GENERAL COMMENTS ON NHANES II RESULTS

Prepared by C. G. Pfeifer

Background

The NHANES II program is a carefully designed statistical survey to gather nutrition and health data reflecting a representative cross section of the U.S. population. The study encompasses four years, from February 1976 to February 1980, covering 64 primary sampling locations - 16 in each of 4 approximately equi-populated regions. At any one point in chronological time at least two, and perhaps three, locations were collecting data. The four-year study was designed and intended to provide a "snapshot" of distributional characteristics for many health and nutritional parameters to be associated with a given point in time, viz, about midstudy or February 1978. Thus, the total four-year sample was intended to be representative of many factors, such as age, race, sex, and income level. However, there was no intention that any period within the study years would in any way be representative of the designed-for demographic factors.

Time Trend Analyses of Blood Lead Within NHANES II

Attempts at estimating time trends of blood lead levels over the duration of NHANES II are being made (Annest and Pirkle, draft; CDC blood lead/gasoline lead correlation). As noted above, the study design does not lend itself to validly inferring trends from these types of analyses. In order to correctly infer from time trends, it would be necessary that within any chronological time period of interest, all demographic factors are representative of the intended population. In this way comparing statistics (eg, arithmetic and geometric means) between time periods would enable valid inferences of trend. Undoubtedly, the NHANES II study design team would agree that if the a priori intention was to use the data to analyze time trends within the study period, then the survey sampling methods used would not be appropriate. In fact, the Plan and Operation of the Second National Health and Nutrition Examination Survey, 1976-1980 by A. McDowell clearly states that a study objective was to monitor changes in health and nutrition over time through successive assessment surveys. There was no intention to monitor within an assessment survey.

To see why comparisons within time periods can provide misleading results, Tables 1 and 2 summarize six month time periods by region sampled according to actual number sampled and population number (in millions) represented by the sampled individuals, respectively. The disparities are evident. There is not one six month period where all four regions were sampled. In several instances, regions were not sampled for an entire year. All samples in 1980 come from the South. Similar disparities are easily displayed for any subdivisions according to age, race and degree of urbanization. What this implies is that observed time trends are partially confounded with the demographic factors. To illustrate this, it is well documented and generally accepted that blacks have higher blood leads than whites. Thus, if a predominantly black area was sampled first followed by a predominantly white area, a downward blood lead trend would result, even if other sources of lead exposure remained unchanged; if the order were reversed, an upward trend would result. Mahaffey, Annest, et al (1979) state that the early study years of NHANES II "overrepresent very young, elderly and low-income individuals and should not be used as normative data." It is well known that the young and low income individuals tend to have higher blood lead levels. Consequently, since the latter part of the study must be underrepresented by these groups, any time trends would be presumably accentuated. Thus, the degree of confidence that can be put into trend analysis is questionable.

The extent of the partial confounding could be partly addressed, but not necessarily sorted out, if sampling times and demographic data were made available for each of the 64 primary sampling units; this information is considered confidential, however. Thus, analyses predicting the percent decrease in blood lead levels (Annest and Pirkle) clearly should be used with great caution. It is unfair to quote statistical inferential results when using unrepresentative data within short time periods. There is little doubt that the overall weight of the evidence strongly suggests that mean blood leads have decreased over the study period. What the NHANES II data cannot tell us is how much and why.

The preceding comments further undermine the appropriateness of generating correlations with gasoline lead consumed. Even if the correlations were accurate, there is no way to conclude from this study that gasoline lead reduction caused all or part of a decrease in blood lead. One can only speculate as to what if any portion of the suspected blood lead decrease is due to gasoline lead as opposed to other reasons such as eliminating sources of lead-based paint poisoning.

Risk of Lead Poisoning

The other side of the coin is, if one believes the blood lead time trend, then one must believe the results from the latter part of the study. Table 3 summarizes some key statistics from NHANES II blood lead data collected after January 1, 1979, on the high risk age group 6 months to 6 years. These results are based on lead weights provided on the data tape and demonstrate

- Geometric mean blood lead levels are all well below 15 µg/dl, the established safe level.

- One percent of the 6 month to 6 year subpopulation have blood leads that exceed 30 µg/dl.

The implications of the results are that by all health criteria already establshed as of mid-1979 the pediatric population is at a safe level. If one further believes that this safe level was achieved through reduction in gasoline lead, then current blood lead levels must be still smaller since total gasoline lead currently is 30 percent less than levels in 1979.

Furthermore, the air lead standard was established assuming a baseline blood lead (no air lead exposure) of 12 µg/dl and results shown in Table 3 suggest a value of no more than 10 µg/dl is appropriate. Accordingly, if 10 were used in determining the standard, the result would have been 2.5 $\mu g/m^3$ - a value 67 percent higher than the established standard. This suggests that the established standard has an extremely large margin of safety built in and, even so, is being achieved at current gasoline lead levels.

Finally, the estimate of 1 percent at risk in mid-1979 is biased upward according to a study by Lucas (1981), where bias in reported blood leads above 30 µg/dl can be up to 1 percent of the entire population. Also, some of the blood leads truly above 30 are individuals undoubtedly suffering from blood lead poisoning due to ingestion - a state that would exist in the absence of any air lead exposure.

Summary

The foregoing discussion raises two main points. One is that it is extremely questionable as to how much accurate information is possible when developing time trends within the entire NHANES II study period when the short period samples are clearly nonrepresentative. Consequently, due to these limitations, the study is, strictly speaking, useful only when analyzing data not involving chronological time. Although evidence that blood leads have decreased over the study period is strong, it is not useful in accurately estimating blood lead level trends. The other point is that if the trends are to be beleived, then one must accept that the results obtained in the latter part of the study are an accurate portrayal of the U.S. during that period. These results strongly indicate, regardless of the cause of the reduced blood leads, that as of mid-1979 the pediatric population is safely within all accepted health lead criteria guidelines, and therefore current levels based on extrapolation of the trend must be even lower.

CGP:cak
4/13/82

TABLE 1

FREQUENCY COUNTS* BY SAMPLING TIME AND REGION

6 Month Time Periods	Region: Northeast		South		Midwest		West		Total
76-1st half	0	(0)	463	(43)	43	(4)	560	(53)	1,066
76-2nd half	0	(0)	446	(29)	870	(56)	227	(15)	1,543
77-1st half	470	(40)	704	(60)	0	(0)	0	(0)	1,174
77-2nd half	845	(76)	267	(24)	0	(0)	0	(0)	1,112
78-1st half	0	(0)	0	(0)	19	(2)	1,193	(98)	1,212
78-2nd half	424	(35)	0	(0)	776	(65)	0	(0)	1,200
79-1st half	0	(0)	240	(17)	673	(47)	517	(36)	1,430
79-2nd half	353	(36)	330	(33)	308	(31)	0	(0)	991
80-1st half	0	(0)	315	(100)	0	(0)	0	(0)	315
Total	2,092		2,765		2,689		2,497		10,043

TABLE 2

SAMPLE REPRESENTATION* (MILLIONS) BY SAMPLING TIME AND REGION

6 Month Time Periods	Region: Northeast		South		Midwest		West		Total
76-1st half	.0	(0)	9.88	(48)	.38	(2)	10.27	(50)	20.53
76-2nd half	.0	(0)	5.92	(23)	16.28	(63)	3.76	(14)	25.96
77-1st half	10.11	(46)	11.81	(54)	.0	(0)	.0	(0)	21.92
77-2nd half	15.64	(75)	5.09	(25)	.0	(0)	.0	(0)	20.73
78-1st half	.0	(0)	.0	(0)	.30	(1)	24.63	(99)	24.93
78-2nd half	9.55	(43)	.0	(0)	12.72	(57)	.0	(0)	22.27
79-1st half	.0	(0)	2.70	(12)	10.20	(46)	9.17	(42)	22.07
79-2nd half	5.52	(33)	5.45	(32)	5.80	(35)	.0	(0)	16.77
80-1st half	.0	(0)	7.15	(100)	.0	(0)	.0	(0)	7.15
Total	40.82		48.00		45.68		47.83		182.33

TABLE 3

STATISTICS ON RISK TO BLOOD LEAD POISONING
6 MONTHS - 6 YEARS GROUP

	Population No. $\geq$ 30 µg/dl	Population Percent $\geq$ 30 µg/dl	Geometric Mean (µg/dl)	Population Represented	
79-1st half	15,347	.91	12.4	1,677,626	380
79-2nd half	10,450	1.36	11.6	769,585	201
80-1st half	2,563	.65	8.6	392,907	62
ALL	28,360	1.00		2,840,118	

*Parenthesized entries reflect percent of row total.