UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
Environmental Criteria and Assessment Office (MD-52)
Research Triangle Park, North Carolina 27711

June 15, 1983

Dear Chuck

Here is the final report of the NHANES II TIME TREND ANALYSIS REVIEW GROUP as it will appear in the external review draft of the Lead Criteria Document. This is provided for your information only and is not modifiable by comment, suggestion, or discussion. You are encouraged, however, to comment on this appendix and the entire Lead Criteria Document during the 90-day public comment period which will follow release of the external review draft (still anticipated to be in July).

Sincerely,

[signature]

David E. Weil, Ph.D.
Project Manager

Enclosure

+ [illegible]

APPENDIX 11-D

REPORT

OF THE

NHANES II TIME TREND ANALYSIS REVIEW GROUP

June 15, 1983

UNITED STATES DEPARTMENT OF COMMERCE
National Bureau of Standards
Washington, D.C. 20234

June 14, 1983

Dr. Lester D. Grant
Director, Environmental Criteria
and Assessment Office (MD-52)
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Dear Dr. Grant:

I am enclosing the report of the NHANES II Time Trend Analysis Review Group. All of the members of the group have concurred in this report, and we believe it responds to the set of questions addressed to us in your letter of February 17, 1983.

Sincerely,

Joan R Rosenblatt

Joan R. Rosenblatt
Chairman, Review Group

Enclosure

cc: Review Group Members

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
Environmental Criteria and Assessment Office (MD-52)
Research Triangle Park, North Carolina 27711

The materials contained in this report were generated as the result of critical evaluations and deliberations by members (listed below) of the NHANES II Time Trend Analysis Review Group. We, the undersigned members of this Review Group, concur with and endorse the findings and recommendations contained in the present report as representing the collective sense of the Review Group. We also confirm, by our signatures below, that we are neither affiliated with nor receiving or expect to receive any form of financial or other support from any lead industry firm (or research organization) or any other industry-related commercial or research entity likely to be affected by standard setting or other regulatory activities potentially influenced by the present findings and recommendations of this Review Group.

Joan R Rosenblatt

Dr. Joan Rosenblatt (Chairman)
Deputy Director
Center for Applied Mathematics
National Bureau of Standards
Washington, D.C. 20234

Harry Smith, Jr.

Dr. Harry Smith, Professor
Chairman, Department of
Biomathematical Science
Mt. Sinai School of Medicine
New York, New York 10029

Richard Royall (JKR)

Dr. Richard Royall, Professor
Department of Biostatistics
Johns Hopkins University
615 North Wolfe Street
Baltimore, Maryland 21205

R Little

Dr. Roderick Little
American Statical Assoc. Fellow
Bureau of Census
Department of Commerce
Washington, D.C.

J. Richard Landis (J.R.L.)

Dr. J. Richard Landis, Professor
Department of Biostatistics
School of Public Health II
University of Michigan
Ann Arbor, Michigan 18109

Table of Contents

Summary

The Review Group finds strong evidence that there was a substantial decline in the average level of blood lead in the U.S. population during the NHANES II survey period. After adjustment for relevant demographic covariables, the magnitude of the change can be estimated for the total U.S. population and for some major subgroups, provided careful attention is given to underlying model assumptions.

The Review Group also finds a strong correlation between gasoline-lead usage and blood-lead levels. In the absence of scientifically plausible alternative explanations, the hypothesis that gasoline lead is an important causal factor for blood-lead levels must receive serious consideration. Nevertheless, despite the strong association between the decline in gasoline-lead usage and the decline in blood-lead levels, the survey results and statistical analyses do not confirm the causal hypothesis. Rather, this finding is based on the qualitatively consistent results of extensive analyses done in different but complementary ways.

The gasoline lead coefficient in regressions of blood-lead levels on that variable, adjusted for observed covariates, has been used to quantify the causal effect of gasoline lead on blood-lead levels. The Review Group considers that such inferences require strong assumptions about the absence of effects from other unmeasured lead sources, the adequacy of national gasoline lead usage as a proxy for local exposure, and the adequacy of a sample design which does not measure changes in blood-lead levels for individuals in the sample. The validity of these assumptions could not be determined from the NHANES II data or from other data supplied to the Review Group. Furthermore, the Review Group cautions against extrapolation of the observed relationship beyond the limits of the four year period.

Introduction

This Review Group was appointed in February, 1983 by the Director of the Environmental Criteria and Assessment Office, U.S. Environmental Protection Agency (EPA), to consider a series of questions about the interpretation of data from the second National Health and Nutrition Examination Survey (NHANES II) to evaluate relationships over time between blood-lead levels and gasoline lead usage. The questions addressed to the Review Group are listed in full in Appendix D1.

Documents describing NHANES II, analyses of the survey data, and analyses of the relationships between blood-lead values and gasoline lead usage were furnished for review. In two meetings, on March 10-11 and March 30-31, 1983, the Review Group discussed these materials with officials of the EPA, and with specialists from the several institutions that had conducted these studies. The documents provided for review are listed in Appendix D2. The individuals who attended the two meetings are listed in Appendix D3.

The panel members of the Review Group are statisticians with experience in applications of statistics in the physical, biomedical, and social sciences, but had no previous involvement in analyses of data about blood lead or gasoline lead. The affiliations of the panel members are listed in Appendix D3 for identification; views expressed by the panel in this report are their own and not those of the institutions.

Agencies involved in the conduct of the NHANES II were the National Center for Health Statistics (NCHS), the Centers for Disease Control (CDC) where the chemical analyses were done, and the Food and Drug Administration (FDA).

Contributors to the analysis of the association between blood lead and gasoline lead usage, in addition to NCHS and CDC, are E. I. DuPont de Nemours & Co. (DuPont), The Ethyl Corporation (Ethyl), and the EPA Office of Policy Analysis working in collaboration with ICF Incorporated (ICF) and Energy and Resource Consultants, Inc. (ERC).

This report contains two major sections. The first, on time trends in blood-lead levels, addresses a set of questions about the use of NHANES II data to estimate changes over time. The second addresses statistical aspects of evaluating the relationship of changes in blood-lead levels to gasoline lead usage.

Time Trends In Blood-lead Values

At its first meeting on March 10-11, 1983, the Review Group considered only the first of the set of questions presented to it (see Appendix D1), namely questions about the extent to which the NHANES II data could be used to "determine time trends for changes in nationally representative blood-lead values for the years of the study (1976-1980)."

The phrases "define time trends" and "determine time trends ... (1976-1980)" are interpreted throughout this report to mean "estimate changes in blood-lead values during the survey period." In particular, such changes are not to be interpreted as trends that might be extrapolated.

The Group recognized that the survey was designed as a cross-sectional survey, and specifically inquired into three general kinds of possible sources by time-related bias:

- the measurement quality control,
- the nonresponse experience, and
- the survey design.

As would be expected, only incomplete evidence could be made available in each of these areas. The following assessment of this evidence indicates where it depends on the expert opinion of others.

Measurement Quality Control

In order to analyze the time trends in NHANES II data, one must assume that the procedures for collecting, handling, and analyzing blood specimens did not change during the survey years. The Review Group is aware that contamination can produce spuriously high values in determination of trace elements, and sought evidence that quality control procedures were equally stringent at all times.

Although no quality control specimens were prepared at the medical examination sites, the Review Group has been assured that training, periodic retraining, materials, equipment, and procedures were designed to prevent contamination, and not changed. There was some turnover of personnel.

The CDC laboratory established and documented the results of extensive quality control sampling (App. D2, item 14). The data of lead levels in the "blind" samples, from two pools of bovine blood, exhibit essentially constant means and standard deviations. The coefficient of variation for measurement error was found to be about 17 percent for blood-lead levels near 13 µg/dL; it was smaller, about 13 percent, for higher blood-lead levels near 25 µg/dl. Additional evidence of the constancy of quality control is that data from other analyses of the blood specimens (zinc, for example) exhibit little or no change over time.

The Review Group finds no evidence that field and laboratory quality control changes could account for the observed change in blood-lead levels.

Nonresponse

Nonresponse is an important potential source of bias in sample surveys. It is of particular concern in the blood-lead analysis of the NHANES II since the nonresponse rate is high--39.3 percent of sampled persons had missing lead values due to nonresponse at various stages of participation in the survey (App. D2, item 14, p.9). The NCHS attempted to adjust for nonresponse by weighting responding individuals by estimates of the probability of response, calculated within subclasses of the population formed by joint levels of age, income, SMSA/non-SMSA, and region.

This is a standard adjustment method for unit nonresponse in surveys. The method adjusts for differential nonresponse across the subclasses used to calculate the weight, but does not account for residual association between nonresponse and time and blood-lead level, which are the variables of primary interest in the analysis under consideration. Thus there is the possibility that nonresponse bias is a contributory factor to the trend in blood-lead levels across time.

In order for nonresponse to have this effect it is necessary that, after adjusting for the socioeconomic variables used to define the weights, nonresponse be related to blood-lead level, and further that this relationship change over time, so that a differential bias in the mean blood-levels of respondents exists across time. Clearly this question cannot be addressed directly, since the blood-lead levels of nonrespondents are not measured. However, the Review Group considered such an interaction to be highly unlikely, for the following reasons:

- Nonresponse rates did not vary in a consistent way across time. Examination of changes in response rates does not indicate any relationship of importance (App. D2, item 18).

- There does not appear to be evidence that the conditions of the survey changed significantly across time, so that any bias introduced by an association between nonresponse and blood-lead level is unlikely to change across time.

Accordingly, the Review Group rejected nonresponse as a likely explanation for the trend observed in the data.

Survey Design

The NHANES II was designed to provide U.S. national prevalence rates for a wide range of characteristics and health conditions. Due to financial and logistical constraints, the survey design required a four-year data collection period. Consequently, the sample quantities, such as the blood-lead levels, necessarily will provide period prevalence estimators, rather than point prevalence estimators of the underlying population parameters. In general practice, a fundamental assumption underlying the use of period data to generate prevalence estimators is that the condition under investigation remains relatively constant throughout the survey period.

Even though the NHANES II was not designed to detect and estimate changes in prevalence throughout the survey period, one must consider the possibility that the level of a particular target characteristic, such as blood lead, actually may be changing over time. Consequently, one cannot ignore evidence suggesting that the level of lead in blood in the U.S. population was decreasing during the data collection period simply because the survey design was cross-sectional, rather than longitudinal. Rather, the difficult question is to what extent, if any, can these NHANES II data be used to determine time trends.

Although a cross-sectional design such as the one utilized in the NHANES II certainly is not optimal for investigating time trends, one can consider making adjustments within the sample for the effects of relevant covariables such as age, sex, race, residence, and income, if the distributions of these covariables are not highly confounded with time. An additional requirement for making adjustments is that there be reasonably large numbers of sample persons for different covariable levels at various times. These internal adjustments permit one to examine whether the decline in blood-lead levels can be accounted for by differing proportions of individuals from subgroups determined by relevant covariables. The extent of this type of selection bias over time relative to primary demographic characteristics can be summarized (App. D2, item 20, Tables M7, M8 for whites, and M13, M14 for blacks).

The Review Group considered carefully the potential bias due to changing composition of the sample over time, especially since this had been emphasized by Ethyl (App. D2, items 25, 26). The most striking problem occurs with urban vs. rural groups. The fractions of blood samples obtained from white urban residents are shown as follows:

	% urban bloods	Sample size
Jan - Jun 1976	64.2	795
Jul - Dec 1976	36.9	1255
Jan - Jun 1977	44.6	935
Jul - Dec 1977	57.3	1010
Jan - Jun 1978	46.3	1056
Jul - Dec 1978	40.6	981
Jan - Jun 1979	31.6	1228
Jul - Dec 1979	20.7	842
Jan 1980	0.0	267

Thus, there has been a striking decrease in the number of bloods taken from white urbanites across the four years. If one assumes that exposure to lead from gasoline is more prevalent in urban areas, then (without adjustment) the observed mean blood levels across the four years would be biased because of the NHANES II schedule.

Further examination of the CDC tabulation (App. D2, item 20) indicates sparse information on blacks. The numbers are so small that time trend inferences for blacks can be estimated with confidence only for overall mean blood-lead level results without regard to sex, place of residence, and age.

The Review Group finds that despite obvious trends over time for such characteristics as degree of urbanization and the proportion of children aged 0.5 to 5 years, the sample size is distributed across the grid of covariable levels sufficiently to permit reasonable adjustments. In support of this finding, the Review Group notes that similar trends appeared whenever demographic subgroups were examined separately. These subgroups included white males, white females, white children, white teenagers, white adults, and blacks, as well as breakdowns by income and urban-rural status.

Sample Weights

Another possibility is that the sample mean blood-lead level changes resulted from trends in more subtle statistical characteristics of the sample over time, such as characteristics related to the way sample weights are used to calculate averages. But this explanation appears to be inconsistent with the fact that analyses of the unweighted NHANES II data lead to essentially the same results as the weighted data and analysis.

In response to questions raised by both industry representatives and other observers, the Review Group explored the effects of the complex weighting scheme inherent in all the CDC and EPA/ICF analyses. Each sample observation has both a basic weight (related to the probability of selection), a final weight (reflecting additional adjustments to the basic weight accounting for nonresponse patterns of selected demographic subgroups), and a final examined lead subsample weight (corresponding to the entire set of adjustments due to the probability of selection, nonresponse, and post-stratification, and the subsampling of individuals selected for the measurement of blood lead). All the weighted analyses in the CDC and EPA/ICF reports were conducted relative to the final examined lead subsample weight.

One potential problem associated with this final lead subsample weight is the possibility that differential nonresponse patterns for various demographic subgroups may lead to marked differences between the basic weight (without nonresponse adjustments) and this final weight. For that reason, the Review Group requested a data display of the total nonresponse rate and the average blood-lead levels by the 64 separate stands using three different weighting schemes in computing the averages:

i) unweighted;
ii) basic weights;
iii) final lead subsampling weights.

As shown in Table 1, item 18 of App. D2, the average blood-lead levels are quite consistent under each weighting scheme for each of the 64 stands. Furthermore, there is no apparent trend in the nonresponse rate across time. Consequently, one would expect that an analysis of these data under the basic weights also would parallel the results obtained in the CDC and the ICF reports.

These findings, in conjunction with the similarities between the weighted and unweighted analyses, lend additional support to the overall consensus among panel members that these data analyses are not dependent on the particular choice of weights, including the intermediate basic weights.

Estimated Time Trends

There seems to be no doubt that, qualitatively, a downward trend of blood-lead levels has been observed during the NHANES II survey.

The data appear to support reasonably precise estimates of the magnitude of the change for a few major subgroups of the population. In particular, the change in mean blood-lead levels during the survey period can be estimated for the population as a whole and for population sectors grouped by age, sex, race, urban/rural, and income, if each of these demographic categories is considered separately.

For estimating changes in mean blood-lead levels for combinations of demographic factors, sufficient data appeared to be available for white-by-sex and white-by-age breakdowns. These estimated changes, and others that might be considered, can be made on the basis of a linear model that provides adjustments for demographic and socioeconomic covariables that are known or believed to be associated with blood-lead levels.

For finer subdivisions, estimates of change are subject to large sampling error and are sensitive to correct specification of the regression model. Hence, caution must be exercised in their interpretation. It is not possible to show time changes in mean blood levels for specific cities, towns, or locales using the NHANES II data, since no city or locale was sampled more than once. No data which would allow estimates of time trends in mean blood-lead levels for different occupational categories were shown to the Review Group. The only socioeconomic variable considered was income.

Estimates of change, e.g., those reported by CDC (App. D2, item 14, Table 6, page 44), should be accompanied by standard errors. There should be discussions of the use of regression diagnostics to evaluate the adequacy of the model, and the possibility that a few observations exert an excessive influence on the result. The calculation of standard errors should use procedures that take into account the stratification and clustering properties of the survy design. In response to the Review Group's questions, CDC provided a document presenting standard errors and the methodology used to estimate them (App. D2, item 38). The size of these standard errors suggests that there are only weak indications of differences between subgroups with respect to the percent drop in the average blood-lead level.

Summary

Although the survey was not sepcifically designed to measure trends, data from the NHANES II can be used to estimate changes in blood-lead levels during the four-year period, 1976-1980, of the survey. Changes can be estimated for the U.S. population and for major population subgroups, as specified in the previous subsection. Because of sampling error, laboratory measurement error, a high nonresponse rate, and the need to adjust for time-related imbalance in the survey design, such estimated changes should be interpreted with caution.

Correlation Between Blood-Lead and Gasoline-Lead Changes

At its second meeting on March 30-31, 1983, the Review Group considered three sets of studies that examine the association between changes in blood-lead levels estimated from the NHANES II data and changes in the use of leaded gasoline:

- the Ethyl Corp. analysis (App. D2, items 25, 26)
- the ICF/EPA analysis (App. D2, items 11, 22, 23, 24), and
- the CDC/NCHS analysis (App. D2, item 14 and appendices).

The following discussions summarize the Review Group's assessment of the strengths and weaknesses of the analyses.

Preliminary Remarks

The analyses propose and evaluate models for the relationship between blood-lead levels and gasoline-lead usage. All of these analyses rely on multiple linear regression methods, whose limitations with respect to establishing causal relations are well known (See, e.g., reference 1). The statistician-reviewer may adopt one or the other of two approaches in considering the strengths and weaknesses of the several analyses:

(1) Assume (on external authority) the existence of a causal relationship between gasoline lead usage and blood lead levels. Consider the variables and models used to analyze the strength of the association and to estimate the effect of gasoline-lead changes on blood-lead changes. In this approach, the possible effects of other changes over time that affect blood-lead levels are treated as second-order effects. CDC urges this approach.

(2) Adopt a neutral position as to the causal relationships, and examine the associations among the variables studied. In this approach, "time" serves as a proxy for the combined effect of whatever changes affected blood-lead levels and it is left to the interpreter of the analyses to assign relative importance among suggested explanations for changes over time. DuPont and Ethyl suggest this approach.

The ICF and CDC analyses both found a clear relationship between gasoline lead and blood lead. The Ethyl analysis found no evidence of association between these variables. The purpose of this commentary is to discuss the important differences between the analyses and to assess their utility in establishing or contradicting the hypothesized relationship between the decline in blood-lead levels and the decline in gasoline lead emissions over the period of the NHANES II Survey.

Table I (next page) classifies the three analyses by six factors which capture the main differences between them, namely: 1) the choice of measure of gasoline lead, 2) the scale of blood lead variable, raw or logarithm, 3) the unit of analysis, 4) control variables in the regression, and in particular

the inclusion or omission of a time variable, 5) the weighting used in the regressions, and 6) the method used to calculate standard errors. The panel concludes that of these factors only (1) and (4) had a substantial impact on the final results.

Table 1

	CDC	ICF	Ethyl
1) measure of gasoline lead	quarterly	monthly sales x lead conc.	pop. density local lead usage
2) scale of dependent variable	log	raw	raw
3) unit of analysis	individual	individual	individual stage 1 locality stage 2
4) control variables include time	no	time, season, lagged gas	time
5) weighting by selection probs.	both	yes	no
6) design based standard errors	yes	yes	no

The first three factors are discussed under the heading "Variables Used in the Analyses". Factors (4), (5), and (6) are discussed under "Statistical Techniques Used in the Analyses". Factor (4) is considered further in the assessment of "Models Used in the Analyses".

Variables Used in the Analyses

Demographic and socioeconomic covariables were used as defined for the NHANES II Survey. Differences between the analyses occurred in the choice of specific representations for blood-lead levels and gasoline lead usage.

Blood Lead. All the studies used blood-lead values for individuals from the NHANES II Public Use Data Tape, with associated demographic, economic, time, and sampling-weights data.

Ethyl calculated adjusted blood-lead values for its principal analysis by fitting a linear model to adjust for age, sex, race, and income to obtain the residuals from this analysis. Ethyl did not adjust the individual data for the effect of the degree of urbanization, a factor recognized to be related to blood-lead levels. Averages of the adjusted values for 55 of the 64 examination sites were used in the principal (second-stage) analysis.

ICF used the NHANES II blood leads without adjustment or transformation. Adjustment for socio-demographic variables was achieved by including these variables as covariates in regression models for individual blood leads.

CDC adopted a similar approach, but used the natural logarithms of the NHANES II blood leads, on the basis of an analysis showing that the distribution of the values themselves was skewed and that the transformation successfully corrected for the skewness.

The scale of the dependent variable (raw or logarithm) does not appear to have a great influence on the final results. With the exception of race, the blood-lead/gasoline lead slope in the CDC and ICF analyses appeared stable across demographic factors, whether the raw or log scale was used for the dependent variable. The logarithm scale has the advantage of being more likely to yield normal residuals.

The unit of analysis (factor 3) received a considerable amount of discussion by reviewers. In particular, the Ethyl two-stage analysis was subjected to some criticism. At the first stage, the blood lead variable was adjusted for differences in the distributions of demographic variables by an indiviudal level regression on NHANES II data. At the second stage, the adjusted locality mean blood-lead values were regressed on proxies for gasoline lead which had not themselves been adjusted for the demographic variables. This two-step regression procedure leads to bias (see reference 2), but the bias does not appear important, as Ethyl later corrected the analysis with no substantial change in the results.

Gasoline Lead Usage/Exposure. There were several different approaches to defining variables that could be interpreted as indexes of the amount of lead present in the environment at the time when blood samples were taken, as well as during the antecedent months. Clearly, no index number or set of index numbers can serve as an ideal surrogate for a measurement of the exposure experiences of sampled persons. The Review Group recognizes the complexity of the mixture of lead sources and uptake pathways.

The large differences between the results of the ICF/CDC analyses and the Ethyl analysis are caused by different measures of gasoline lead exposure. ICF and CDC used national period measures-quarterly EPA lead additive data for CDC and adjusted monthly gasoline sales data for ICF, whereas Ethyl used two proxy measures for lead exposure at each locality--population density and lead use per unit area.

A fundamental assumption underlying the creation of a local estimate of gasoline lead exposure is the notion that the volume of leaded gasoline consumed locally, with the resulting "fallout", is the primary source of lead in human blood. Although this determination requires substantive expertise beyond that on our Review Group, the choice of a local vs. a global measure of exposure is a pivotal one in all these analyses. If, in fact, lead enters the human blood system via imported fallout through the food chain (and other sources), as well as the inhalation of local "fallout", then ideally one would require a summary measure of exposure which captures both of these sources.

CDC used data from the quarterly EPA Lead Additive Reports (App. D2, item 14, pages 37-40 and Appendix H). These are national values of the total amount (by weight) of lead used in gasoline production. The series exhibits seasonal fluctuations in gasoline production in addition to a general downward trend.

ICF developed a monthly series of national values of the average amount (by weight) per day of lead used in gasoline, as follows: Monthly average gasoline use (liquid volume per day) was obtained from the DOE Monthly Energy Review. Quarterly values of the concentration of lead in gasoline (grams per gallon, based on refiner reports) were obtained from EPA (App. D2, item 11). The product of these produced a monthly series. This series, if aggregated to a quarterly series, would be closely related to the series used by CDC.

The measures of lead use used by CDC and ICF capture the downward trend in gasoline lead over time, but they suffer from specification error in that they are national rather than localized measures of gasoline lead exposure. The defect has two consequences:

(a.) The gasoline lead use variable does not capture variation in gasoline lead exposure between localities.

(b.) The lead use variable can be only partially adjusted for correlations with the demographic covariates.

The CDC analysis partially corrects for (a) by aggregating the gasoline lead exposure over all sampled localites in a six month period of sampling. The second problem remains, however. The panel does not believe that these deficiencies invalidate the qualitative findings of a relationship between lead usage and blood lead. However, the impact on the coefficient of lead usage in the CDC analysis is not clear.

Ethyl adopted a different approach, seeking to represent gasoline-lead usage at the survey locations and also to consider separately the effects of lead in air and lead fallout. The variables used to represent the two kinds of lead exposure were, respectively, population density and gasoline lead usage per square mile for the sampled localities.

The Review Group applauded the intention of the Ethly effort, but the variables selected appear to be inappropriate. In the Ethyl discussion (App. D2, item 26, Appendix page A-3) it is pointed out that population density is strongly related to degree of urbanization, a factor for which adjustment is made in the CDC and ICF analyses, but not in the Ethyl analysis. Furthermore, Ethyl calculated population density by interpolation between censuses and it is doubtful that it would reflect changes (if any) in the concentration of lead in air within the four-year survey period.

Ethyl represented lead usage per unit area by annual values by state. Department of Transportation reports of annual gasoline sales (by state) and annual Ethyl estimates of the amount of lead in gasoline being sold (by state) produced state estimates of annual totals of lead used. These were then divided by the area of the state. Examination of the resulting values (App. D2, item 26, Table 6, page 23) reveals anomalies. For example, the 1979 lead usage value for Washington, DC, is 5 times larger than that for any other

location. The second-largest value is the one for New Jersey in 1977, used for locations adjacent to New York City; it is more than 4 times the 1977 value used for both New York City and its Westchester County suburbs. As another example, the computed exposure for Houston, TX (ID no. 28) is 101, compared to 7174 for Washington, DC (ID no. 33). The naive implication of these two data points is that persons living in Washington, DC received a 71-fold (7174/101) increase in dosage of air-lead (or food chain lead) compared to persons living in Houston, TX. Whether we view this dosage as exposure through air or food, this extreme differential is highly unlikely. This variable appears to represent chiefly the statewide average population density. The Review Group cannot accept it as an indicator of gasoline lead usage at the sample locations.

Statistical Techniques Used in the Analyses

All final models reported by EPA/ICF and CDC were fitted to the NHANES II data using the SURREGR procedure available in SAS. This computing software permits sample weights and cluster design effects to be incorporated into the variance-covariance estimators of the model parameters. Although unweighted and weighted ordinary least squares model fitting provided the same conclusions, SURREGR provides better estimates of standard errors for these complex survey data. This estimation and hypothesis testing strategy is the most conservative approach, since it will produce larger standard errors for the parameter estimates due to the clustering in the data. Extensive empirical investigations of the role of weights and design effects in the NHANES I survey demonstrated that test statistics are decreased when including weights, and decreased even further when adjusting for design effects (see reference 3).

The two-stage procedure adopted by Ethly was described in the preceding subsection.

Models Used in the Analyses

There is no unique correct approach to analyzing the relationships within the NHANES II data or between the NHANES II and other data sets. For this reason, it has been useful to compare and contrast a variety of approaches and models.

All of the models have the general character that a measure of blood lead is expressed as a linear combination of a measure (or measure) of exposure to gasoline lead with various demographic and socioeconomic covariables and (sometimes) time.

The primary difficulty with the Ethyl analyses (App. D2, item 26) lies in the choice of constructed gasoline-lead variables. Neither the population density variable (C19) nor the lead usage variable (C16) is an acceptable measure of gasoline lead exposure.

The Ethyl report concludes with the observation

> In summary, our analysis of the NHANES II data has shown that time (T) is the major contributor to differences in blood lead between

> 1976 and 1980 ... The major contribution of time to the decrease in blood lead indicates that other factors that vary with time are the major causes of the 1976 to 1980 decrease in blood lead and not gasoline lead usage.

Ironically, national gasoline lead usage (as defined in the CDC or ICF analysis) is such a variable that varies with time and is known to be causative of some portion of the lead in blood. The constructed variable (C16) does not display a similar relationship with time.

The CDC and ICF/EPA analyses are similar in their general approach. In each case, a variety of models was considered (adding and deleting various subsets of the covariables and interaction terms). These variations had only minor impact on the value of the coefficient for the lead usage variable.

Although both the CDC and EPA/ICF analyses used national data on leaded gasoline sales, the EPA/ICF models utilized a gasoline lead use variable which was estimated at each month of the survey (App. D2, item 11, Table 1, pp. 13-14). Consequently, since the data survey spanned across two months, the gasoline lead use variable could, and in some cases did, assume two different values for the same site, according to the month of examination. Investigations of the same site, according to the month of examination. Investigations of the relationships between time and blood-lead levels involved comparisons within sites (due to spanning two months), as well as among sites. Thus, even though there is a high degree of correlation between time and gasoline lead usage, these two variables are not completely confounded with the 64 different sites.

It is, nevertheless, a significant question whether the time variable is included in the model as a covariate. The ICF analysis included a linear time covariable and seasonal effects in the model, "to give the models the ability to attribute temporal variations in blood lead to effects other than gasoline lead" (App. D2, item 11, p. 8). Variables for time and gasoline lead were not included simultaneously in the CDC analysis.

The intent of the ICF procedure is reasonable, but the confounding between time and gasoline lead in the data make the simultaneous inclusion of these variables in the model questionable. The data do not allow the relationship between gasoline lead and blood lead to be estimated at any particular time point. Thus the attempt to adjust for time is highly dependent on the specification of the time effects in the model. Despite these problems, two aspects of the CDC analysis yielded some circumstantial evidence that gasoline lead is an important agent of the trend in blood lead. The gasoline lead variable accounted for seasonal variation in blood lead, and the lagged gasoline lead variables provided a plausible lag structure: the one-month lagged variable had the strongest association with blood lead.

Gasoline Lead as a Causal Agent for the Decline in Blood-Lead Levels

The CDC and ICF analyses provide strong evidence that gasoline lead is a major contributor to the decline in blood lead over the period of the NHANES study. DuPont stressed the limitations of statistical theory and methods as tools for assessing causal relationships.

Analysis of the NHANES II data cannot prove whether changes in the use of leaded gasoline caused a change in average blood-lead levels. Variables X and Y can be correlated because changes in X cause changes in Y, or vice versa, or because some third factor, Z, affects both X and Y. There are many other possibilities as well, but these are enough for this discussion. If X stands for some measure of average blood lead concentration and Y stands for the amount of lead in gasoline, we can dismiss the first possibility as absurd. But the relative plausibility of the other two is a matter for expert scientific judgement. To date, no hypothesis of the third form which could explain the NHANES II data has been presented to the panel. One hypothesis of this form has been discussed. This hypothesis has Z representing regulatory changes and publicity aimed at reducing lead exposure generally. This could result in reductions in gas lead, lead in food, lead in paint, etc., and it could be that the gas lead change had little effect on blood-lead levels -- the blood-lead changes might have been caused by the other factors (food, paint, etc.). Although this hypothesis cannot be disregarded entirely, it → does not seem to explain the blood-lead drop adequately. We have seen little evidence that food lead has dropped by a factor large enough to explain a sizable part of the drop in blood lead. In fact, the FDA diet lead values shown in the ICF Report (App. D2, item 11, Table 2) were increasing during the study period. That changes in exposure to leaded paint caused the decrease in blood-lead observed over all age and sex groups seems highly unlikely. The → existence of influences (other than gasoline lead usage) that are not included in the models must be recognized as a limiting factor in the evaluation of all of the analyses.

Use of NHANES II Data for Forecasting Results of Alternative Regulatory Policies

Regression models have been used in all three analyses to see if the NHANES II time trend in average blood-lead levels can be explained in terms of changes in demographic variables or in terms of changes in gas and lead usage. Extension of the use of these and other statistical techniques "to estimate the distribution of blood-lead levels of whites, blacks, and black children and to forecast the results of alternative regulations," as in Section III of the ICF Report of December, 1982 (App. D2, item 11), raises questions and → involves assumptions that go much further than those the Review Group was able to consider. In general, the Review Group would warn that the weaknesses that have been discussed in the context of analyzing relationships within the four-year survey period become enormously greater in any attempt to extrapolate beyond that period. For example, the cautions mentioned in the ERC review (App. D2, item 22, p. 6) of the ICF analysis probably do not go far enough.

Summary

In general, there is a significant correlation between gasoline-lead levels and blood-lead levels in persons examined in the NHANES II Survey. Major obstacles interfere with the use of the available data to describe the relationship. They are: the need to perform model-based adjustments to compensate for imbalance in the design of the NHANES II, the possibility of specification error in the regression models, and the lack of a satisfactory measure of individual or local exposure to gasoline lead, in addition to sampling error, laboratory measurement error, and the high nonresponse rate.

The Review Group finds that the Ethyl analyses contribute little to understanding the association between blood lead and gasoline lead because the variables adopted to represent lead exposure are deemed inappropriate.

The CDC and ICF/EPA analyses relating the NHANES II blood-lead data to a national measure of the amount of lead used in gasoline indicate that the drop in average blood-lead levels can be explained, in large part, by the concurrent drop in gasoline lead. This by no means confirms the hypothesis that the blood lead decrease was caused by the decrease in gasoline lead but, in the absence of scientifically plausible alternative explanations, that hypothesis must receive serious consideration.

References

Literature cited in this report, in addition to the documents furnished by the EPA which are listed in Appendix D2.

(1) Ling, R. F. (1982). A review of Correlation and Causation by David A. Kenny, John Wiley & Sons. J. Am. Statis. Assoc. 77, 490-491.

(2) Goldberger, A. S. (1961). Step wise Least Squares: Residual Analysis and Specification Error. J. Am. Statis. Assoc. 56, 998-1000.

(3) Landis, J. R., Lepkowski, J. M., Eklund, S. A. and Stehouwer, S. A. (1982). A General Methodolody for the Analysis of Data from the NHANES I Survey. Vital and Health Statistics, NCHS Series 2- No. 92. DHHS Publ No. (PHS) 82-1366. Washington. U.S. Government Printing Office.

Appendix D1

Questions for the Review Group

The following questions were stated in letters to members of the Review Group form Dr. Lester D. Grant, Director of the EPA Environmental Criteria and Assessment Office, February 17, 1983.

1. To what extent is it valid to use the NHANES II data to determine time trends for changes in nationally representative blood-lead values for the years of the study (1976-1980)? More specifically, to what extent can the NHANES II data appropriately be used to define time trends for blood-lead levels (aggregated on an annual, semiannual, or any other time-related basis) for the total NHANES II sample (all ages, sexes, races, etc.) or for subsamples defined by the following demographic variables: (1) age (e.g., children <6 years old, children 6-12 years old, adults by 10- or 20- year age groups); (2) sex; (3) race; (4) geographic location (e.g., urban vs. rural residence; Northeast vs. Southeast, Midwest, or other large regional areas of the U.S.; residence in specific cities, towns, or rural locales); (5) socioeconomic status; (6) occupation of respondants or their parents/head of household at main residence; or (7) any combination of such demographic variables (e.g., black children <6 years or white children <6 years old living in urban or rural areas, etc.).

2. If it is indeed possible to derive such time trends from the NHANES II data, to what extent can the changes in NHANES II blood-lead levels over time be correlated credibly with changes in the usage of leaded gasoline over the same time period (i.e., the years 1976-1980)? Several analyses of this type have already been conducted and submitted to us, and we would appreciate your evaluation of those analyses.

3. Are there any other appropriate credible statistical approaches or analyses, besides those alluded to as already having been done, that might be carried out with the NHANES II data to evaluate relationships over time between blood-lead levels and gasoline lead usage?

Appendix D2

Documents Considered by
NHANES II TIME TREND ANALYSIS REVIEW GROUP

1. Plan and Operation of the Second National Health and Nutrition Examination Survey. (1976-1980) National Center for Health Statistics, Series 1, No. 15. July, 1981.

2. Public Use Data Tape Documentation. Hematology and Biochemistry, catalog number 5411. NHANES II Survey, 1976-1980, NCHS. July, 1982.

3. NHANES II Weight Deck (one record for each SP). Deck #502. Attachment I, NCHS.

4. NHANES II Sampling Areas. Document furnished by NCHS during site visit, March 10, 1983.

5. Steps in Selection of PSU's for the NHANES II Survey. Document furnished by NCHS during site visit, March 10, 1983.

6. Location of Primary Sampling Units (PSU) chronologically by pair of caravans: NHANES II Survey, 1976-80. Document furnished by NCHS during site visit, March 10, 1983.

7. Annest, J. L. et al. (1982) Blood lead levels for person 6 months - 74 years of age: United States, 1976-1980. NCHS ADVANCEDATA, No. 79, May 12, 1982.

8. Mahaffey, K. R. et al. (1982) National estimates of blood lead levels: United States, 1976-1980. Association with selected demographic and socioeconomic factors. New England Journal of Medicine 307: 573-579.

9. Average Blood Lead Levels for White Persons, 6 months - 74 years stratified chronologically by PSU's: NHANES II, 1976-80 by caravan. "Graph" furnished by NCHS, March 17, 1983.

10. Schwartz, J. The use of NHANES II to investigate the relationship between gasoline lead and blood lead. Memo to David Weil (ECAO) (March 3, 1983).

11. ICF Report: The Relationship between Gasoline Lead Usage and Blood Lead Levels in Americans: A Statistical Analysis of the NHANES II Data. December 1982.

12. Annest, J. L. et al. (1983) The NHANES II study. Analytic error and its effect on national estimates of blood lead levels.

13. Pirkle, J. L. Comments on the Ethyl Corp. analysis of the NHANES II data submitted to EPA October 8, 1982 (Feb. 26, 1983).

14. Pirkle, J. L. Chronological trend in blood lead levels of the second NHANES, Feb. 1976-Feb. 1980 (Feb. 26, 1983).

15. Lynam, D. R. Letter to David Weil dated October 15, 1982 containing additional comments on NHANES II data.

16. E. I. DuPont de Nemours & Co., Inc. Supplementary statement presented to EPA in the matter of regulation of fuel and fuel additives - lead phasedown regulations proposed rulemaking (Oct. 8, 1982).

17. Pirkle, J. L. An expanded regression model of the NHANES II blood lead data including more than 100 variables to explain the downward trend from Feb., 1976-Feb., 1980 (Dec. 23, 1982).

18. Annest, J. L. et al. Table 1. Average blood lead levels and total nonresponse rates for persons ages 6 months - 74 years stratified chronologically by primary sampling unit (PSU): NHANES II, 1976-1980 (Corrected version; April 8, 1983).

19. Pirkle, J. L. (1983). Duplicate measurements differing by more than 7 mg/dl in the lead measurements done in NHANES II Survey. Document furnished by CDC at Panels request, March 18, 1983.

20. Pirkle, J. L. Appendix M: Tabulation by demographic variables (March 18, 1983).

21. Pirkle, J. L. Appendix N: Regression analysis of urban and rural population subgroups (March 18, 1983).

22. Miller, C. and Violette, D. Comments on studies using the NHANES II data to relate human blood lead levels to lead use as a gasoline additive (March, 1983).

23. Miller, C. and Violette, D. (March 4, 1983). The Usefulness of the NHANES II Data for Discerning the Relationship between Gasoline Lead Levels and Blood Lead Levels in Americans and a Review of ICF's Analysis using the NHANES II Data. Energy and Resource Consultants, Inc.; Boulder, Colorado.

24. Schwartz, J. Analysis of NHANES II data to determine the relationship between gasoline lead and blood lead. Memo to David Weil (ECAO). (March 18, 1983).

25. Excerpt - (Section I. C. - "Discussion of NHANES II Blood Lead Data") from the Ethyl submission to the EPA's docket on the Lead Phasedown dated May 14, 1982.

26. Excerpt - (Section III. A. - entitled "Correlation of Blood Lead to Gasoline Lead" and Appendix "Discrete Linear Regression Study") from the Ethyl submission to EPA's docket on the Lead Phasedown. (October 8, 1982)

27. Ethyl Analyses of the NHANES II Data. This item was distributed at the Criteria Document meeting held on January 18-20, 1983.

28. Comments by Dr. Norman R. Draper on Ethyl Corporation's comments and ICF, Inc.'s comments.

29. Comments by Dr. Ralph A. Bradley entitled "A Discussion of Issues and Conclusions on Gasoline Lead Use and Human Blood Lead Levels".

30. Comments by Dr. Ralph A. Bradley in a letter to B. F. Fort. (Ethyl Corp.)

31. Ethyl Corp. NHANES II - blood lead data correlation with air lead concentration data.

32. Ethyl Corp. Summary of analyses of the NHANES II blood lead data (January, 1983).

33. E. I. DuPont de Nemours & Co. Comments submitted March 21, 1983.

34. E. I. DuPont de Nemours & Co. Comments by R. Snee and C. Pfieffer on paper by Annest et al. on analytic error (see item #5).

35. Pirkle, J. L. The relationship between EPA air lead levels and population density. (March, 1983).

36. Pirkle, J. Consecutive numbering of points on plots of 6-month average NHANES II blood lead levels versus 6-month total lead used in gasoline (April 11, 1983).

37. Pirkle, J. L. Distribution of the NHANES II lead subsample "weight" variable (April 11, 1983).

38. Pirkle, J. L. Appendix O: Propagation of error in calculating the percent decrease in blood lead levels over the NHANES II survey period (April 11, 1983).

39. Pirkle, J. L. Appendix P: Regressing ln (blood lead) on the demographic covariates and then regressing the residuals on GASQ compared to regressing ln (blood lead) simultaneously on the demographic covariates + GASQ (April 11, 1983).

40. Pirkle, J. L. Appendix Q: Regression of ln (blood lead) on the demographic covariates only and subsequently adding GASQ: F statistics, R square and Mallows $C_{(p)}$ (April 11, 1983).

Appendix D3

List of Attendees at March 10-11 and March 30-31, 1983
meeting of
NHANES II TIME TREND ANALYSIS REVIEW GROUP

Panel Members

Joan Rosenblatt (Chairman)
National Bureau of Standards

J. Richard Landis
University of Michigan

Roderick Little
Bureau of the Census

Richard Royal
Johns Hopkins University

Harry Smith
Mt. Sinai School of Medicine

David Weil (Co-chairman)
U.S. EPA

Observers

Dennis Kotchmar*
U.S. EPA

Vic Hasselblad
U.S. EPA

Allen Marcus
U.S. EPA

George Provenzano
U.S. EPA

Joel Schwartz
U.S. EPA

Earl Bryant*
NCHS

Trena Ezzote*
NCHS

J. Lee Annest
NCHS

Mary Kovar*
NCHS

Bob Casady*
NCHS

Jean Roberts*
NCHS

Robert Murphy
NCHS

Vernon Houk†
Centers for Disease Control

James Pirkle
Centers for Disease Control

Don Lynam
Ethyl Corporation

Ben Forte
Ethyl Corporation

Jack Pierrard*
DuPont

Chuck Pfieffer
DuPont

Ron Snee
DuPont

Asa Janney
ICF

Kathryn Mahaffey*
FDA

*attended March 10-11 meeting only.
†attended March 30-31 meeting only.