OR-6828-D REV. 5-64

E. I. DU PONT DE NEMOURS & COMPANY
INCORPORATED

PETROLEUM LABORATORY
WILMINGTON, DELAWARE 19898

TELEPHONE
AREA CODE 609-299-3500
540-2874

December 5, 1983

Professor Ralph A. Bradley
325 Hickory Hill Drive
Watkinsville, GA 30677

Dear Professor Bradley:

Accompanying this letter are portions of a draft report on analysis of the NHANES II blood lead study, for your review prior to our meeting December 15 in Atlanta arranged by Ben Fort.

The statistical analysis is primarily the work of Dr. Charles G. Pfeifer of the Du Pont Engineering Department. Dr. Pfeifer will attend the meeting on December 15, as will Dr. Ronald D. Snee, also of the Engineering Department.

This draft material is intended to provide sufficient detail for your critique of the approach we have adopted, and its execution. Should you note any critical omission, please call me or Chuck Pfeifer (302-366-3540).

I am also enclosing a copy of a report (PLMR 58-83) which presents part of our analysis in less detail than in the draft material. Perhaps it will be useful to give you an overview of our line of argument.

I look forward to meeting you on the 15th.

Very truly yours,

JMPierrard

John M. Pierrard
Engineering Fellow
Environmental Activities

JMP/er
Enc.

cc: ~~R. D. Snee~~
C. G. Pfeifer
D. Lynam - Ethyl
B. Fort - Ethyl

N33844

| | L1 | L2 | L3 |
|---|---|---|---|
| % B | 70 | 50 | 10 |
| Ga Pb | O | O | O |
| n | 100 | 82 | 193 |

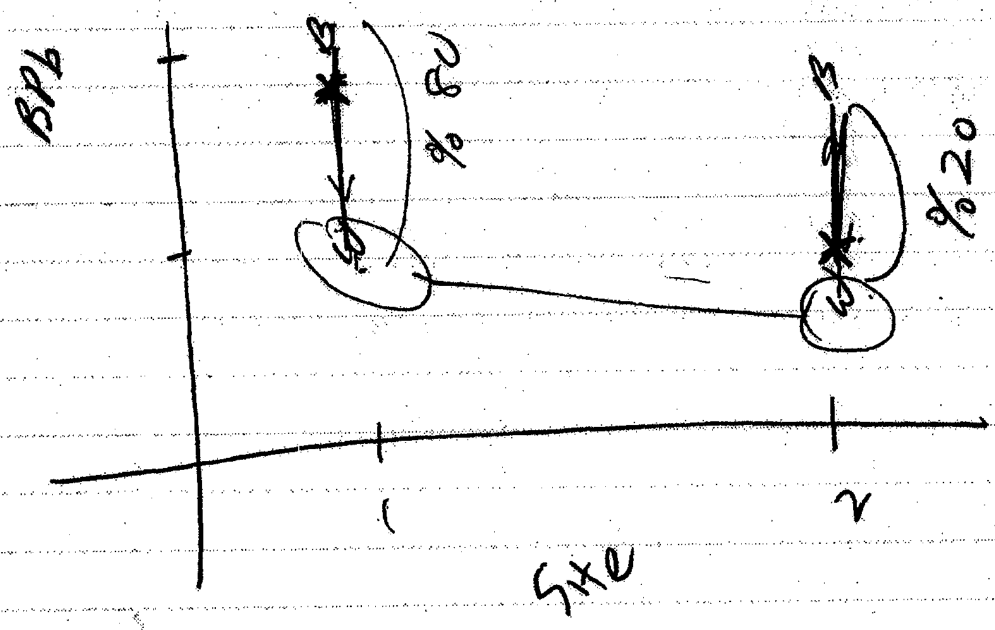
$$b_W (X_{ij} - X_{i.}) + b_B (X_{i.} - X_{..})$$

$$Y = \mu + b_W (X_{ij} - X_{i.}) + b_B (\quad) + c \, \underline{\text{Ga Pb}}$$

TEH 0532680

N33844.01

DUP050033950




$$y_i = \mu_i + b_W (X_{ij} - X_{i\cdot}) + b_B (X_{i\cdot} - X_{\cdot\cdot})$$

$$= \mu_i + b_W$$

$y_{ij} = \mu + R_i + L_j + e_{ij}$

| | | L1 | L2 | L3 | L4 |
|---|---|---|---|---|---|
| ~~Locations~~ | B | 3 | 2 | 0 | 0 |
| | W | 2 | 2 | 0 | 0 |
| ~~Quly~~ | %W ~~Race~~ | | | | |

~~Race~~

| | |
|---|---|
| Location | 3 |
| Race | 1 |
| RxL | 3 |
| ~~Quly~~ (LxR) | |

4 / 180 45

| | L1 | L2 | L3 | L4 | AVG |
|---|---|---|---|---|---|
| B | 20 | 30 | 50 | 80 | 45 |
| W | 80 | 70 | 50 | 20 | 55 |

$b_W$ = Race Diff

$b_B$ = effect of % Black

| $R_1$ | $L_1$ | $L_2$ | $L_3$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 |
| 1 | 0 | 0 | 1 |

| $n$ | $R_1$ | $L_1$ | $L_2$ | $L_3$ |
|---|---|---|---|---|
| 20 | 0 | 0 | 0 | 0 |
| 30 | 0 | 1 | 0 | 0 |
| 50 | 0 | 0 | 1 | 0 |
| 90 | 0 | 0 | 0 | 1 |
| 80 | 1 | 0 | 0 | 0 |
| 70 | 1 | 1 | 0 | 0 |
| 50 | 1 | 0 | 1 | 0 |
| 10 | 1 | 0 | 0 | 1 |

$$y_{\cdot} = b_0 + b_1 R_1 + b_2 L_1 + b_3 L_2 + b_4 L_3$$

$$y_i = b_0 + b_1 R_i + b_2 L_1 + b_3 L_2 + b_4 L_3$$

| | $L_1$ | $L_2$ | $L_3$ | $L_4$ | |
|---|---|---|---|---|---|
| B | 80 | 105 | 40 | 40 | 55 |
| W | 20 | 45 | 40 | 160 | |
| TOTAL | 100 | 150 | 80 | 200 | 45 |
| %B | 80 | 70 | 50 | 20 | $\boxed{55 = \text{AVG}}$ |

$$y_j = b_0 + b_1 R_j + b_2 L_1 + b_3 L_2 + b_4 L_3$$

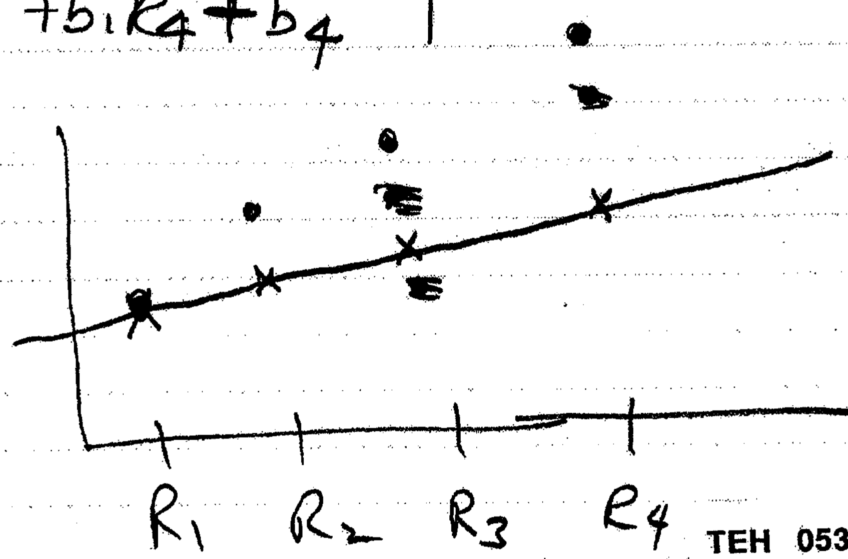
Site 1 $\hat{y} = b_0 + b_1 R_1$

Site 2 $\hat{y} = b_0 + b_1 R_2 + b_2$

Site 3 $\hat{y} = b_0 + b_1 R_3 + b_3$

Site 4 $\hat{y} = b_0 + b_1 R_4 + b_4$

$$\hat{y}_i = b_0 + b_1 R_i + a_1 R_i = b_0 + (b_1 + a_1) R_i$$



12/14/83

Issues

- Role of SURREG ← MODEL / weights
- Lead Exposure
  - Local
  - national (Food chain etc)

Is this the

SURREG ANOVA ?

| | | | |
|---|---|---|---|
| Total | N-1 | Regn Coeff | |
| Strata | 31 | | |
| PSU(S) | 32 | | ERROR 1 |
| Subjects | N-1-31-32 | | ERROR 2 |

SURREG

→ weights are estimated

Should we try to fit the PSU(Strata) terms in our model

within Location time trends

what do the distribution of 64 slopes look like

TEH 0532685

DUP050033955

why is drok present no matter how you ~~segre~~ segregate the data?

RDS Several ($p \approx 5$) variables have equal effects ~~&~~ change over time ∴ Total effect with change over time

CBP

- X-Covariate Within Site
- Z-Covariate Between Site

$\beta_{within}$ $\sigma_w$ $\beta_{Between}$ $\sigma^2_{BSN} = \sigma^2_B + \sigma_w$

Pechile
$\beta_w = \beta_B$
$\sigma_B = 0$

Pirkle used a single var term model when time was in the model

$$y = Dem + b_L T$$

$$y - Dem = b_R T$$

$$\boxed{b_L = VIF\,(b_R)} \leftarrow \text{Basis ?}$$

$VIF = 3.5$

$\cancel{\frac{1}{1-3.5^2}}$

Could Ridge Regr help out here

$$\frac{1}{1-r^2} = 3.5$$

$$1 = 3.5 - 3.5r^2$$

$$r^2 = \frac{3.5 - 1}{3.5} = \underline{71\%}$$

4

$$\frac{.000298}{.0000816} = \frac{2.98 \times 10^{-4}}{8.16 \times 10^{-5}}$$

$$\doteq 3/8 \times 10^{1} = \underline{3.75} \equiv VIF$$

$$= \sim 3.40$$

* How does Ethyls gas variables compare with DuPonts gas variables?

* we need a graphical display to show how DuPont ~~Our Pb~~ Gas Pb time trend come about

* Could we sell urban as Dust + Dirt Pb?

* Look at air Pb trends in air Quality criteria Document

TEH 0532688

DUP050033958

Ethyle (Clayalit Leighfoot) has data on food Pb which comes from FDA — we need to develop these dietary lead exposure

Effects of uncertainty in X-variables

why the diff between
DuPont (2σ model)
VS
$\boxed{\text{single } \sigma \text{ model}}$

identify adj function

CBP

note with ~~the~~ adj doesn't have much of an effect

TEH 0532689

DUP050033959

why is $\beta_s \neq \beta_{m}$?

They are assu with different variables

CGP out the 3rd wk of January

CGP did some ridge Regr need to discuss

1/6 1/9 next mtg?

TEH 0532690

DUP050033960

# ANALYSIS OF NHANES II

# BLOOD LEAD STUDY

PLMR-58-83

J. M. PIERRARD
PETROLEUM LABORATORY
E. I. DU PONT DE NEMOURS & CO.
SEPTEMBER 21, 1983

The U.S. Government's second National Health and Nutrition Examination Survey - known as NHANES II - was designed to assess nutritional status of the population based on questionaires and examinations administered in 64 sample areas over a 4 year period. As an added part of the study, venous blood lead measurements were made on 9936 persons (1).

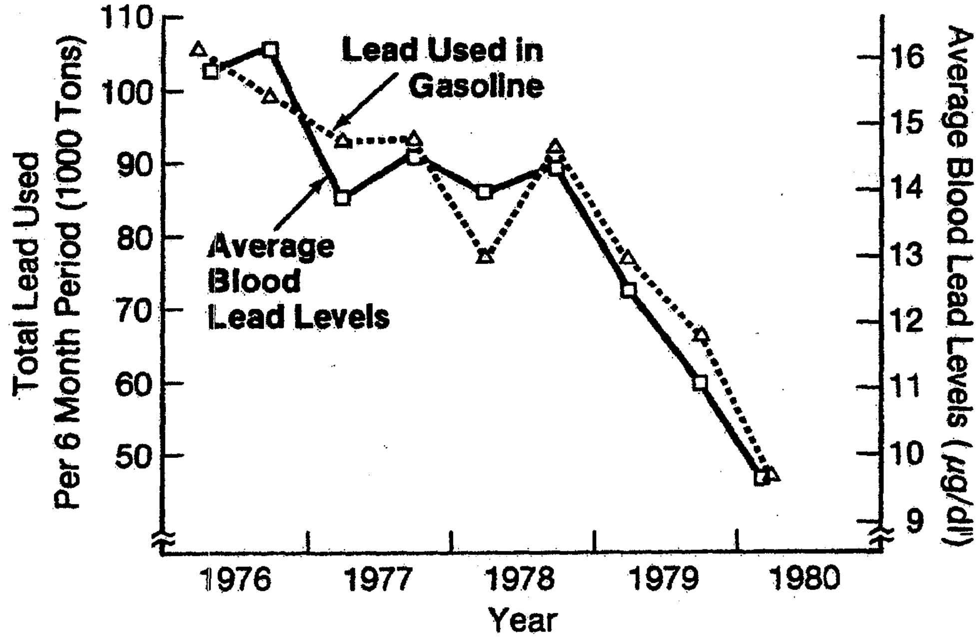
The study was designed for cross-sectional analysis - that is, to obtain a value for each measured variable representative at the study midpoint - and not for chronological analysis. However, an apparent blood lead decrease of 37% over the 4 year study period prompted attempts to analyze the time trend (2).

## NHANES II

- Assessment of National Health Status
- February 1976 to February 1980
- 64 Sampling Sites
- 6 Months to 74 Years of Age
- 27,801 Persons Selected — 20,322 Examined
- 9,936 Venous Blood Lead Levels Determined
- Not Designed for Chronological Trend Analysis
- Apparent 37 Percent Reduction in Blood Lead

Attention was drawn to this representation of the parallel declines of six month average blood lead and total U.S. gasoline lead use (3). This figure might well have led some to infer that the blood lead decline was solely due to the decline in total gasoline lead use.

# Gasoline Lead and NHANES II Blood Lead Levels (Feb. 1976 — Feb. 1980)




TEH 0532695

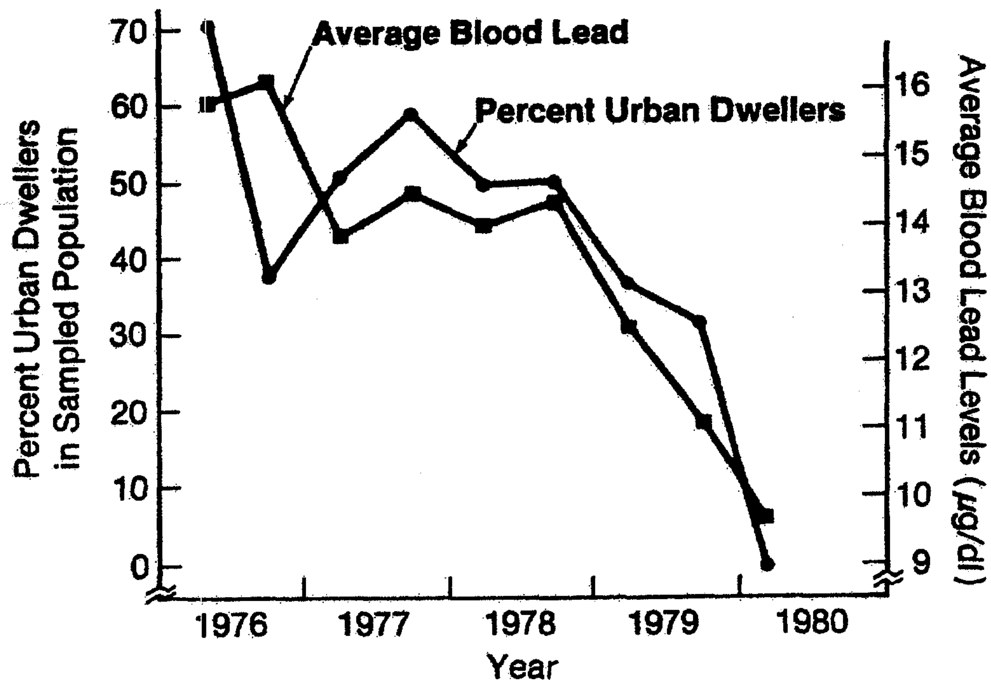
DUP050033965

Once the basic data became available, further examination showed that there is substantial confounding because of nonuniformity of subject characteristics over the 4 year course of the NHANES II study.

This figure shows one example of such confounding (4).

These considerations resulted in a number of more detailed analyses, including our own. To contrast the features of the various analyses, it is helpful to consider the identifiable factors which might affect apparent blood lead.

## Urban Dwellers Sampled and NHANES II Blood Lead Levels (Feb. 1976 — Feb. 1980)




TEH 0532697

DUP050033967

It is generally agreed that exposure to lead may occur via a number of routes. In several analyses of the NHANES II data conducted or sponsored by government groups, it was assumed that only exposure to lead from gasoline decreased during the study period (5,6,7). Their assumption that food lead was unchanged was justified on the basis of calculated lead content for a constant market basket of foods (8). However, 64% of the U.S. households surveyed in 1979 by the Economics and Statistics service of the U.S. Department of Agriculture reported diet changes during the period 1977 to 1979 (9). Furthermore, aluminum cans were displacing soldered cans during this period (10). Therefore, we made provision in our analysis for possible changes due to these or other unquantifiable factors.

Returning to gasoline lead, it should be noted that the other analyses referred to used total national gasoline lead use as the index of exposure. Yet, the 64 sites visited during NHANES II had a wide range of lead exposures as shown by air lead data. To account for this variability, we calculated the site-specific gasoline lead density for each location visited. Each site consisted of one or more counties. The gasoline lead consumed in the site county at the time of the visit was calculated from state gasoline consumption and lead dosage, and county and state populations, using the fact that gasoline consumption is highly correlated with population in most of the U.S. (11). The site-specific lead <u>density</u> then was calculated by dividing county lead use by county area. This approach allows for exposure to gasoline lead by ingestion as well as inhation.

The study design factors listed in the chart were treated as classification variables. Caravan itinerary covers region of the country and season. Personal variables are sex, age, and race, and residence variables are size of city or town and in or outside of the core city of a Standard Metropolitan Statistical Area.

Quality control data on the laboratory analysis of blood leads for the last 3 years of NHANES II do not indicate any significant problem. However, there were no check samples carried through the sequence of field collection, storage, shipment and final analysis at the home laboratory, so we allowed for a possible trend contribution from this source.

## Identifiable Factors Potentially Influencing Blood Lead Trend

**Exposure**
- Air Lead from Gasoline
- Food and Beverages
- Water
- Paint
- Dust

**Study Design**
- Caravan Itinerary
- Personal Variables
- Residence Characteristics

**Study Execution**
- Laboratory Analysis
- Sample Handling

We investigated the time trend contribution attributable to the demographic factors alone, for demographics and gasoline lead, and for demographics, gasoline lead and time, which was used as a surrogate for unquantifiable time-related blood lead effects. Details are given in our paper (12) which is appended to this report.

Of particular interest is this comparison of national gasoline lead use and site-specific lead density. National gasoline use is not statistically significant as a blood lead predictor in the presence of the time surrogate for other time-related blood lead effects. On the other hand, site-specific lead density and its logarithm retain their significance.

## Significance of National Gasoline Lead and Site Specific Lead Density

| Gasoline Lead Exposure Variable | Significance Level* Lead Exposure Variable | Time |
|---|---|---|
| National | 0.8850 | ≤ 0.0001 |
| Site Density | 0.0022 | ≤ 0.0001 |
| Log (Site Density) | 0.0019 | ≤ 0.0001 |

*All models included the full set of 30 demographic terms.

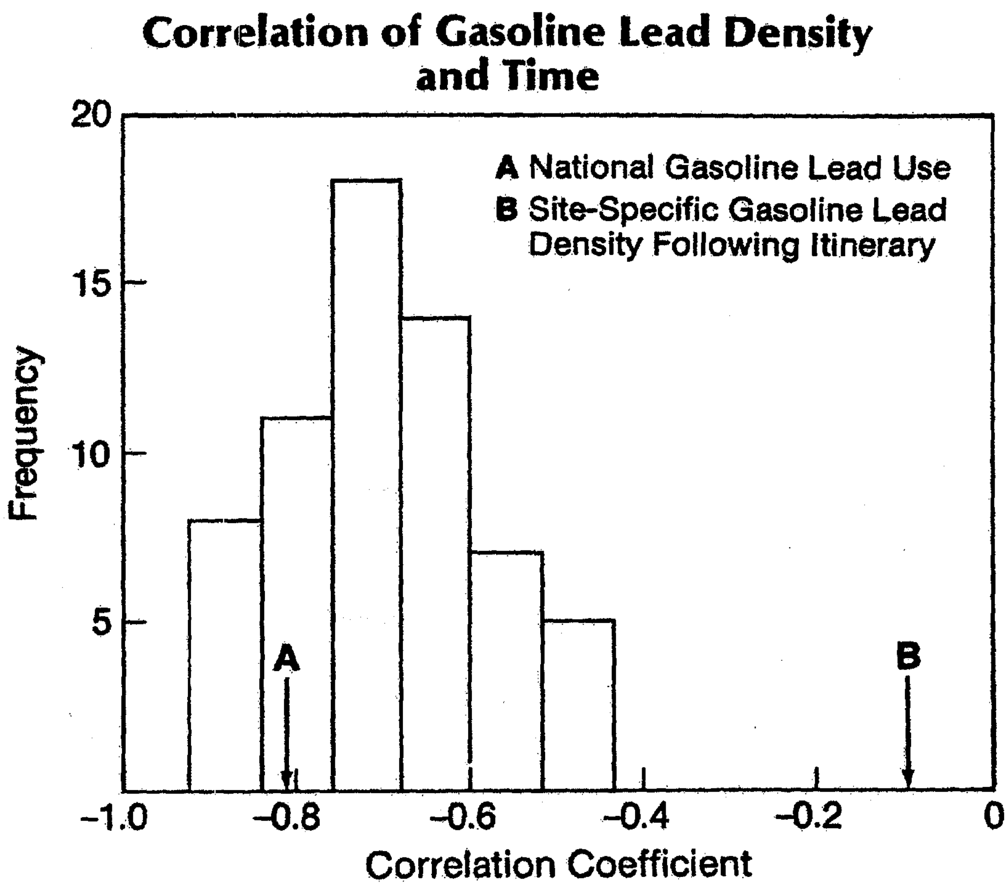
TEH 0532701

DUP050033971

The histogram shows the frequency distribution of the correlations with time of the site-specific lead densities at each of the NHANES II sites over the 4 years of the study.

The arrow at A shows that total national gasoline lead use is highly correlated with time. But, as indicated by the arrow at B, the site-specific lead densities at the sites when they were actually visited is poorly correlated with time.

Thus, national gasoline lead use not only fails to reproduce the time history of lead exposure as the caravans moved from site to site, but it also obscures the influence of other potential time-related effects because it is highly correlated with time.



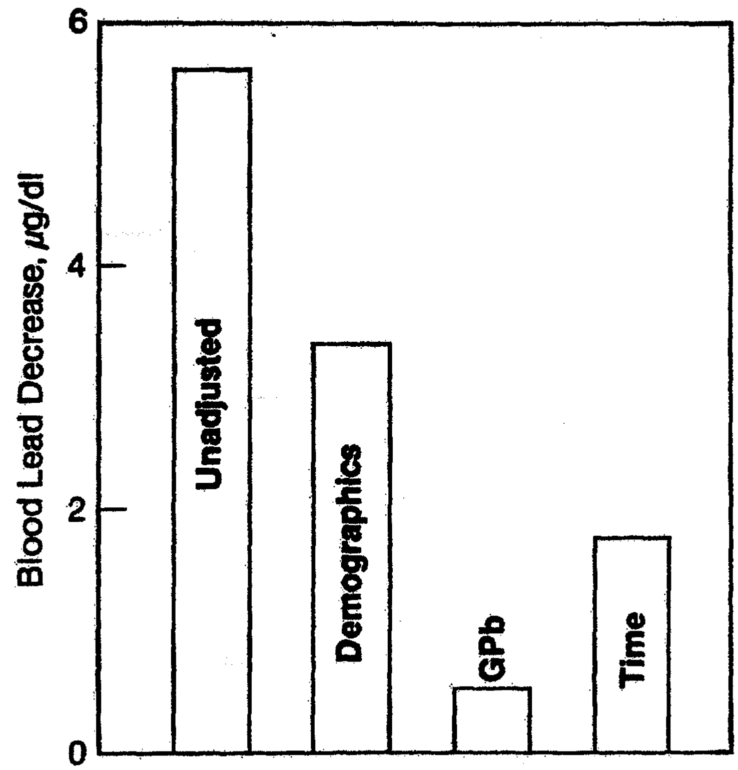

TEH 0532703

DUP050033973

Using site-specific lead density as the measure of exposure to gasoline lead, stepwise regression analysis showed that

- Over half the apparent blood lead decrease can be accounted for by changing demographics.
- About one-tenth of the apparent blood lead decrease can be accounted for by gasoline lead exposure.
- A significant time trend still remains after accounting for demographics and gasoline lead exposure.

## Contributions to Blood Lead Decrease 1976-80




1

TEH 0532705

DUP050033975

To complement the stepwise regression analysis reported in (12), we analyzed the NHANES II blood lead data by the method of split-plot covariance analysis (13). This method takes its name from agricultural applications, where a field or "plot" is "split" and planted with two varieties of plant. To validly compare performance of the two varieties grown under the same conditions "within-plot" corrections - such as for different numbers of the plant varieties - must be made. To compare performance of the varieties in different fields, or plots - say under different climatic conditions - "between-plot" corrections are required.

The situation is analogous for the NHANES II blood lead data set. Here, corrections are required for within-site differences in the characteristics of the subjects. The within-site analysis adjusts the raw observed blood lead value of each individual at a site to the value he would have if his personal and residence characteristics were the average values for the site.

The between-site analysis amounts to an adjustment of the site-adjusted blood leads to the value each would have if every site had the same value of each of the personal, residence and itinerary variables. After these adjustments all observations are on a common basis relative to demographics, and the remaining time trend is attributable to lead exposure factors and/or other trend influences.

By omitting or including site-specific lead density in the adjustment process, we were able to estimate the contribution of gasoline lead to the blood lead decrease.

## Steps in Blood Lead Trend Analysis

| | |
|---|---|
| Within-Site Analysis | Adjust for Differing Characteristics of Individual Subjects |
| ↓ Between-Site Analysis | Further Adjust for Differing Site Characteristics — With and Without Gasoline Lead |
| ↓ Residual Time Analysis | Quantify Unexplained Time Trend |
| ↓ Gasoline Lead Contribution Analysis | Compare Unexplained Time Trend With and Without Gasoline Lead Exposure |

The split-plot analytical approach was applied to 216 models for blood lead as listed.

An important result is that use of the NHANES II blood lead weighting factors, or the basis of residual weighting, or the transformation of the dependent variable blood lead did not affect the overall results, which were consistent throughout.

## Analyses for Residual Time Trend

| Factor | Options |
|---|---|
| BPb and ln BPb | 2 |
| P, R, I and 26 combinations of current and lagged site specific gasoline lead density and their logarithms | 27 |
| With and without BPb weighting factors | 2 |
| Residuals weighted according to $1/S^2$ and N | 2 |

Number of models = 2 x 27 x 2 x 2 = 216

**Consistent results regardless of BPb weights, residual weighting, and BPb or ln BPb**

TEH 0532709

DUP050033979

Because the results were consistent, the overall picture can be grasped by examining a few cases.

For untransformed blood lead as the dependent variable, before adjustment to remove any effects, the 4-year decrease is 5.8 ug/dl. After within-site and between-site adjustment for itinerary, personal, and residence (DEM) variations, this is reduced to about 1.6 $\mu$g/dl, with $R_A^2$=0.53.

Including current site-specific lead density (G Pb) or its lagged values (G $Pb_{30}$ - 30 days lag, G $Pb_{45}$ - 45 days lag) in the model accounts for up to 0.5 $\mu g/m^3$ of the 4-year decrease.

Note that the addition of lagged lead density terms after the first one has little effect on either the blood lead decrease attributable to gasoline lead or on $R_A^2$. Also note in the last two lines that making the adjustment for gasoline lead before any adjustment for demographics reduces the 4-year decrease by less than 1 $\mu$g/dl, but with $R_A^2$ of only 0.20 to 0.21.

# Typical Model Results*

| Effects Removed | Blood Lead, µg/dl: Remaining Four Year Decrease | Attributable to Gasoline Lead | $R^2_A$ |
|---|---|---|---|
| None | 5.80 | | |
| DEM | 1.62 | | 0.53 |
| DEM, GPb | 1.27 | 0.35 | 0.59 |
| DEM, GPb, $GPb_{30}$ | 1.21 | 0.41 | 0.59 |
| DEM, GPb, $GPb_{30}$, $GPb_{45}$ | 1.19 | 0.43 | 0.58 |
| DEM, log GPb | 1.48 | 0.14 | 0.57 |
| DEM, log GPb, log $GPb_{30}$ | 1.10 | 0.52 | 0.60 |
| DEM, log GPb, log $GPb_{30}$, log $GPb_{45}$ | 1.09 | 0.53 | 0.59 |
| GPb | 5.17 | | 0.20 |
| log GPb | 4.93 | | 0.21 |

*Dependent variable BPb, residuals weighted by $1/S^2_i$ no BPb weights

TEH 0532711

DUP050033981

From the site-by-site analysis (12) and from the additional split-plot analysis of covariance on the individual observations just discussed, we conclude that national gasoline lead use is not a valid index because it does not reflect the exposures due to lead use at the NHANES II sites at the times they were sampled. Further, confounding of total national gasoline lead use with time precludes estimation of other unquantifiable, time-related effects.

We found that over half of the apparent blood lead decrease is traceable to changes in site subject demographics.

From a large number of models we estimate that 5 to 10% (or 0.3 to 0.6 $\mu$g/dl) of the blood lead decrease is attributable to gasoline lead exposure.

This estimate is consistent with the existing literature. It agrees with the estimate based on an air lead decrease of about 0.3 $\mu g/m^3$, an $\alpha$ value in the range 1 to 2 (14,15). It is also consistent with the Frankfurt study (16).

Finally, after adjustment for both demographics and gasoline lead exposure, a significant blood lead time trend remains.

## Conclusions

- National gasoline lead use is not a valid index of exposure at NHANES II sites
- Over 50% of the apparent blood lead decrease is attributable to demographics
- 5 to 10% of the apparent blood lead decrease is attributable to gasoline lead exposure
- The blood lead decrease of 0.3 to 0.6 $\mu g/dl$ attributable to gasoline lead is consistent with the observed air lead decrease
- A significant blood lead time trend remains after accounting for demographics and gasoline lead exposure

## REFERENCES

1. United States. National Center for Health Statistics. Plan and operation of the second National Health and Nutrition Examination Survey, 1976-1980. Programs and Collection Procedures, Series 1, No. 15.

2. United States. Center for Disease Control. Morbidity and Mortality Weekly Report 30, 132 (1982).

3. Houk, V. Statement at Environmental Protection Agency Hearing on Regulation of Fuel and Fuel Additives, Lead Phasedown Regulation, Notice of Proposed Rulemaking, April 15, 1982.

4. Petroleum Chemicals Division, E. I. du Pont de Nemours & Co., Inc. Statement presented to Environmental Protection Agency, Science Advisory Board, Environmental Health Committee at an Open Meeting on Public Health Impacts Associated with Lead in Gasoline Phasedown Program, July 7, 1982.

5. Schwartz, J. U.S. EPA Memorandum "Health Effects of Gasoline Lead Emissions", May 11, 1982.

6. ICF Incorporated "The Relationship Between Gasoline Lead Usage and Blood Lead Levels in Americans: A Statistical Analysis of the NHANES II Data", Report to EPA, December, 1982.

7. Pirkle, J.L. U.S. CDC Memorandum "Comments on the Dupont (sic) Analysis of the NHANES II Blood Lead Data", December 1, 1982.

8. Jelinek, C.F. "Levels of Lead in the United States Food Supply", J. Assoc. Off. Anal. Chem. 65, 942-6 (1982).

9. "Growing Concern for Nutrition", Food Processing, p.40, March, 1981.

10. "Beverage Can Shipments", Modern Metals, p. 34, July, 1980.

11. Pierrard, J.M., et al, "Vehicle Emissions Controls and Ambient Air Quality", SAE Australia, Jubilee Year Conference, Melbourne (1977).

12. Pierrard, J.M. et al, "Assessment of Blood Lead Levels in the U.S.A. from NHANES II Data", Proceedings of the International Conference on Heavy Metals in the Environment, p. 421-4, September, 1983.

13. Kempthorne, O. Design and Analysis of Experiments, John Wiley and Sons, New York, 1952.

14. U.S. EPA Office of Air Quality Standards. National Air Quality and Emissions Trend Report, 1981.

15. Snee, R.D. "Evaluation of Studies of the Relationship Between Blood Lead and Air Lead", Int. Arch. Occup. Environ. Health 48, 219-42 (1981).

16. Sinn, W. "On the Relationship Between Lead in Air and Blood Lead Content of Persons Living and Working in the Centre of a City (Frankfurt Blood Lead Study)", Int. Arch. Occup. Environ. Health 47, 93-118 (1980), 48; 1-23 (1981).