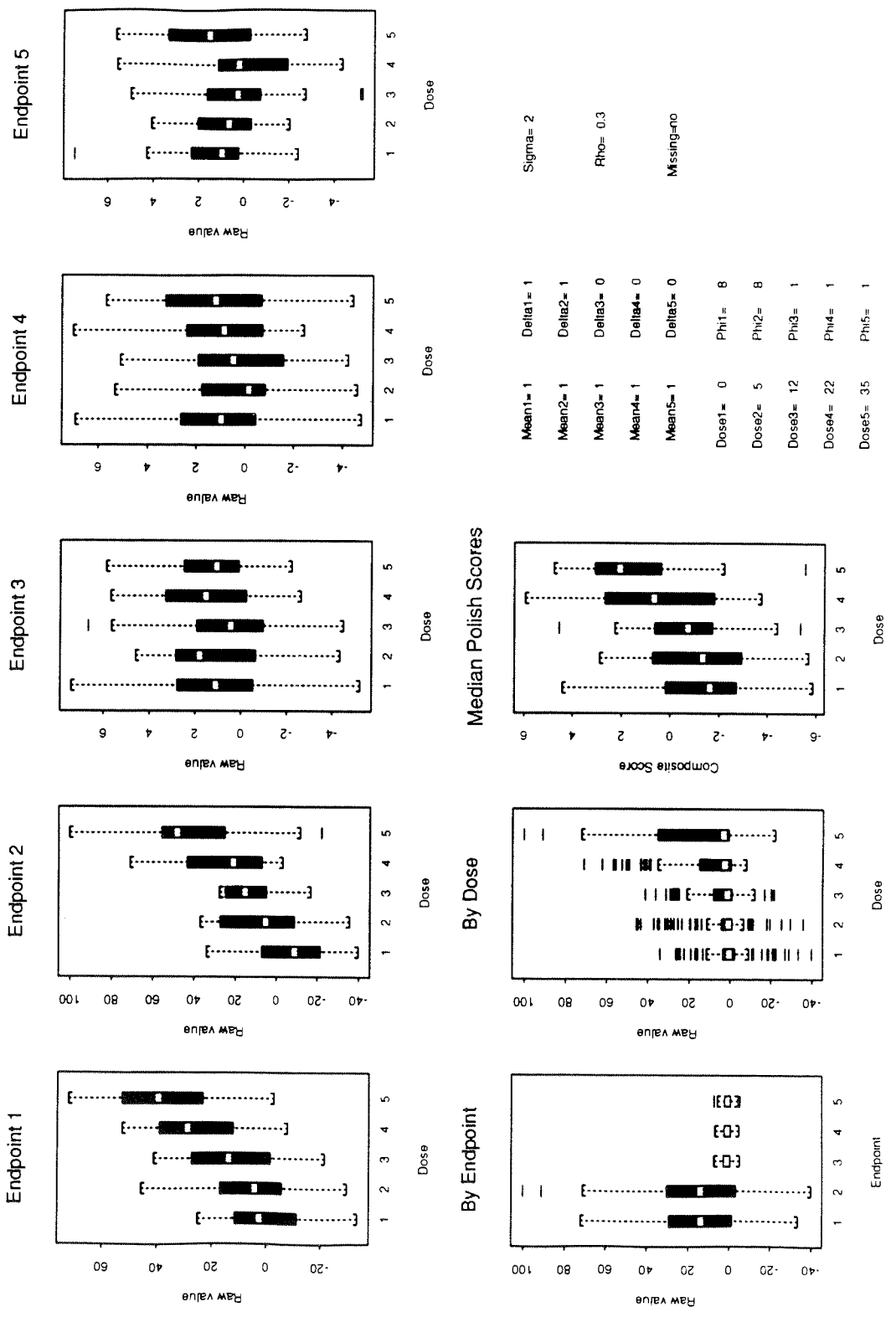

Figure 36. Simulation detail: two of five endpoints are affected with within-dose variances that are unequal and larger in the affected endpoints.

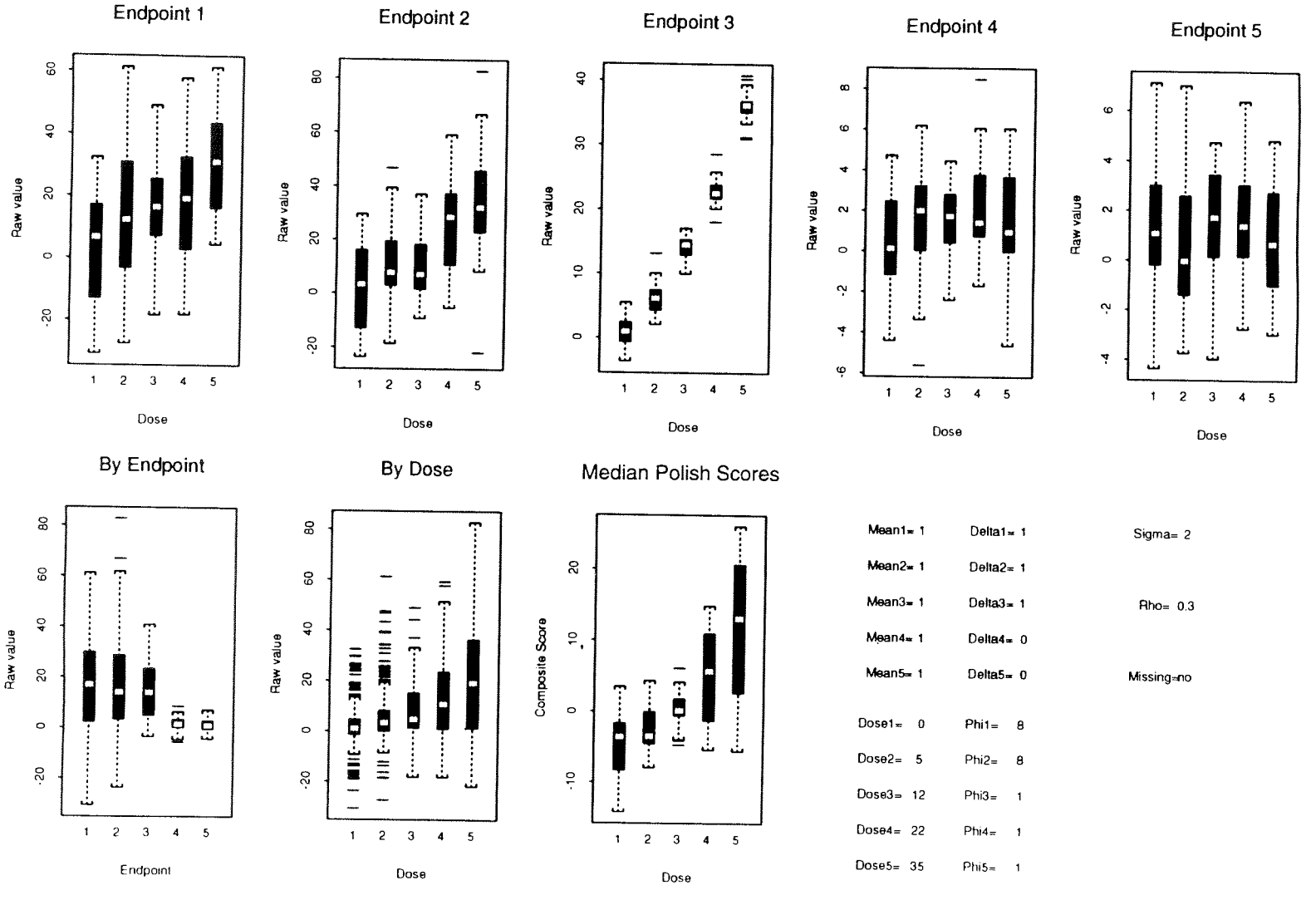

Figure 37. Simulation detail: three of five endpoints are affected with within-dose variances that are unequal and larger in the affected endpoints.

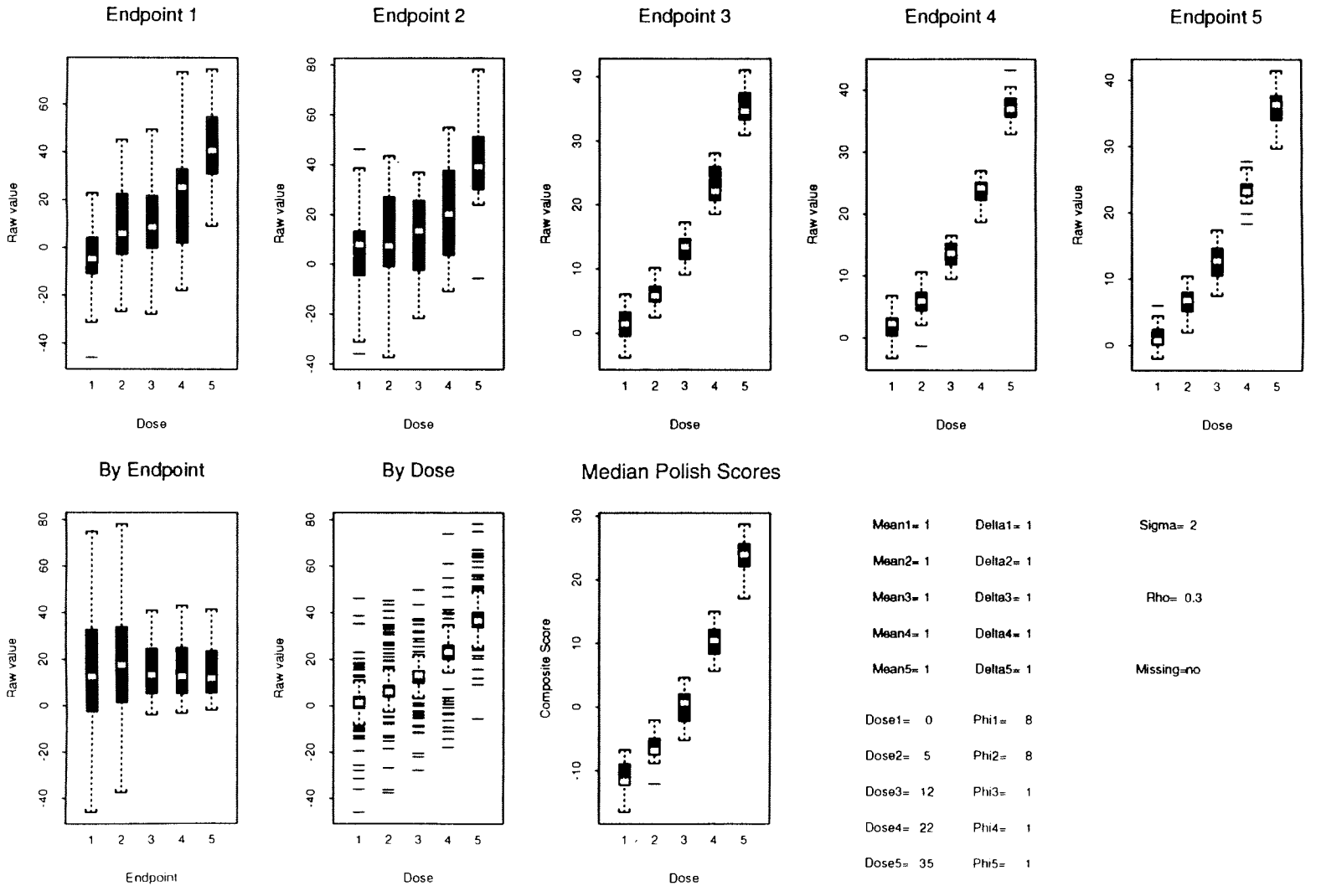

Figure 38. Simulation detail: all five endpoints are affected with within-dose variances that are unequal and larger in the affected endpoints.

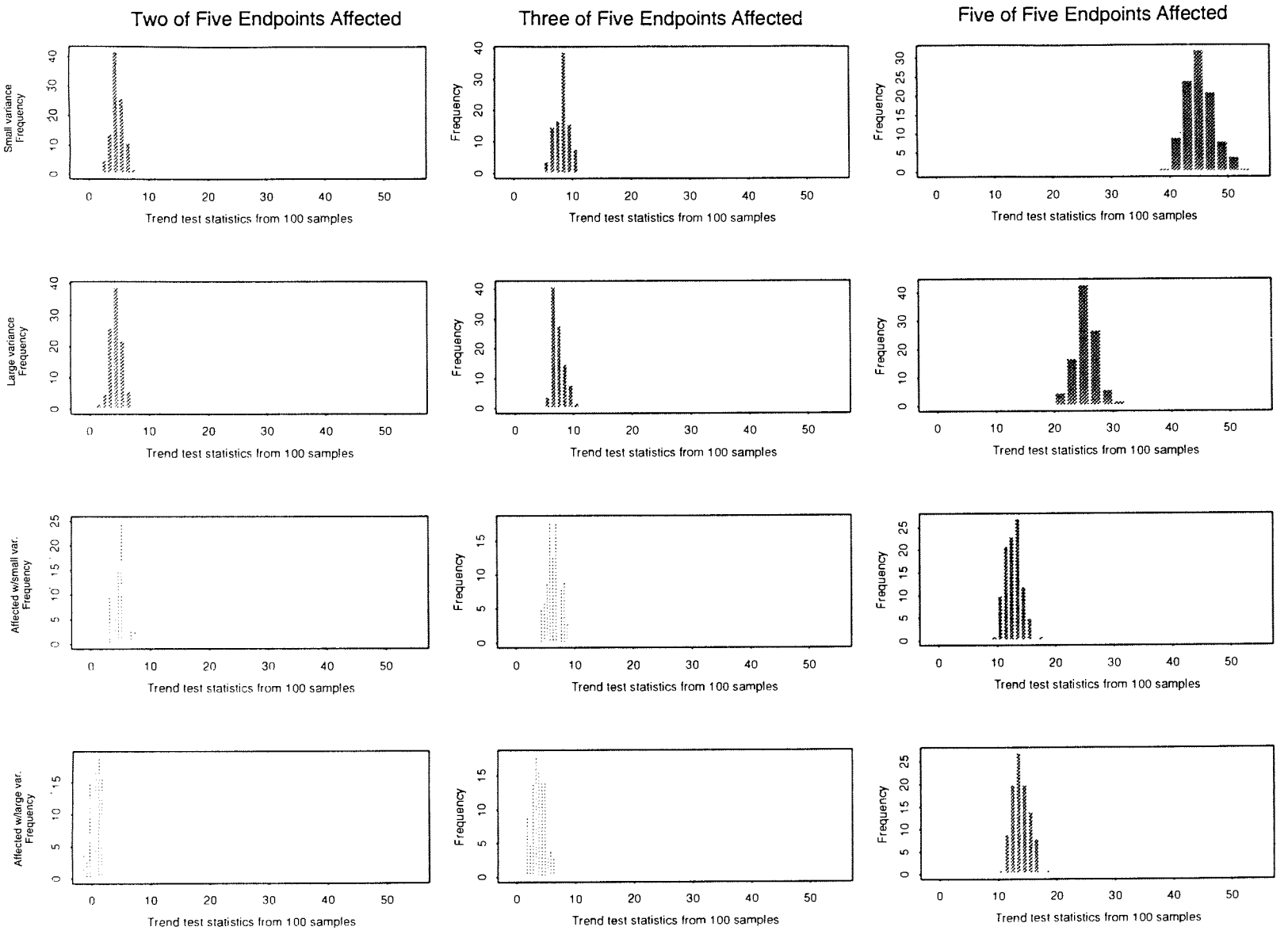

Figure 39. Simulation trend test statistics under 12 scenarios with missing data.

multivariate response of each animal. The initial standardization of the variables (before they are subjected to median polish analysis) provides a convenient interpretation of the composite scores. The score for a particular animal measures, in SD units, the distance of that animal's composite score from the average composite score (which is zero). Animals that are extreme (positive, say) in all of their individual measures will have a composite score that is large and positive (the argument is similar for negative scores). Animals that are average in their individual endpoints or that have a mixture of negative and positive extremes will tend to have composite scores near zero.

The median polish analysis does not address the effects of ozone exposure or any other covariate on the responses; it merely establishes a procedure for obtaining a summary of each animal's responses. Once the composite scores are obtained, the next phase of analysis investigates the relations between the composite scores and ozone exposure level, gender, date animals were killed, and all other concomitant study variables. It is at this phase of analysis that the question of the impact of exposure on response is directly addressed. This portion of the analysis follows more closely standard techniques for assessing concentration-response data. Parametric models can be fit to the composite scores, they can be plotted by exposure group, and parametric or nonparametric tests comparing each group with the control group can be obtained as well as tests of trend. More sophisticated multivariate analyses can also be conducted when study variables are evaluated and used as adjustment cofactors in a standard regression model setting.

The proposed method raises several challenging and interesting issues. It relies on the user to specify the endpoints that should be included in the construction of robust composite scores. The user must also correctly specify the sign associated with each variable: that is, the variables should have a positive sign if they are anticipated to increase with increasing exposure, a negative sign if they are anticipated to decrease. Once specified for inclusion in construction of the composite scores, all variables are given equal weight. Ideally, one would like to apply a data reduction technique, such as principal components or factor analysis, that allowed for differential weighting of variables and that allowed the data to determine the appropriate sign. Furthermore, it might be preferable to have some kind of data-driven technique (like principal components) to determine which variables should be included in the score construction. As discussed previously, however, these classic techniques cannot be easily adapted to handle the complex patterns of missing data encountered in this study. Development of appropriate methods would be an interesting and worthwhile topic for statistical research.

Some might argue that the features of median polish analysis outlined in the previous paragraph are serious limitations. From a different perspective, however, median polish analysis may have an advantage over data-driven techniques such as principal components analysis. Principal components analysis tends to group together variables that behave similarly from individual to individual. This tendency is an advantage when the goal is to find the combination of variables that best predict the behavior of separate individual variables. However, this tendency is a disadvantage when analyzing variables that have purposely been selected to give, collectively, the best representation of a chosen phenomenon. In this situation, median polish analysis, by synthesizing all the selected variables, may have an advantage over principal components analysis. Hence, one could argue that even if principal components analysis had been technically feasible (which it wasn't because of missing data), median polish analysis might still have been preferable.

REFERENCES

American Lung Association. 1989. Breath in Danger. Estimation of Populations-at-Risk of Adverse Health Consequences in Areas Not in Attainment with National Ambient Air Quality Standards (NAAQS) of the Clean Air Act, pp. 2–3. American Lung Association, New York, NY.

Boorman GA, Catalano PJ, Jacobson BJ, Kaden DA, Mellick PW, Nauss KM, Ryan LM. 1995. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part VI, Background and Study Design. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Catalano PJ, Chang L-Y, Harkema J, Last JA, Kaden DA, Mellick PW, Parks WC, Pinkerton KE, Radhakrishnamurthy B, Ryan LM, Szarek JL. 1995. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part XI, Integrative Summary. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Chang, L-Y, Stockstill B, Menache MG, Mercer RR, Crapo JD. 1995. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part VIII, Morphometric Analysis of Structural Alterations in Alveolar Regions. Research Report Number 65. Health Effects Institutes, Cambridge, MA.

Cleveland WS. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74:829–836.

Fink AM. 1988. How to polish off median polish. SIAM J Sci Statist Comput 9:932–940.

Harkema J, Mauderly JL. 1994. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part V, Effects on Pulmonary Function. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Harkema J, Morgan KT, Bermudez EG, Catalano PJ, Griffith WC. 1994. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part VII, Effects on the Nasal Mucociliary Apparatus. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Hoaglin DC, Mosteller F, Tukey JW. 1983. Understanding Robust and Exploratory Data Analysis. John Wiley & Sons, New York, NY.

Johnson RA, Wichern DW. 1992. Applied Multivariate Statistical Analysis, 3rd ed. Prentice Hall, Englewood Cliffs, NJ.

Last JA, Gelzleichter TR, Harkema J, Hawk S. 1994. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part I, Content and Cross-Linking of Lung Collagen. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Liang K-Y, Zeger SL. 1986. Longitudinal data analysis using generalized linear models. Biometrika 73:13–22.

Morrison, DF. 1990. Multivariate Statistical Methods, 3rd ed. McGraw-Hill, New York, NY.

Mosteller F, Tukey JW. 1977. Data Analysis and Regression: A Second Course in Statistics. Addison-Wesley Publishing Co., Reading, MA.

Pinkerton KE, Ménache MG, Plopper CG. 1995. In: Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part IX, Changes in the Tracheobronchial Epithelium, Pulmonary Acinus, and Lung Antioxidant Enzyme Activity. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Radhakrishnamurthy B. 1994. Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies, Part III, Effects on Complex Carbohydrates of Lung Connective Tissue. Research Report Number 65. Health Effects Institute, Cambridge, MA.

APPENDIX A. Individual Animals and Endpoints for the Three Disease Surrogates Investigated

Table A.1. Centriacinar Fibrosis Endpoints by Animal[a]

Animal	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H1	*							*	*	*
H2	*					*	*	*	*	*
H3	*					*	*	*	*	*
H4	*					*	*	*	*	*
H5	*					*	*	*	*	*
H6	*					*	*	*	*	*
H7						*	*	*	*	*
H8						*	*	*	*	*
H9						*	*	*	*	*
H10						*	*	*	*	*
H11								*	*	*
H12								*	*	*
H13								*	*	*
H14								*	*	*
H15								*	*	*
H16								*	*	*
H17								*	*	*
H18								*	*	*
H19	*							*	*	*
H20	*							*	*	*
H21	*							*	*	*
H22	*									
H23	*							*	*	*
H24	*							*	*	*
H25						*	*			
H26						*	*			
H27						*	*			
H28						*	*			
H29										
H30						*	*			
H31						*	*			
H32						*	*			
H33						*	*			
H34						*	*			
H35						*	*			

(Table continues next page.)

[a] See Table 1 for an explanation of the codes used in column headings.

Table A.1. Centriacinar Fibrosis Endpoints by Animal[a] *(continued)*

Animal	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H36						*	*			
H37		*	*	*	*					
H38		*	*	*	*					
H39		*	*	*	*					
H40		*	*	*	*					
H41		*	*	*	*					
H42		*	*	*	*					
H43		*	*	*	*					
H44		*	*	*	*					
H45		*	*	*	*					
H46		*	*	*	*					
H47		*	*	*	*					
H48		*	*	*	*					
H49		*	*	*	*					
H50		*	*	*	*					
H51		*	*	*	*					
H52		*	*	*	*					
H53								*	*	*
H54		*	*		*			*	*	*
H55								*	*	*
H56								*	*	*
H57		*	*	*	*			*	*	*
H58		*	*	*	*			*	*	*
H59								*	*	*
H60								*	*	*
H61		*	*		*			*	*	*
H62		*	*		*			*	*	*
H63								*	*	*
H64										
H65		*	*	*	*			*	*	*
H66		*	*	*	*			*	*	*
H67				*				*	*	*
H68								*	*	*
H69										
H70										
H71										
H72										
H73										
H74										
H75										

(Table continues next page.)

[a] See Table 1 for an explanation of the codes used in column headings.

Table A.1. Centriacinar Fibrosis Endpoints by Animal[a] *(continued)*

Animal	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H76										
H77										
H78										
H79										
H80										
H81										
H82										
H83										
H84										
H85	*							*	*	*
H86										
H87	*							*	*	*
H88	*							*	*	*
H89	*							*	*	*
H90										
H91	*							*	*	*
H92	*							*	*	*
H93										
H94										
H95										
H96										
H97										
H98										
H99										
H100										
H101										
H102										
H103										
H104										
H105										
H106										
H107										
H108										
H109										
H110										
H111										
H112										
H113										
H114										
H115										

(Table continues next page.)

[a] See Table 1 for an explanation of the codes used in column headings.

Table A.1. Centriacinar Fibrosis Endpoints by Animal[a] *(continued)*

Animal	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H116										
H117		*	*	*	*					
H118		*	*	*	*					
H119		*	*	*	*					
H120		*	*	*	*					
H121		*	*	*	*					
H122		*	*	*	*					
H123		*	*	*	*					
H124		*	*	*	*					
H125		*	*	*	*					
H126		*	*	*	*					
H127		*	*	*	*					
H128		*	*	*	*					
H129		*	*		*					
H130		*	*	*	*					
H131		*	*	*	*					
H132		*	*	*	*					
H133	*					*	*	*	*	*
H134	*					*	*	*	*	*
H135	*									
H136	*					*	*	*	*	*
H137	*					*	*	*	*	*
H138	*					*	*	*	*	*
H139	*					*	*	*	*	*
H140	*					*	*	*	*	*
H141	*									
H142	*									
H143	*									
H144	*									
H145	*									
H146	*									
H147	*									
H148	*									
H149						*	*	*	*	*
H150						*	*	*	*	*
H151										
H152						*	*			
H153						*	*			
H154						*	*	*	*	*
H155								*	*	*

(Table continues next page.)

[a] See Table 1 for an explanation of the codes used in column headings.

Table A.1. Centriacinar Fibrosis Endpoints by Animal[a] *(continued)*

Animal	COLT	COLM	FV	IV	IMV	C6S	DS	RV	QCC	D_{CO}
H156								*	*	*
H157	*							*	*	*
H158	*									
H159	*							*	*	*
H160	*							*	*	*
H161	*							*	*	*
H162	*									
H163	*							*	*	*
H164	*									

[a] See Table 1 for an explanation of the codes used in column headings.

Table A.2. Airway Disease Endpoints by Animal[a]

Animal	MMEF	MUCT	MUCD	MUCR	ASMS	ASML	BRON	ED	CICL	NCCT	NCCD	NCCR	TENS	TENL	ECS	ECL	PGE_2	EV
H1	*																	
H2	*																	
H3	*																	
H4	*																	
H5	*																	
H6	*																	
H7	*				*	*							*	*	*	*	*	
H8	*				*	*							*	*	*	*	*	
H9	*				*	*							*	*	*	*	*	
H10	*																	
H11	*																	
H12	*																	
H13	*				*	*							*	*	*	*	*	
H14	*				*	*							*	*	*	*	*	
H15	*				*	*							*	*	*	*	*	
H16	*																	
H17	*																	
H18	*																	
H19	*				*	*							*	*	*	*	*	
H20	*																	
H21	*				*	*							*	*	*	*	*	
H22																		
H23	*				*	*							*	*	*	*	*	
H24	*				*	*							*	*		*	*	
H25																		

(Table continues next page.)

[a] See Table 2 for an explanation of the codes used in column headings.

Table A.2. Airway Disease Endpoints by Animal[a] *(continued)*

Animal	MMEF	MUCT	MUCD	MUCR	ASMS	ASML	BRON	ED	CICL	NCCT	NCCD	NCCR	TENS	TENL	ECS	ECL	PGE_2	EV
H26																		
H27																		
H28																		
H29																		
H30																		
H31																		
H32																		
H33																		
H34																		
H35																		
H36																		
H37		*	*	*			*	*	*	*	*	*						*
H38			*	*			*	*	*	*	*	*						*
H39		*	*	*			*	*	*	*	*	*						*
H40		*	*	*			*	*	*	*	*	*						*
H41		*	*	*			*	*	*	*	*	*						*
H42		*	*	*			*	*	*	*	*	*						*
H43		*	*	*			*	*	*	*	*	*						*
H44		*	*	*			*	*	*	*	*	*						*
H45		*	*	*			*	*	*	*	*	*						*
H46		*	*	*			*	*	*	*	*	*						*
H47		*	*	*			*	*	*	*	*	*						*
H48		*	*	*			*	*	*	*	*	*						*
H49		*	*	*			*	*	*	*	*	*						*
H50		*	*	*			*	*	*	*	*	*						*
H51		*	*	*			*	*	*	*	*	*						*
H52		*	*	*			*	*	*	*	*	*						*
H53	*																	
H54	*						*		*									
H55	*																	
H56	*																	
H57	*						*	*	*									*
H58	*						*		*									
H59	*																	
H60	*																	
H61	*						*		*									
H62	*						*		*									
H63	*																	
H64																		
H65	*						*		*									

(Table continues next page.)

[a] See Table 2 for an explanation of the codes used in column headings.

Table A.2. Airway Disease Endpoints by Animal[a] *(continued)*

Animal	MMEF	MUCT	MUCD	MUCR	ASMS	ASML	BRON	ED	CICL	NCCT	NCCD	NCCR	TENS	TENL	ECS	ECL	PGE_2	EV
H66	*						*	*	*									*
H67	*							*										*
H68	*																	
H69																		
H70																		
H71																		
H72																		
H73					*	*							*	*	*	*	*	
H74					*	*							*	*	*	*	*	
H75					*	*							*	*	*	*	*	
H76					*	*							*	*	*	*	*	
H77																		
H78																		
H79																		
H80																		
H81																		
H82																		
H83																		
H84																		
H85	*				*	*							*	*	*	*	*	
H86																		
H87	*				*	*							*	*	*	*	*	
H88	*				*	*							*	*	*	*	*	
H89	*				*	*							*	*	*	*	*	
H90																		
H91	*																	
H92	*																	
H93																		
H94																		
H95																		
H96																		
H97																		
H98																		
H99																		
H100					*	*							*	*	*	*	*	
H101																		
H102																		
H103																		
H104					*	*							*	*	*	*	*	
H105																		

(Table continues next page.)

[a] See Table 2 for an explanation of the codes used in column headings.

Table A.2. Airway Disease Endpoints by Animal[a] *(continued)*

Animal	MMEF	MUCT	MUCD	MUCR	ASMS	ASML	BRON	ED	CICL	NCCT	NCCD	NCCR	TENS	TENL	ECS	ECL	PGE_2	EV
H106																		
H107																		
H108																		
H109																		
H110																		
H111					*	*							*	*	*	*	*	
H112					*	*							*	*	*	*	*	
H113																		
H114																		
H115																		
H116																		
H117		*	*	*			*	*	*	*	*	*						*
H118		*	*	*			*	*	*	*	*	*						*
H119		*	*	*			*	*	*	*	*	*						*
H120		*	*	*			*	*	*	*	*	*						*
H121		*	*	*			*	*	*	*	*	*						*
H122		*	*	*			*	*	*	*	*	*						*
H123		*	*	*			*	*	*	*	*	*						*
H124		*	*	*			*		*	*	*	*						
H125		*	*	*			*	*	*	*	*	*						*
H126		*	*	*			*	*	*	*	*	*						*
H127		*	*	*			*	*	*	*	*	*						*
H128		*	*	*			*	*	*	*	*	*						*
H129		*		*			*		*	*	*	*						
H130		*	*	*			*	*	*	*	*	*						*
H131		*	*	*			*	*	*	*	*	*						*
H132		*	*	*			*	*	*	*	*	*						*
H133	*																	
H134	*																	
H135																		
H136	*																	
H137	*																	
H138	*																	
H139	*																	
H140	*																	
H141					*	*							*	*	*	*	*	
H142					*	*							*	*	*	*	*	
H143					*	*							*	*	*	*	*	
H144					*	*							*	*	*	*	*	
H145																		

(Table continues next page.)

[a] See Table 2 for an explanation of the codes used in column headings.

Table A.2. Airway Disease Endpoints by Animal[a] *(continued)*

Animal	MMEF	MUCT	MUCD	MUCR	ASMS	ASML	BRON	ED	CICL	NCCT	NCCD	NCCR	TENS	TENL	ECS	ECL	PGE_2	EV
H146																		
H147																		
H148																		
H149	*				*	*							*	*	*	*	*	
H150	*				*	*							*	*	*	*	*	
H151																		
H152					*	*							*	*	*	*	*	
H153																		
H154	*				*	*							*	*	*	*	*	
H155	*																	
H156	*																	
H157	*				*	*							*	*	*	*	*	
H158																		
H159	*				*	*							*	*	*	*	*	
H160	*				*	*							*	*	*	*	*	
H161	*				*	*							*	*	*	*	*	
H162					*	*							*	*	*	*	*	
H163	*																	
H164																		

[a] See Table 2 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a]

Animal	SMLW	MFLW	NSC
H1			
H2			
H3			
H4	*	*	
H5	*	*	
H6	*	*	
H7			
H8			
H9			
H10	*	*	
H11	*	*	
H12	*	*	
H13			
H14			
H15			

(Table continues next column.)

[a] See Table 3 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a] *(continued)*

Animal	SMLW	MFLW	NSC
H16	*		
H17	*	*	
H18	*		
H19	*	*	
H20			
H21	*	*	
H22			
H23	*		
H24	*	*	
H25			
H26			
H27			
H28	*	*	
H29	*	*	
H30	*	*	

(Table continues next page.)

[a] See Table 3 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a] *(continued)*

Animal	SMLW	MFLW	NSC
H31	*	*	
H32	*	*	
H33	*	*	
H34	*	*	
H35	*		
H36	*	*	
H37			*
H38			
H39			
H40			
H41			
H42			
H43			*
H44			
H45			
H46			
H47			
H48			
H49			
H50			
H51			
H52			
H53			
H54			
H55			
H56			
H57			
H58			
H59			
H61			
H62			
H63			
H64			*
H65			
H66			
H67			
H68			
H69			
H70			
H71			

(Table continues next column.)

[a] See Table 3 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a] *(continued)*

Animal	SMLW	MFLW	NSC
H72			
H73			
H74			
H75			
H76			
H77			
H78			
H79			
H80			
H81			
H82			
H83			
H84			
H85	*	*	
H86			
H87	*	*	
H88	*		
H89	*	*	
H90			
H91	*	*	
H92	*	*	
H93			
H94			
H95			
H96			
H97			
H98			
H99			
H100			
H101			
H102			
H103			
H104			
H105			
H106			
H107			
H108			
H109			
H110			

(Table continues next page.)

[a] See Table 3 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a] *(continued)*

Animal	SMLW	MFLW	NSC
H111	*	*	
H112	*	*	
H113	*	*	
H114			
H115	*	*	
H116	*	*	
H117			*
H118			*
H119			*
H120			*
H121			*
H122			*
H123			
H124			*
H125			*
H126			*
H127			*
H128			*
H129			*
H130			*
H131			*
H132			*
H133	*	*	
H134	*	*	
H135			
H136			
H137	*	*	
H138		*	
H139			
H140	*	*	
H141			
H142			
H143			
H144			
H145			
H146			
H147			
H148			
H149	*	*	
H150	*	*	

(Table continues next column.)

[a] See Table 3 for an explanation of the codes used in column headings.

Table A.3. Chronic Rhinitis Endpoints by Animal[a] *(continued)*

Animal	SMLW	MFLW	NSC
H151			
H152			
H153	*		
H154	*	*	
H155	*	*	
H156	*	*	
H157			
H158			
H159			
H160			
H161	*	*	
H162	*	*	
H163			
H164	*	*	

[a] See Table 3 for an explanation of the codes used in column headings.

ABOUT THE AUTHORS

Paul J. Catalano is an Assistant Professor of Biostatistics at the Harvard School of Public Health and the Dana-Farber Cancer Institute. He received a doctoral degree in biostatistics from the Harvard School of Public Health in 1991. His research interests focus on development and application of statistical methodology to environmental problems including multivariate modeling, analysis of animal bioassay data, and methods for risk assessment. Dr. Catalano is also involved in the statistical design and analysis of cancer clinical trials through the Eastern Cooperative Oncology Group.

John Rogus is a doctoral student in the Department of Biostatistics at the Harvard School of Public Health. He received a B.A. in mathematics and economics from Gettysburg College, Gettysburg, PA, in 1989.

Louise M. Ryan is an Associate Professor of Biostatistics at the Harvard School of Public Health and the Dana-Farber Cancer Institute. She received her Ph.D. in statistics from Harvard University in 1983. Her primary research interests include statistical methods for the design and analysis of carcinogenicity, developmental toxicity, and other toxicological experiments. She is particularly interested in statistical methods for multiple outcomes, especially in the context of dose-response modeling. Dr. Ryan also works on statistical analysis of cancer clinical trials, as well as epidemiological studies in fertility and human reproduction.

LIST OF ABBREVIATIONS

AB/PAS	Alcian blue/periodic acid–Schiff
ANOVA	analysis of variance
ASML	large airway smooth muscle
ASMS	small airway smooth muscle
BRON	percent bronchiolarization
C6S	chondroitin 6-sulfate
CICL	shift from ciliated to Clara cells
COLM	morphometric collagen volume
COLT	total lung collagen
D_{CO}	carbon monoxide diffusing capacity
DS	dermatan sulfate
EC_{50}	estimated concentration of stimulus that elicits a half-maximal response
ECL	large airway EC_{50} (responsiveness)
ECS	small airway EC_{50} (responsiveness)
ED	epithelial distance
EV	epithelial volume
FV	fibroblast volume
GAG	glycosaminoglycan
GEE	generalized estimating equation
IMV	interstitial macrophage volume
IV	interstitial volume
MAD	median absolute deviation
MFLW	nasal mucous flow rate, lateral wall
MMEF	mean midexpiratory flow
MUCD	caudal region stored AB/PAS-positive material
MUCR	cranial region stored AB/PAS-positive material
MUCT	central region stored AB/PAS-positive material
NCCD	caudal region, percentage of nonciliated cells
NCCR	cranial region, percentage of nonciliated cells
NCCT	central region, percentage of nonciliated cells
NSC	nasal secretory cells
NTP	National Toxicology Program
PGE_2	prostaglandin E_2 (mediator)
ppm	parts per million
QCC	quasistatic chord compliance
RV	residual volume
SD	standard deviation
SMLW	stored mucus, nasal lateral wall
T_{max}	maximal tension parameter
TENL	large airway maximal tension (responsiveness)
TENS	small airway maximal tension (responsiveness)

INTRODUCTION

Ozone is a highly reactive gas found in two regions of the atmosphere. In the stratosphere, miles above the earth's surface, it provides a protective shield by filtering out the sun's harmful ultraviolet radiation. In the troposphere, it is an ambient air pollutant to which people, agricultural crops, forests, and ecosystems are exposed (McKee 1993).

There are substantial uncertainties about the potential for prolonged exposure to ozone to contribute to or exacerbate chronic lung diseases (reviewed by Lippmann 1993; U.S. Environmental Protection Agency 1993). The widespread exposure of the population to this pollutant in many areas of the United States led the Health Effects Institute and the National Toxicology Program (NTP)* to enter into a collaboration to evaluate the effects of long-term exposure to low, medium, and high concentrations of ozone on laboratory animals.

The HEI component of the NTP/HEI Collaborative Ozone Project consisted of independent research studies designed to address the pathologic and physiologic consequences of prolonged exposure to ozone. Four groups of investigators measured changes in lung biochemical constituents, and two groups evaluated structural and cellular changes in the airways. Two types of functional measurements were made: in vivo pulmonary function measurements on anesthetized rats and in vitro stress and tension measurements on large and small airways isolated from rats. One team of investigators assessed nasal structure and function. In addition, Dr. Paul Catalano and colleagues developed an animal allocation scheme that allowed several investigators to measure endpoints on the same subset of study animals, assisted the individual investigators with the analyses of their data, and developed a statistical approach for analyzing multiple endpoints across the individual studies. Details of the NTP/HEI Collaborative Ozone Project can be found in the Introduction to this Report and in HEI Research Report Number 65, Part VI (Boorman et al. 1995). The results of the individual studies also have been published as Parts I through V and Parts VII through IX of Research Report Number 65.

This multiple-investigator project generated a large data base containing more than 240 measured endpoints. Although each study was designed to be interpreted as an individual entity, the Project provided a unique opportunity to look across the results of the individual studies and to evaluate the effects of ozone exposure on a broad spectrum of outcomes. Such an evaluation presented a problem because, in the field of toxicology, data reduction techniques have not been routinely applied to combine large numbers of related outcomes from independent studies. This report describes how Dr. Catalano and colleagues used the data collected in the NTP/HEI Collaborative Ozone Project to develop new procedures for combining and analyzing a complex data set containing multiple outcomes. The results of their analysis provide one way of viewing endpoints related to ozone exposure; they are not intended to replace the conclusions drawn from the individual studies, but to complement them.

During the course of the project and after the laboratory studies had been completed, the investigators, together with outside consultants, selected endpoints from their studies that were grouped together to represent three disease surrogates: centriacinar fibrosis (defined as fibrosis at the junction between the conducting airways and the alveolar gas-exchange area), airway disease (defined as conditions that contribute to a decreased ability of the conducting airways to transport air to and from the alveoli), and chronic rhinitis (defined as inflammation and irritation of the nasal passages). When the studies were completed, Dr. Catalano and colleagues applied a multivariate analysis to the grouped endpoints constituting each disease surrogate. The purpose of this Commentary is to evaluate this technique and compare it with other methods of multivariate analysis.

REGULATORY BACKGROUND

The U.S. Environmental Protection Agency (EPA) sets standards for oxidants (and other pollutants) under Section 202 of the Clean Air Act, as amended in 1990. Section 202(a)(1) directs the Administrator to "prescribe (and from time to time revise) . . . standards applicable to the emission of any air pollutant from any class or classes of new motor vehicles or new motor vehicle engines, which in his judgment cause, or contribute to, air pollution which may reasonably be anticipated to endanger public health or welfare." Sections 202(a), (b)(1), (g), and (h) and Section 207(c)(4)–(6) impose specific requirements for reductions in motor vehicle emissions of certain oxidants and other pollutants. In some cases, they provide the EPA with limited discretion to modify those requirements.

Section 109 of the Clean Air Act provides for the establishment of National Ambient Air Quality Standards (NAAQS) to protect the public health. Ozone's potentially harmful effects on respiratory function led the EPA to promulgate the NAAQS for ozone of 0.12 parts per million (ppm), a level not to be exceeded for more than one hour once per

* A list of abbreviations appears at the end of the Investigators' Report for your reference.

year. Section 181 of the Act classifies the 1989 nonattainment areas according to the degree that they exceed the NAAQS standards and assigns a date by which the primary standard must be attained for each classification.

The current ozone standard relies heavily on data derived from controlled human exposure studies that have demonstrated lung dysfunction following short-term exposure of exercising human subjects to ozone. Such studies do not address the potential for long-term ozone exposure to cause health effects. Determining the appropriate standards for emissions of oxidants and their precursors depends, in part, on an assessment of their risks to health. Research into the effects of long-term exposure of the airways to ozone, such as that supported by the NTP/HEI Collaborative Ozone Project, is therefore essential to the informed regulatory decision-making required by the Clean Air Act.

OBJECTIVE AND STUDY DESIGN

Because of the complexity of the NTP/HEI Collaborative Ozone Project, which generated a large data base containing multiple endpoints, a major objective of Dr. Catalano and colleagues was to develop a statistical framework for an integrative analysis of the results of the entire project. This required a multivariate analysis that could analyze jointly the endpoints from eight independent studies that used a total of 144 rats exposed to 0, 0.12, 0.50, and 1.0 ppm ozone for six hours per day, five days per week, for 20 months.

The investigators used a technique called median polish analysis that draws on the ideas of Tukey (1977). The technique produces composite variables, each of which is a function of a number of variables. The composite variables can be used to test for (1) trends across ozone exposure concentrations, and (2) the statistical significance of the paired comparisons between the rats exposed to the three concentrations of ozone and the control rats exposed to 0 ppm ozone for each of the three disease surrogates (centriacinar fibrosis, airway disease, and chronic rhinitis).

BACKGROUND TO THE STATISTICAL ANALYSIS

Reducing multiple outcomes to a few composite variables is a problem encountered frequently in statistical analysis. Methods that are often used to achieve this are principal components analysis and factor analysis (Cureton and D'Agostino 1983). These more general forms of methods examine the correlative structure of the original multiple outcomes and, in a two-step procedure, produce the composite variables. First, the correlations among the variables are examined and the number of dimensions of the data is determined. The number of dimensions is the number of composite variables that are needed to extract all the information or variance in the original set of variables. Second, given the number of dimensions, composite variables are generated that have substantive interpretations, which implies that the composite variables are a meaningful representative of a substantive field of study, e.g., centriacinar fibrosis. This is achieved by a series of mathematical transformations of the data that are continued until composite variables with the "best" substantive interpretation are produced. When these transformations have been completed, composite variables are produced that are functions of the original set of variables. The methods are successful if the number of composite variables is substantially smaller than the number of original variables, if they can be interpreted substantively, and if they are more precise and more reliable than the individual variables. These composite variables are then used as variables in other statistical analyses, such as regression analyses. These methods require a complete or near complete data set with at most only a few missing values.

MEDIAN POLISH ANALYSIS

Dr. Catalano and colleagues wanted to produce a few composite variables that could be used to analyze the effects of ozone on the endpoints examined in the individual studies. The composite variables the were used for trend analyses and paired comparisons.

The major reason that techniques such as principal components and factor analysis could not be applied in this project is that those methods do not easily accommodate the substantial number of data that were missing in the present studies, in which each investigator did not measure all variables on all rats. Also, those methods are sensitive to the influence of outliers, a potential problem in interpreting the data from this Project. The authors' stated goal was to develop a data summary that accommodated missing data points and was robust to the presence of outliers. Dr. Catalano and colleagues chose median polish analysis to generate the composite variables because they concluded that this technique satisfied these requirements. As applied in this study, the method resembles computations performed in a two-way analysis of variance more than it does principal components or factor analysis. For a given set of variables measured on subjects (rats), median polish produces a subject effect for each subject and a variable effect for each variable. The subject effects are used as the summary composite variable in subsequent statistical analyses. For each rat and each disease surrogate, the subject effect can be viewed as an estimate of the "average" for that rat

that combines the information from all the individual variables included in the analysis; it also adjusts for the fact that a particular rat was not used to measure each variable in the analysis.

Median polish analysis proceeds as follows: First, variables that relate to a substantive area are selected a priori. The investigators in the NTP/HEI Collaborative Ozone Project selected three major groupings of endpoints related to centriacinar fibrosis, airway disease, and chronic rhinitis (described in the Introduction). These endpoints are listed in Tables 1, 2, and 3 of the Investigators' Report. Second, for each disease surrogate, the variables are arranged in the same direction, positive or negative, to ensure that they do not cancel each other. In the present study, this was accomplished by changing the sign of variables that were expected to decrease with increasing ozone levels. For example, in the analysis of centriacinar fibrosis, the investigators anticipated that most variables would increase with increasing levels of ozone; however, residual volume was expected to decline with increasing ozone levels, so the sign was reversed for median polish analysis.

In the third step, the individual variables are standardized by subtracting the median from each observation and dividing by the median absolute deviation, which is defined as the median distance around the median. These data are then "polished." Polishing is an iterative smoothing of the data accomplished in part by continuously computing the medians for the subjects and variables; it is described in detail in the section of the Investigators' Report entitled "Robust Composite Scores Based on Median Polish Analysis." After the polishing was completed, the subject (rat) effects were obtained. Although they are far from being linear combinations, these scores can be considered to be composite variables because they were generated from all of the variables and from all of the rats on which the variables were measured. In the present study, two composite variables were based on data obtained from several investigators.

TECHNICAL EVALUATION

USE OF THE COMPOSITE VARIABLES

Dr. Catalano and colleagues used the composite variables as surrogates for the three disease surrogates (centriacinar fibrosis, airway disease, and chronic rhinitis) to analyze trends across the ozone concentrations and for paired comparisons of rats exposed to ozone and control rats exposed to 0 ppm ozone. Linear regression was used for the trend analysis, and *t* tests, without correction for multiple testing, were used for the paired comparisons. The investigators first performed these analyses using data from all rats on which endpoints for a particular disease surrogate had been tested. Then they repeated the analyses separately for male and female rats. Analyzing the residuals (the differences between the actual data and the median polish effects) for each disease surrogate indicated that data points from rats with leukemia may have been outliers; therefore, the analyses were repeated excluding all rats with leukemia and again repeated excluding only rats with advanced leukemia.

RESULTS

The major results obtained by median polish analysis are displayed in the Investigators' Report in Figures 5 through 7 for centriacinar fibrosis, Figures 18 through 20 for airway disease, and Figures 22 through 24 for chronic rhinitis.

The results for centriacinar fibrosis indicated a marginally significant trend toward an increased response associated with increasing ozone levels ($p = 0.091$) when the data for all rats were included, and a significant trend for females alone ($p = 0.038$). However, if the data for females are adjusted for multiple testing by Bonferroni's method, this result becomes at least marginally significant ($p = 0.076$) (the investigators did not include such an adjustment in their analyses). The results of pairwise comparisons of the rats exposed to ozone with control rats were not statistically significant.

Tables 12 and 13 in the Investigators' Report illustrate how the results for centriacinar fibrosis are affected by the form of the variable used and by the inclusion or exclusion of different variables. For example, the investigators planned to include collagen content as an endpoint in the composite score for fibrosis. However, the process by which collagen content is standardized affects the significance of the analysis. Deciding how to best standardize that variable required ongoing discussions among the investigators and Dr. Catalano and colleagues. The marginally significant or significant trends mentioned above were revealed when total collagen was expressed as collagen content per lung lobe; in contrast, if the collagen content had been expressed on the basis of lung weight, the results would not have been statistically significant. The analysis of centriacinar fibrosis also showed statistically significant changes, both in the trend analysis and in the pairwise comparisons of 0.5 and 1.0 ppm ozone exposures, when the sign for glycosaminoglycans was changed from that originally anticipated (an increase), and when the glycosaminoglycan data were eliminated from the analysis. The latter results are interesting, but, as the authors clearly state, they are post hoc and should not replace the original method of data analysis. The importance of the form of the variable used for the analyses and the inclusion or exclusion of variables have implications for the sensitivity of the method and are commented on below.

The results for airway disease showed statistically significant trends toward an increased response associated with increasing ozone levels ($p = 0.014$) for all rats and for male rats alone ($p = 0.033$). Again, adjustment for multiple testing reduced the trend in the males to marginal significance ($p = 0.066$). The response of male rats exposed to 0.12 ppm ozone was lower than that of males exposed to 0 ppm ozone. However, the sample size was considerably smaller in the group exposed to 0.12 ppm ozone, which weakens the statistical power of this pairwise analysis. In contrast, the responses of male rats exposed to 0.5 or 1.0 ppm ozone were slightly elevated, compared with those of control males. Although these pairwise comparisons did not achieve statistical significance, they were responsible for the statistical significance of the of the trend analysis. Because the response of male rats to 0.12 ppm ozone is statistically uncertain, the results of the trend analysis must be interpreted with caution and cannot be used to draw conclusions about the group exposed to 0.12 ppm ozone.

The results for chronic rhinitis showed a highly statistically significant trend toward an increased response associated with increasing ozone levels ($p < 0.0001$) for both male and female rats. Examination of the data indicates that they do not display a clear monotonically increasing trend. Rats exposed to 0.12 ppm ozone did not differ from control rats; however, the small sample size makes it difficult to interpret this pairwise comparison. The responses of the groups exposed to 0.5 or 1.0 ppm ozone were clearly greater than those of controls, and these differences were statistically significant. Thus, chronic rhinitis showed the most pronounced changes resulting from prolonged exposure to ozone.

Finally, when the analyses were repeated removing all rats with leukemia and again repeated removing rats with advanced leukemia, there were no changes in the authors' conclusions regarding ozone's effect on any of the three disease surrogates.

ISSUES CONCERNING MEDIAN POLISH ANALYSIS

Median polish has appeared in the literature for a number of years; however, it has not been applied to the problems of data analysis presented by the NTP/HEI Collaborative Ozone Project and, in this context, must be considered an experimental method. Its use here is an innovative application that was carefully applied by Dr. Catalano and colleagues, and it is sure to receive much attention in the statistical and health effects literature. However, it requires careful examination because it has several features that greatly distinguish it from methods such as principal components and other forms of factor analysis.

First, in median polish analysis all variables and their appropriate sign must be selected a priori, and all of the variables were assigned equal attention in determining the composite variable. Second, median polish analysis produces only one composite variable and provides no measure of how well it summarizes the individual variables. In principal components and other forms of factor analysis, variables that do not correlate well together will not be grouped together and given equal attention, even if only one composite variable is obtained. This reduces the importance of the a priori selections. In addition, if variables are not substantially correlated, these analyses will usually apply the correct sign to the composite variable. Finally, in principal components and other forms of factor analysis, characteristic values, called eigenvalues, indicate how well one composite variable explains the variance or correlations in the original variables, and correlation measures quantify how well each individual variable correlates to the composite measure. These measurements allow the analyst to judge whether one composite is a sufficient summary of the original variables.

The features of median polish and the potential problem of the method that may affect it need to be identified and addressed. Some of these issues are as follows:

1. Do summary composite variables produced for the different disease surrogates have face validity? Factors that enter into the determination of face validity for median polish and factor analysis include the appropriateness of the variables entered into the analysis, the number of investigators contributing to these variables, the extent of missing data, and the number of subjects (rats) studied. For example, an analysis that included data from 100 (of 144) rats produced by six of the seven principal investigators would have a powerful face validity as a representation and summary of the full study.
2. Does the method accurately estimate effects and produce summaries that are biologically meaningful?
3. How sensitive is the method to the inclusion or exclusion of variables?
4. How do the results obtained by this method compare with those obtained by other methods that could be applied to these data?

Face Validity

Tables 1 through 3 in the Investigators' Report list the variables selected for the determination of the composite variables for each of the three disease surrogates and identify the investigators that contributed data to each composite variable.

An issue in interpreting the results of the median polish analysis is the appropriateness of the variables selected for each disease surrogate. The endpoints selected by the investigators are generally appropriate for each disease surrogate as it was defined. For example, the 10 endpoints that entered into the summary of centriacinar fibrosis included 4 biochemical, 3 structural, and 3 functional measurements that are all "fibrosis-related" endpoints. However, although the 18 variables selected for the airway disease surrogate are, strictly speaking, related to "conditions that contribute to a decreased ability of the conducting airways to transport air to and from the alveoli," a large percentage of the endpoints for this disease surrogate are derived from morphometric studies. Clinical assessment of airway disease relies heavily on functional parameters; in this study, only one measurement of in vivo pulmonary function (mean midexpiratory flow) entered into the composite score. Because endpoints are given equal attention in median polish analysis, the functional measurements, which are generally considered to be critical in clinical diagnosis of airway disease, have less overall input in this analysis than the morphometric measurements.

Tables 7 through 9 list each of the 144 rats used in the studies and indicate which of the variables that entered into the composite variable for the three disease surrogates were measured on each rat. Tables 10 and 11 provide information for each disease surrogate concerning the number of rats that were studied, the number of variables (endpoints) measured per animal, and the number of individual studies per animal.

The information provided in these tables can be used to judge the face validity of the method. For example, in chronic rhinitis, only three variables were measured on 65 of the 144 rats and the data were generated from one study. Thus, this composite variable is not a composite over the entire collaborative Project; however, it is a composite over a single well-performed study that correlated the effects of ozone on nasal function and structure. In contrast, the data for airway disease are derived from 18 variables that were measured on 107 rats. Multiple variables were measured on many rats, and these data were generated by four of the seven investigators. Thus, this does have the spirit of being a summary over the entire project.

Biological Significance of the Composite Scores

The authors conducted simulations to evaluate whether their analyses produced scientifically reasonable summaries. The simulations were designed to demonstrate that median polish analysis detects the true biological effects even if there is noise in the data. Such noise could result from the presence of variables that have large variation, variables that have no association with ozone exposure, and large numbers of missing values. These were carefully addressed by the investigators in a series of studies discussed in the section of the Investigators' Report entitled "Evaluation of the Method via Simulation." These simulations were well conceived and executed. Their results indicate that the method has reasonable power to detect trends, even under difficult conditions. However, as with all simulation studies, one can never be certain that all aspects have been tested.

Sensitivity

The concern here is that this method, like others, may be sensitive to the inclusion or exclusion of a particular variable or variables. The investigators performed a series of analyses by removing one variable at a time and repeating the median polish on the reduced set of variables. For example, the results for airway disease (Table 15 in the Investigators' Report) indicated that, except for the pulmonary function endpoint, removing a single variable did not appreciably change the p values for statistical significance. However, this was not the case with the analysis for centriacinar fibrosis. Tables 12 and 13 summarize the data illustrating that the results were sensitive to the parameter by which collagen content was normalized, by the inclusion (as planned) of glycosaminoglycan changes, by post hoc exclusion of glycosaminoglycan changes, or by reversing the expected direction of change in glycosaminoglycans. Nonsignificant results became highly significant if the authors reversed the signs of the glycosaminoglycans chondroitin 6-sulfate and dermatan sulfate or excluded glycosaminoglycans entirely. These differing results indicate that the median polish is sensitive; however, they do not indicate which analyses are correct. A scientifically defensible option is to retain the variable that was decided upon a priori, by the investigators; this is what the authors chose to do.

Comparison with Other Methods

Another statistical approach that could have been applied to these data is the generalized estimating equation (GEE) procedure. This method also requires that the appropriate signs of the variables be "assigned" to the standardized variables so that all endpoints within a disease surrogate have the same expected direction with respect to ozone exposure. When the authors used this method to perform trend tests for the three disease surrogates, they obtained results that were similar to those obtained by median polish analysis, thus providing some assurance of the robustness of the results.

CONCLUSIONS AND DISCUSSION

The investigators produced a multivariate analysis that could analyze jointly the multiple outcomes from seven principal investigators who performed eight independent studies using a total of 144 rats exposed to ozone or clean air for 20 months. Median polish analysis produced three composite variables that were used to analyze the effect of ozone exposure on endpoints related to centriacinar fibrosis, airway disease, and chronic rhinitis.

The clearest ozone-induced changes were seen in chronic rhinitis. Statistically significant differences were seen by trend analysis and by pairwise comparisons of the rats exposed to 0.5 or 1.0 ppm ozone with control rats. Statistically significant or marginally significant changes were seen by trend analysis of the centriacinar fibrosis and airway disease surrogates; however, pairwise comparisons between control rats and rats exposed to ozone did not produce statistically significant differences.

Median polish analysis has not been used extensively in the present context; therefore, it must still be considered experimental. Three major concerns are associated with median polish analysis as applied here. First, the correct direction of variables had to be selected a priori. For example, the a priori assumption that glycosaminoglycans would increase in centriacinar fibrosis produced results that were statistically insignificant; however, post hoc analysis indicated that reversing the signs of the changes in individual glycosaminoglycans produced statistically significant results. Second, the results of the statistical analyses were sensitive to the set of variables included in the composite, as was seen when the glycosaminoglycan variable was removed from the analysis of centriacinar fibrosis. Third, the analysis gave equal attention to all variables. Thus, the composite variable for airway disease gives more attention to morphometric, rather than in vivo, pulmonary function variables.

Even with these caveats in mind, we can conclude that Dr. Catalano and colleagues successfully produced composite variables for some disease surrogates that appeared to have biological meaning and identified significant differences across ozone exposure levels.

The investigators provided evidence that median polish analysis maintained good power (1) in the presence of variables that had no relation with exposure, (2) in the presence of variables with different variance within the concentration-response patterns, and (3) when a substantial number of data points were missing. These are all strong points, and the method shows great promise. The authors' use of median polish analysis should make a substantial impact on the fields of toxicology and statistics.

ACKNOWLEDGMENTS

The Health Review Committee wishes to thank the ad hoc reviewer for help in evaluating the scientific merit of the Investigators' Report, and is grateful to Dr. Bernard Jacobson for assistance in preparing the Commentary. The Committee also acknowledges Virgi Hepner and Valerie Carr for overseeing the publication of this report, and Diane Foster and Mary L. Stilwell for their editorial and administrative support.

REFERENCES

Boorman GA, Catalano PC, Jacobson B, Kaden DA, Nauss KM, Ryan LM. 1995. Consequences of Prolonged Inhalation of Ozone on F344 Rats: Collaborative Studies, Part VI, Background and Study Design. Research Report Number 65. Health Effects Institute, Cambridge, MA.

Cureton EE, D'Agostino RB. 1983. Factor Analysis: An Applied Approach. Erlbaum Associates, Hillside, N.J.

Lippmann M. 1993. Health effects of tropospheric ozone: A review of recent research findings and their implications to ambient air quality standards. J Expos Anal Environ Epidemiol 3:103–129.

McKee DJ. 1993. Tropospheric ozone: What it is and how it is formed. In: Tropospheric Ozone: Human Health and Agricultural Impacts (McKee DJ, ed.) pp 3–7. Lewis Publishers, Boca Raton, FL.

Tukey JW. 1977. Exploratory Data Analysis. Addison-Wesley Publishing Co., Reading, MA.

U.S. Environmental Protection Agency. 1993. National Air Quality and Emissions Trends Report, 1992. Office of Air Quality Planning and Standards, Research Triangle Park, NC.

Report No.	Title	Principal Investigator	Publication Date
Toxicity and Carcinogenesis			
Research Reports			
1	Estimation of Risk of Glucose 6-Phosphate Dehydrogenase-Deficient Red Cells to Ozone and Nitrogen Dioxide	M. Amoruso	1985
6	Effect of Nitrogen Dioxide, Ozone, and Peroxyacetyl Nitrate on Metabolic and Pulmonary Function	D. M. Drechsler-Parks	
11	Effects of Ozone and Nitrogen Dioxide on Human Lung Proteinase Inhibitors	D.A. Johnson	1987
14	The Effects of Ozone and Nitrogen Dioxide on Lung Function in Healthy and Asthmatic Adolescents	J.Q. Koenig	1988
22	Detection of Paracrine Factors in Oxidant Lung Injury	A.K. Tanswell	1989
37	Oxidant Effects on Rat and Human Lung Proteinase Inhibitors	D.A. Johnson	1990
38	Synergistic Effects of Air Pollutants: Ozone Plus a Respirable Aerosol	J.A. Last	1991
44	Leukocyte-Mediated Epithelial Injury in Ozone-Exposed Rat Lung	K. Donaldson	1991
45	The Effects of Exercise on Dose and Dose Distribution of Inhaled Automotive Pollutants	M.T. Kleinman	1991
48	Effects of Ozone on Airway Epithelial Permeability and Ion Transport	P.A. Bromberg	1991
50	The Role of Ozone in Tracheal Cell Transformation	D.G. Thomassen	1992
54	Oxidant Injury to the Alveolar Epithelium: Biochemical and Pharmacologic Studies	B.A. Freeman	1993
60	Failure of Ozone and Nitrogen Dioxide to Enhance Lung Tumor Development in Hamsters	H. Witschi	1993
65	Consequences of Prolonged Inhalation of Ozone on F344/N Rats: Collaborative Studies		
	Part I: Content and Cross-Linking of Lung Collagen	J. Last	1994
	Part II: Mechanical Properties, Responses to Bronchoactive Stimuli, and Eicosanoid Release in Isolated Large and Small Airways	J.L. Szarek	1994
	Part III: Effects on Complex Carbohydrates of Lung Connective Tissue	B. Radhakrishnamurthy	1994
	Part IV: Effects on Expression of Extracellular Matrix Genes	W.C. Parks	1994
	Part V: Effects on Pulmonary Function	J.R. Harkema	1994
	Pat VII: Effects on the Nasal Mucociliary Apparatus	J.R. Harkema	1994
70	Oxidant and Acid Aerosol Exposure in Healthy Subjects and Subjects with Asthma		1994
	Part I: Effects of Oxidants, Combined with Sulfic or Nitric Acid, on the Pulmonary Function of Adolescents with Asthma	J.Q. Koenig	
	Part II: Effects of Sequential Sulfuric Acid and Ozone Exposures on the Pulmonary Function of Healthy Subjects and Subjects with Asthma	M.J. Utell	
HEI Communications			
1	New Methods of Ozone Toxicology: Abstracts of Six Pilot Studies	L-Y.L. Chang R.A. Floyd W.C. Parks K.E. Pinkerton D.A. Uchida R. Vincent	1992

(continued on next page)

Report No.	Title	Principal Investigator	Publication Date
Development of Methods to Measure Exposure			
Research Reports			
39	Noninvasive Determination of Respiratory Ozone Absorption: Development of a Fast-Responding Ozone Analyzer	J.S. Ultman	1991
63	Development of Samplers for Measuring Human Exposure to Ozone		1994
	Active and Passive Ozone Samplers Based on a Reaction with a Binary Reagent	J. Hackney	
	A Passive Ozone Sampler Based on a Reaction with Nitrate	P. Koutrakis	
	A Passive Ozone Sampler Based on a Reaction with Iodine	Y. Yanagisawa	
69	Noninvasive Determination of Respiratory Ozone Absorption: The Bolus-Response Method	J.S. Ultman	1994
HEI Communications Number 3: Environmental Epidemiology Planning Project			1994
•	Introduction to Working Group on Tropospheric Ozone	I.B. Tager	
•	Use of Human Lung Tissue for Studies of Structural Changes Associated with Chronic Ozone Exposure: Opportunities and Critical Issues	M. Lippmann	
•	Examining Acute Health Outcomes Due to Ozone Exposure and Their Subsequent Relationship to Chronic Disease Outcomes	B.D. Ostro	
•	Detection of Chronic Respiratory Bronchiolitis in Oxidant-Exposed Populations: Analogy to Tobacco Smoke Exposures	D. Bates	
•	The Role of Ozone Exposure in the Epidemiology of Asthma	J.R. Balmes	
•	Identification of Subpopulations That Are Sensitive to Ozone Exposure: Use of Endpoints Currently Available and Potential Use of Laboratory-Based Endpoints Under Development	R.B. Devlin	
•	Design and Analysis of Studies of the Health Effects of Ozone	A. Muñoz	
•	Summary of Papers and Research Recommendations of Working Group on Tropospheric Ozone	I.B. Tager	

Copies of these reports can be obtained by writing or calling the Health Effects Institute, 141 Portland Street, Suite 7300, Cambridge, MA 02139. Phone (617) 621-0266. FAX (617) 621-0267. Request a Publications and Documents booklet for a complete listing of publications resulting from HEI-sponsored research.

HEI

The Board of Directors

Archibald Cox *Chairman*
Carl M. Loeb University Professor (Emeritus), Harvard Law School

William O. Baker
Chairman (Emeritus), Bell Laboratories

Douglas Costle
Chairman of the Board and Distinguished Senior Fellow, Institute for Sustainable Communities

Donald Kennedy
President (Emeritus) and Bing Professor of Biological Sciences, Stanford University

Walter A. Rosenblith
Institute Professor (Emeritus), Massachusetts Institute of Technology

Health Research Committee

Bernard Goldstein *Chairman*
Director, Environmental and Occupational Health Sciences Institute

Joseph D. Brain
Chairman, Department of Environmental Health, and Cecil K. and Philip Drinker Professor of Environmental Physiology, Harvard University School of Public Health

Glen R. Cass
Professor of Environmental Engineering and Mechanical Engineering, California Institute of Technology

Seymour J. Garte
Professor and Deputy Director, Department of Environmental Medicine, New York University Medical Center

Leon Gordis
Professor and Chairman, Department of Epidemiology, Johns Hopkins University, School of Hygiene and Public Health

Stephen S. Hecht
Director of Research, American Health Foundation

Meryl H. Karol
Professor of Environmental and Occupational Health, University of Pittsburgh, Graduate School of Public Health

Robert F. Sawyer
Class of 1935 Professor of Energy (Emeritus), University of California at Berkeley

Gerald van Belle
Chairman, Department of Environmental Health, School of Public Health and Community Medicine, University of Washington

Health Review Committee

Arthur Upton *Chairman*
Clinical Professor of Pathology, University of New Mexico School of Medicine

A. Sonia Buist
Professor of Medicine and Physiology, Oregon Health Sciences University

Ralph D'Agostino
Professor of Mathematics/Statistics and Public Health, Boston University

Gareth M. Green
Associate Dean for Education, Harvard School of Public Health

Donald J. Reed
Professor and Director, Environmental Health Sciences Center, Oregon State University

David J. Riley
Professor of Medicine, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School

Herbert Rosenkranz
Chairman, Department of Environmental and Occupational Health, Graduate School of Public Health, University of Pittsburgh

Robert M. Senior
Dorothy R. and Hubert C. Moog Professor of Pulmonary Diseases in Medicine, Washington University School of Medicine

Frederick A. Beland *Special Consultant to the Committee*
Director for the Division of Biochemical Toxicology, National Center for Toxicological Research

Henry A. Feldman *Special Consultant to the Committee*
Senior Research Scientist, New England Research Institute

Edo D. Pellizzari *Special Consultant to the Committee*
Vice President for Analytical and Chemical Sciences, Research Triangle Institute

Officers and Staff

Daniel S. Greenbaum *President*
Richard M. Cooper *Corporate Secretary*
Kathleen M. Nauss *Director for Scientific Review and Evaluation*
Elizabeth J. Story *Director of Finance and Administration*
Jane Warren *Director of Research*
Maria G. Costantini *Senior Staff Scientist*
Chester A. Bisbee *Staff Scientist*
Aaron J. Cohen *Staff Scientist*
Bernard Jacobson *Staff Scientist*
Debra A. Kaden *Staff Scientist*
Martha E. Richmond *Staff Scientist*
Lata Shirnamé-Moré *Staff Scientist*

Gail V. Allosso *Senior Administrative Assistant*
Valerie Anne Carr *Publications Production Coordinator*
Diane C. Foster *Editorial Assistant*
L. Virgi Hepner *Managing Editor*
Teresina McGuire *Accounting Assistant*
Jean C. Murphy *Research Associate*
Jacqueline C. Rutledge *Controller*
Malti Sharma *Publications Assistant*
Mary L. Stilwell *Administrative Assistant*
Stacy Synan *Administrative Assistant*
Susan J. Walsh *Receptionist*

HEI HEALTH EFFECTS INSTITUTE

141 Portland Street, Cambridge, MA 02139 [617] 621-0266

Research Report Number 65

January 1995